ATAS DA 15ª CONFERÊNCIA

Associação Portuguesa de Sistemas de Informação

Lisboa, 2 e 3 outubro de 2015

ISCTE Instituto Universitário de Lisboa

TÍTULO

Atas da 15ª Conferência da Associação Portuguesa de Sistemas de Informação

SUB-TÍTULO

Os sistemas de Informação da Gestão de Conhecimento, na Educação e na Formação

EDIÇÃO

ISCTE Instituto Universitário de Lisboa

COORDENAÇÃO

Henrique O'Neill, Carlos J. Costa, Luisa Domingues, Bráulio Alturas, Carlos Serrão

setembro 2015

ISSN: 2183489X

COMISSÃO DE CONFERÊNCIA

Isabel Ramos, Presidente da Comissão da Conferência, Centro Algoritmi/Universidade do Minho e Associação Portuguesa de Sistemas de Informação

Henrique O'Neill, ISCTE-Instituto Universitário de Lisboa

João Samartinho, Escola Superior de Gestão e Tecnologia/Instituto Politécnico de Santarém

Luís Silva Rodrigues, Instituto Superior de Contabilidade e Administração do Porto/Instituto Politécnico do Porto

COMISSÃO DE PROGRAMA

Carlos Costa, Presidente da Comissão do Programa, Instituto Universitário de Lisboa (ISCTE-IUL)

Portugal Abílio Oliveira, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Agostinho Sousa Pinto, Instituto Politécnico do Porto (IPP), Portugal

Alberto Carneiro, U. Europeia, Portugal

Alberto Silva, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Álvaro Rocha, University of Coimbra, Portugal

Ana Mendes, Departamento de Sistemas de Informação, Escola Superior de Ciências Empresariais, Instituto Politécnico de Setúbal, Portugal

Ana Moreira, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Ana Paula Afonso, Faculdade de Ciências, Universidade de Lisboa, Portugal

Anabela Mesquita, Instituto Superior de Contabilidade e Administração, Instituto Politécnico do Porto, Portugal

Antonio Guerreiro, Escola de Ciências Sociais, Universidade de Évora, Portugal

António Lucas Soares, Faculdade de Engenharia, Universidade do Porto, Portugal

António Palma-Dos-Reis, ISEG - Universidade de Lisboa, Portugal

António Serrano, Escola de Ciências Sociais, Universidade de Évora, Portugal

Antonio Carlos Maçada, Universidade Federal do Rio Grande do Sul, Brasil

Artur Silva, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Artur Marques, Escola Superior de Gestão e Tecnologia, Instituto Politécnico de Santarém, Portugal

Bráulio Alturas, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Carla Pereira, Escola Superior de Tecnologia e Gestão de Felgueiras, Instituto Politécnico do Porto e INESC TEC, Portugal

Carlos Serrão, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Carlos Coutinho, Caixa Mágica Software, Portugal

Carlos Morais, Instituto Politécnico de Bragança, Portugal

Carlos Ferreira, Universidade de Aveiro, Portugal

Carlos B. Sousa Pinto, Universidade do Minho, Portugal

Carlos Zorrinho, Escola de Ciências Sociais, Universidade de Évora, Portugal

Célio Gonçalo Marques, Escola Superior de Gestão, Instituto Politécnico de Tomar, Portugal

Cristiane Pedron, Universidade Nove de Julho - UNINOVE, São Paulo, Brasil

Delfina Soares, Departamento de Sistemas de Informação, Universidade do Minho, Portugal

Elsa Cardoso, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Esther Lage, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Eurico Lopes, Instituto Politécnico Castelo Branco, Portugal

Fernando Brito E Abreu, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Fernando Bacao, NOVAIMS, Universidade Nova de Lisboa, Portugal

Filipe Santos, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Filipe Madeira, Escola Superior Gestão Tecnologia Santarém, Politécnico de Santarém, Portugal

Filomena Castro Lopes, Universidade Portucalense, Portugal

Francisco Antunes, Universidade da Beira Interior e INESC Coimbra, Portugal

Francisco José García Peñalvo, Universidad de Salamanca, Espanha

Gabriel David, Faculdade de Engenharia, Universidade do Porto / INESC-Porto, Portugal

Gonçalo Dias, Universidade de Aveiro, Portugal

Guilhermina Lobato Miranda, Universidade de Lisboa, Portugal

Henrique O'Neill, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Henrique Gil, ESE, Instituto Politécnico de Castelo Branco/ CAPP-ISCSP, Universidade de Lisboa, Portugal

Henrique Santos, Universidade do Minho, Portugal

Henrique S. Mamede, Universidade Aberta,Portugal

Isabel Ramos, Universidade do Minho, Portugal

Isabel Alexandre, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Ivo Sousa, Universidade Aberta, Portugal

Joao Guerreiro, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

João Samartinho, Escola Superior de Gestão e Tecnologia, Instituto Politécnico de Santarém, Portugal

João Varajão, Universidade de Tás-os-Montes e Alto Douro, Portugal

João Costa, Faculdade de Economia, Universidade de Coimbra, Portugal

Joao Alvaro Carvalho, Escola de Engenharia, Universidade do Minho, Portugal

Jorge Oliveira E Sá, Departamento de Sistemas de Informação, Universidade do Minho, Portugal

Jorge Loureiro, Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Viseu, Portugal

Jorge Coelho, SisConsult / Universidade do Minho, Portugal

Jose Pina Miranda, MULTICERT - Serviços de Certificação Electrónica, S.A., Portugal

Jose Tribolet, Instituto Superior Tecnico, Instituto Superior Técnico, Universidade de Lisboa, Portugal

José Dias Coelho, Faculdade de Economia, Universidade Nova de Lisboa, Portugal

Jose Carlos Nascimento, Departamento de Sistemas de Informação, Universidade do Minho, Portugal

Leonel Morgado, INESC TEC / Universidade Aberta, Portugal

Leonilde Reis, Escola Superior de Ciências Empresariais, Instituto Politécnico de Setúbal, Portugal

Luis Amaral, Departamento de Sistemas de Informação, Universidade do Minho, Portugal

Luis Silva Rodrigues, Instituto Superior de Contabilidade e Administração do Porto, Instituto Politécnico do Porto, Portugal

Luisa Domingues, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Luísa Miranda, Instituto Politécnico de Bragança, Portugal

Manuel Santos, Departamento de Sistemas de Informação, Universidade do Minho, Portugal

Marco Painho, NOVAIMS, Universidade Nova de Lisboa, Portugal

Maria Bernardo, Universidade Aberta, Portugal

Maria Helena Monteiro, ISCSP, Universidade de Lisboa, Portugal

Maria Manuela Cruz Cunha, Instituto Politécnico do Cávado e Ave, Portugal

Maribel Yasmina Santos, Departamento de Sistemas de Informação, Universidade do Minho, Portugal

Mário Caldeira, ISEG, Universidade de Lisboa, Portugal

Mário Romão, ISEG, Universidade de Lisboa, Portugal

Mário J. Silva, INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Miguel Mira Da Silva, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Miguel Baptista Nunes, The University of Sheffield, Department. of Information, Western Bank, Sheffield, S10 2TN, UK, Reino Unido

Mirian Oliveira, Pontifícia Universidade Católica do Rio Grande do Sul, Brasil

Nuno Melão, Instituto Politécnico de Viseu, Portugal

Nuno Guimarães, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

Orlando Belo, Escola de Engenharia, Universidade do Minho, Portugal

Patrícia Leite Brandão, Escola Superior de Tecnologia, Instituto Politecnico Cavado e Ave, Portugal

Paula Serdeira Azevedo, ESGHT, Universidade do Algarve, Portugal

Paulo Quaresma, Escola de Ciências e Tecnologia, Universidade de Evora, Portugal

Paulo Resende Da Silva, Escola de Ciências Sociais, Universidade de Évora, Portugal

Paulo Rupino Da Cunha, Departamento de Engenharia Informática, Universidade de Coimbra, Portugal

Paulo Tomé, Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Viseu, Portugal

Pedro Rangel Henriques, Universidade do Minho, Portugal

Pedro Isaias, Universidade Aberta, Portugal

Pedro Fernandes, Universidade de Coimbra, Portugal

Ramiro Gonçalves, Universidade de Trás-os-Montes e Alto Douro, Portugal

Raul Vidal, Faculdade de Engenharia, Universidade do Porto, Portugal

Rodrigo Baroni, Pontíficia Universidade Católica de Minas Gerais, Portugal

Rui Gomes, Instituto Politécnico de Viana do Castelo, Portugal

Rui Dinis de Sousa, DSI/UM - Dep. Sistemas de Informação/Universidade do Minho, Portugal

Rui Rijo, IPL - Instituto Politécnico de Leiria, Portugal

Rui Quaresma, Escola de Ciências Sociais, Universidade de Évora, Portugal

Rui Pedro Lourenço, INESC Coimbra / Faculdade de Economia, Universidade de Coimbra, Portugal

Salvador Abreu, Escola de Ciências e Tecnologia, Universidade de Évora, Portugal

Tiago Oliveira, NOVAIMS, Universidade Nova de Lisboa, Portugal

Vitor Santos, NOVAIMS, Universidade Nova de Lisboa, Portugal

Winnie Ng Picoto, Instituto Superior de Economia e Gestão, Universidade de Lisboa, Portugal

COMISSÃO ORGANIZADORA

Henrique O'Neill, Presidente da Comissão Organizadora, Instituto Universitário de Lisboa (ISCTE-IUL)

Bráulio Alturas, Instituto Universitário de Lisboa (ISCTE-IUL)

Carlos Serrão, Instituto Universitário de Lisboa (ISCTE-IUL)

António Palma dos Reis, Instituto Superior de Economia e Gestão/Universidade de Lisboa

Marco Painho, Nova IMS Information Management School /Universidade Nova de Lisboa

Mário J. Gaspar Silva, Instituto Superior Técnico/Universidade de Lisboa

PARCEIROS, APOIOS E PATROCINADORES

Instituto Universitário de Lisboa (ISCTE-IUL)

Alumni Clube ISCTE-IUL

ISCTE ACM Student Chapter

ISCTE-IUL IEEE Students Branch

Portuguese Chapter of the Association for Information Systems

Fundação para a Ciência e Tecnologia

Noesis

Lidel FCA

PREFÁCIO

A 15a Conferência da Associação Portuguesa de Sistemas de Informação regressou a Lisboa em 2015 para a sua 15ª. edição. A CAPSI 2015 pretende continuar a afirmar-se como um espaço de partilha de experiências entre os membros da comunidade de sistemas de informação, provenientes das instituições do ensino superior, dos centros de investigação e das empresas em Portugal e nos países lusófonos.

Foi com grande satisfação que assumimos a responsabilidade de organizar esta edição, por motivos de natureza institucional e de natureza pessoal.

O ISCTE-IUL é uma instituição onde os sistemas de informação constituem uma área relevante em termos de ensino e investigação. Embora inicialmente associado fundamentalmente à Gestão e às Ciências Sociais, a temática dos sistemas de informação desenvolveu-se posteriormente associada às Tecnologias da Informação. É ainda de destacar que a licenciatura de Informática e Gestão de Empresas, cuja primeira edição se iniciou em 1989, correspondeu a um marco importante nesse desenvolvimento.

Em termos pessoais, satisfaz-nos contribuir para o aprofundamento desta área de conhecimento, consolidando o espírito de comunidade que tem vindo a ser construído no âmbito das CAPSIs.

"Os Sistemas de Informação na Criação de Conhecimento, na Educação e na Formação" é o tema central da CAPSI 2015, materializado fundamentalmente na intervenção de oradores convidados, reconhecendo a importância que os sistemas e as tecnologias da informação têm na disseminação do conhecimento e na formação das pessoas, nas instituições de ensino e de investigação, bem como nas empresas. Nesse contexto, a CAPSI 2015 estimula a apresentação de artigos de natureza académica e empresarial, em diversas áreas temáticas e domínios de aplicação.

Este ano, o processo de avaliação dos artigos resultou da conjugação de esforços de uma vasta comissão de programa que se empenhou em criticar, analisar, dar sugestões, selecionar os artigos submetidos. Os artigos científicos selecionados foram objeto de apresentação oral ou de apresentação com poster.

Cada um dos artigos enviados para a Conferência foi sujeito a um processo de revisão e avaliação por dois membros de comissão de programa, de forma anónima. Tendo sido recebidos 62 artigos, foram aprovados 42, o que dá uma taxa de aceitação de 68%.

Os artigos envolvem um leque alargado de temáticas, incluindo Teoria, Modelos, Governance, Gestão de Risco, Ensino, Processos: da modelação à auditoria, Business Intelligence, Gestão de Conhecimento, Big Data e Outsourcing.

- Para além dos artigos científicos, a conferência CAPSI 2015 tem um Doctoral Consortium com os seguintes objetivos:
- Encorajar os estudantes a elaborar um artigo que descreva o seu projeto de doutoramento e a apresentá-lo perante os membros da Comissão Científica.
- Oferecer a cada estudante comentários e novas perspetivas sobre o seu trabalho, exteriores à sua instituição.
- Reforçar a divulgação do trabalho através da exposição pública de um poster.
- Promover o desenvolvimento de uma comunidade de académicos com um espírito colaborativo de investigação.
- Contribuir para os objetivos da CAPSI através da interação com os outros investigadores, de uma forma integrada com os diversos eventos da conferência.

O Fórum de Profissionais de SI pretende ser um espaço de partilha de experiências e de reflexão sobre os desafios que se colocam a todos aqueles que exercem uma atividade profissional no domínio dos SI, nas empresas e em organismos da administração pública.

Neste livro de atas estão reunidos todos os artigos que constam no programa final da conferência.

Esta Conferência só foi possível porque a Comunidade dos Sistemas de Informação partilhou recursos, conhecimentos e contactos. Neste sentido, temos de agradecer a determinante colaboração e apoio dos colegas pertencentes à comissão do programa, que colaboraram na revisão dos artigos científicos, artigos do Doctoral Consortium e artigos profissionais.

Queremos destacar ainda outras pessoas e entidades que foram determinantes, e que, com o seu tempo, dedicação e esforço, tornaram possível a organização deste evento, nomeadamente:

- Os serviços do ISCTE-IUL e da do ISTAR-IUL que ajudaram operacionalmente para o sucesso da Conferência e acolheram a CAPSI 2015 nas suas instalações.
- As empresas e instituições que patrocinaram financeiramente a conferência, designadamente a NOESIS, a a FCT - Fundação para a Ciência e Tecnologia e a FCA - Editora de Informática.
- A APSI que assegurou o apoio administrativo, bem como as instituições universitárias parceiras: o IST, o ISEG e a NOVA IMS.
- Todos os que enviaram artigos e, em especial, todos os participantes na Conferência.

A todos vós gostaríamos de manifestar o nosso reconhecimento, na expectativa de que a CAPSI 2015 contribua para abrir novos horizontes e perspetivas quanto ao modo como poderemos continuar a intervir na geração de conhecimento, na educação e na formação através dos sistemas de informação.

Henrique O'Neill, Carlos J. Costa, Luisa Domingues, Bráulio Alturas, Carlos Serrão

ÍNDICE DE ARTIGOS

(APRESENTAÇÃO ORAL)

A Public Organization Career Progression Analysis Using Process Mining 25
Nuno Silva, Marcelo Silva, Carlos Mendes, Miguel Mira Da Silva

Adoção de Tecnologias da Informação: Uma proposta de Modelo de Análise Explorando o modelo UTAUT para Ambientes Virtuais de Aprendizagem para Suporte ao Ensino Presencial 47
Ivo Pedro Gonzalez Junior, Ernani Santos

Análise da relação entre processos de gestão do conhecimento e de gestão de projetos. 67
Fernanda Gomes, Mírian Oliveira

Analysis of Enterprise Architecture Models with Description Logics Reasoning and SPARQL 91
Gonçalo Antunes, Artur Caetano, Jose Borbinha

As Tecnologias da Informação no processo de ensino-aprendizagem da música 109
Luís Costa, Bráulio Alturas, Abílio Oliveira

Auditoria do Alinhamento entre Sistemas de Informação e Negócio 119
Gilberto Capeca, Filipe De Sá-Soares

Autopoiesis and organizations: a biological perspective 139
Ana S. Costa, José Tribolet

Categorização do Consumo de Energia em Sistemas de Povoamento de Data Warehouses 151
Miguel Guimarães, João Saraiva, Orlando Belo

Column-Based Databases: Estudo Comparativo no âmbito das Bases de Dados NoSQL 163
José Pedro Cunha, José Luís Pereira

Cruzamentos e Encruzilhadas – a importância de um SI para um serviço de psicologia escolar 183
Pedro Miguel de Oliveira, Abílio Oliveira, António Martins

Document-Based Databases: Um estudo exploratório no âmbito das bases de dados NoSql 191
Gonçalo Sousa, José Pereira

Ferramentas Gratuitas para Desenvolvimento de Soluções de Business Intelligence 209
Tiago Silva, Jorge Loureiro

Information management in supply networks: a framework to support strategic decision-making 233
Dario Messina, Ana Barros, António Lucas Soares

Information systems design and user satisfaction: A banking case study 241
Bruno Padinha, Henrique O'Neill

Medição de usabilidade e eficiência de dashboards: framework desenvolvida usando os princípios de desenho de Business Intelligence e Interação Homem-Máquina 263
Miguel Oliveira, Elsa Cardoso, Marina Santana

Metodologia para Análise de Oportunidades de Melhoria do Sistema de Gestão de Transporte de Doentes 287
Constança Roquette, Mónica Duarte Oliveira, José Tribolet, Lúcia Costa, Paulo Marques

Mobile ERP: Melhoria do processo de inventário em loja 305
Alexandra Monteiro, Carlos J. Costa

Modelação de Processos de Gestão de Recursos Humanos 319
Sebastião Pedro, Vitor Santos

Modelação de Processos de Negócio: Análise Comparativa de Linguagens 327
Diogo Silva, José Luis Pereira

Modelo de Avaliação do Desempenho das Concessões Automóveis 347
Jorge Pires, Bráulio Alturas

Modelos de Simulação de Processos por Junção de Componentes 369
Manuel Almeida, Jose Luis Pereira

O impacto da utilização de RFID: Um estudo de caso numa empresa de manutenção e engenharia 385
David Ribeiro, Winnie Ng Picoto, Paulo Gonçalves

O Impacto do Sistema de Workflow no Licenciamento de Redes e Serviços de Radiocomunicações – Estudo de caso na ANACOM 401
Manuel Gomes, Luisa Domingues

O Twitter como agente facilitador de recolha e interpretação de sentimentos 421
Carina Andrade, Maribel Santos

Os perfis de infraestrutura e a partilha do conhecimento 443
Mírian Oliveira, Antonio Maçada, Carla Curado, Felipe Nodari

Outsourcing de Serviços de Sistemas de Informação na Banca em Portugal 455
Cidália Pereira, João Varajão, Luis Amaral, Delfina Soares, António Cunha

Pontos de verificação para contratação de serviços de Outsourcing em Segurança da Informação de TI 481
Teofilo Branco Junior

Portuguese Public Procurement Governance Model 495
Isabel da Rosa, José Tribolet, Miguel Mira da Silva

Sistema de gestão e auditoria fiscal na nuvem 521
Luís Rolo, Alexandre Fonte, Eurico Lopes

Sistema Inteligente de Recolha, Armazenamento e Visualização de Informação proveniente do Twitter 535
Gaspar Brogueira, Fernando Batista, Joao Paulo Carvalho

Um Método para Identificar a Não Conformidade entre os Modelos de Processos de Negócio e o Conceitual de Informação 557
Rafael Duarte, Denis Silveira, Fernando Wanderley

Uma reflexão sobre a evolução das funcionalidades de tecnologias e sistemas de informação na perspectiva da evolução natural 575
Jairo Dornelas, Joao Alvaro Carvalho , Denis Silveira

Value Focused Approach to Information Systems Risk Management 595
Sergio Nunes, Gurpreet Dhillon, Mario Caldeira

ÍNDICE DE ARTIGOS

(APRESENTAÇÃO EM POSTER)

Alinhamento do negócio com os sistemas de informação: Médias e pequenas empresas portuguesas do setor Vitivinícola 613
Paulo António Salvador, Ramiro Gonçalves, Vítor Santos

Contributo para a Implementação de uma ferramenta de service desk no Instituto da Mobilidade e dos Transportes, I.P. 629
Carlos Carvalho, Henrique O'Neill

Does IS/IT Investments Impact on Firm Financial Performance: "Stakeholder" versus "Shareholder" Orientation 643
António Guerreiro

Estruturas Organizacionais para Gestão do Valor das TI: Um Modelo Conceptual para as Universidades Públicas em Portugal 661
Cristiano Pereira, Carlos Ferreira, Luis Amaral

GeoMarketing: Expansion and Development 679
Giovanna Monteiro , Ana Mafalda Barata , Carlos J. Costa

Gestão do Conhecimento nas Organizações 691
Maria da Conceição de Mello Sampayo de Freitas, José Tribolet

On the evolution of hyperlinking 703
Artur Marques

Tecnologias Informáticas em Organizações de Saúde: o que ainda deve ser repensado 715
Luís Costa, Alberto Carneiro

ÍNDICE DO *DOCTORAL CONSORTIUM*

Arquiteturas Informacionais para a Organização da Informação em Redes Colaborativas Empresariais 729
Solange Francisca Mazzaroto, António Manuel Lucas Soares

Information, knowledge and collaborative networks in internationalisation processes of SMEs 737
Eric Costa, António Lucas Soares, Jorge Pinho de Sousa

Information management in supply networks: a framework to support risk management decisions 745
Dario Messina, Ana Cristina Barros, António Lucas Soares

Governança de TI em universidades públicas: Proposta de um modelo 753
Isaías Bianchi, Rui Dinis Sousa

The emotion-aware city: Development of a framework based on social media towards emotion & stress mapping 761
Tiago H. Moreira de Oliveira, Marco Painho

Um modelo de confiança para a gestão dos Sistemas de Informação no ambiente de Computação em Nuvem 771
Teófilo Branco Júnior, Henrique Manuel Dinis dos Santos

ÍNDICE DOS ARTIGOS PROFISSIONAIS

9 Tips on How to Pitch a Project - The Collectors Bridge case study 781
João Diogo Ramos

Angariação e retenção de clientes numa maternidade privada 787
Maria Rodrigues, Guilherme Victorino, Roberto Henriques

Qual o Impacto das aplicações em "Software as a Service" (SaaS) no negócio das PME ? 793
Jorge Carrola Rodrigues

Technical Writing for Portuguese Product Companies 799
Joaquim Baptista

ARTIGOS

(APRESENTAÇÃO ORAL)

A Public Organization Career Progression Analysis Using Process Mining

Nuno Silva[1], Marcelo Silva[2], Carlos Mendes[2], Miguel Mira da Silva[1]

1) Instituto Superior Técnico, Portugal

nuno.silva@ist.utl.pt, mms@ist.utl.pt

2) INOV INESC Inovação, Portugal
marcelo.silva@ist.utl.pt, carlos.mendes@ist.utl.pt

Abstract

Process discovery provides a set of methods and techniques to help understand the reality of operating processes in an organization. Process Mining is a well-recognized method by academia and the industry that focuses on workflow discovery. By analysing event data from running processes it is possible to discover workflow models that, despite being executed, were not completely known and understood in detail by the stakeholders. In this paper we describe a field study in which we applied the Process Mining approach to a different target than the usual business processes: career progression based on categories. The field study took place in a public entity with 585 workers. After a results analysis that included an interview with a stakeholder we found the advantages of applying Process Mining techniques to this context. Some of these findings were somewhat predictable, while others raised some interesting thoughts.

Keywords: Process Mining, Process Discovery, Workflow Management, Field Study, Career progression

1. Introduction

Nowadays the term workflow management is widely used throughout the world either in organizations as means of conducting the management of its business processes or in academia where new modelling techniques and capabilities concerning this area are proposed or existing ones improved. By making graphical process definitions, one can configure workflow management systems to support business processes [Van Der Aalst 2004].

Workflow mining, most commonly known as Process Mining, is to construct a workflow model (as means of a Petri net) given a workflow log with all the instances (cases), activities, timestamps, resources and other data that are part of a process. These event logs are provided by workflow management systems.

Process Mining is useful for finding out how people and/or procedures really work and for Delta analysis, i.e., comparing the actual process with some predefined process. Such a model specifies how people and organizations are assumed/expected to work. By comparing the descriptive or prescriptive process model with the discovered model, discrepancies between both can be detected and used to improve the process [Van Der Aalst 2004].

The Process Mining techniques were chosen in the context of the First Sight Model (FSM) project to extract the processes of the applications developed using Genio. Genio is the rapid development solution from Quidgest, the FSM project promoter. These processes correspond to the layer with highest level of abstraction of the framework that was developed in the FSM project. By using Process Mining we saved time in the process elicitation step, since it gave us the initial process models that were used as the basis of the workflow engine.

The contribution of this paper is not focused on the Process Mining algorithm used, but on the application of the most used algorithms in an unusual context: career paths. Therefore, we are not focused on finding new algorithms or proposing improvements to existing ones, instead we are analysing which are the results of applying existing Process Mining techniques to career path progress data. Accordingly, this paper objective is to answer to following question:

- **Is Process Mining suitable as a process discovery method to analyse career progress?**

In order to answer this question we applied Process Mining techniques to career progress data of a public institute. We retrieved results of the field study concerning career progress speed (average time between career categories), most current career paths, outlier career paths**,** number of employees that passed between categories, and between which categories the progression is faster.

The paper is structured as follows. We start by describing the related work of this research (Section 2). In Section 3, we describe the field study at a Portuguese public organization. In Section 4, we explain the evaluation process and discuss the results. Finally, we present our conclusions (Section 5).

2. Related Work

In this section we make an overview on the principle of Process Mining (also known as Workflow Mining) and the main methods and applications supporting this principle from which value can be gathered.

2.1. Process Mining

Workflow management offers a set of capabilities necessary for automating structured business processes. Workflow management systems are tools for supporting workflow management that can be configured through the use of models describing the life-cycle of a workflow instance. Besides pure workflow management systems, other systems have adopted

workflow technology such as ERP (Enterprise Resource Planning) systems like SAP, CRM (Customer Relationship Management) software, etc. [Van Der Aalst 2004].

Some issues typically arise when applying this type of workflow technology. One of the problems is that these systems require a workflow design. Modeling a workflow is not a trivial task. To design a proper model one requires deep knowledge of the workflow language and understanding of what is required through stakeholder discussion. Therefore, instead of starting by the workflow design, we start by gathering information about the workflow processes in runtime. For that one assumes that it is possible to record events such that [Van Der Aalst 2004]:

1. "each event refers to a task (i.e., a well-defined step in the workflow),"

2. "each event refers to a case (i.e., a workflow instance), and"

3. "events are totally ordered (i.e., in the log events are recorded sequentially, even though tasks may be executed in parallel)."

Transactional systems (e.g., ERP, CRM), or workflow management systems offer this information in some form. This information called workflow logs are then used to construct a process specification, which adequately models the behaviour registered. The term Process Mining is used for the method of discovering a structured process description from a set of real executions [Van Der Aalst 2004].

Process Mining is useful for finding out how people and/or procedures really work. Take for example processes supported by an ERP system like SAP (e.g., a procurement process). Such a system logs all transactions, but in many cases does not enforce a specific way of working. In such an environment, Process Mining could be used to gain insight in the actual process [Van Der Aalst 2004].

Furthermore, Process Mining is a useful tool for Delta analysis, i.e., comparing the actual process with some predefined process. In many situations, there is a descriptive or prescriptive process model. Such a model specifies how people and organizations are assumed or expected to work. By comparing the defined process model with the discovered model, discrepancies between both can be detected and used to improve the process [Van Der Aalst 2004].

There are many other benefits of using Process Mining related to process resources and other types of process data that can help increase process awareness, finding bottlenecks caused by resources, etc. This and the comparison between actual processes and predefined processes can be used to trigger Business Process Reengineering (BPR) efforts [Van Der Aalst 2004].

In the next subsections we discuss in more detail some of the methods and techniques of Process Mining.

2.2. Methods and techniques of Process Mining

Process Mining provides a wide range of methods and techniques from which valuable insight concerning processes can be achieved. In the next subsections we briefly describe some of the most popular techniques.

2.2.1. Conformance Checking

Conformance dictates, in the Process Mining context, whether or not there is a good match between the recorded events and the model [Rozinat 2008]. Consider Figure 1, where conformance checking is positioned in the broader context of Process Mining techniques. While discovery aims at the automatic extraction of a process model from log data, conformance checking is concerned with the comparison of an existing process model and a corresponding log.

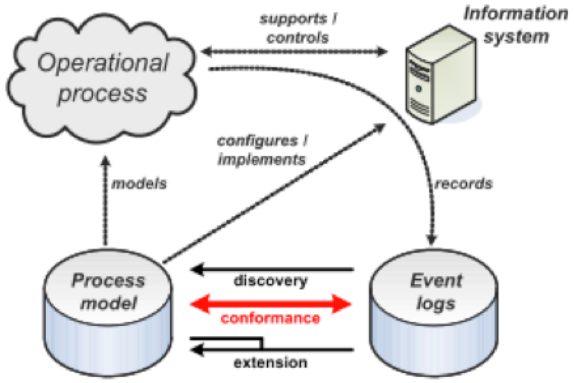


Figure 1 - Conformance checking in the boarder context of Process Mining (from [Rozinat 2008])

An organization typically documents its processes in some form. As stated in the previous section, a process model may be of a descriptive or of a prescriptive nature. Descriptive models try to capture existing processes without being normative. Prescriptive models describe the way that processes should be executed. Even if the process model and the event log are fully compliant, it is interesting to see how frequent specific parts in the model are actually used, and to potentially remove obsolete parts which otherwise would need to be maintained [Rozinat 2008].

The most dominant requirement for conformance is fitness. An event log and a Petri net (i.e., process model) “fit” if the Petri net can generate each trace in the log [Rozinat 2008]. Fitness is a quantifiable metric ranging from 0 to 1. Unfortunately, a good fitness does not imply

conformance. It is easy to construct Petri nets that are able to parse any event log. Although such Petri nets have a fitness of 1 they do not provide meaningful information as illustrated by Figure 2 [Rozinat 2008].

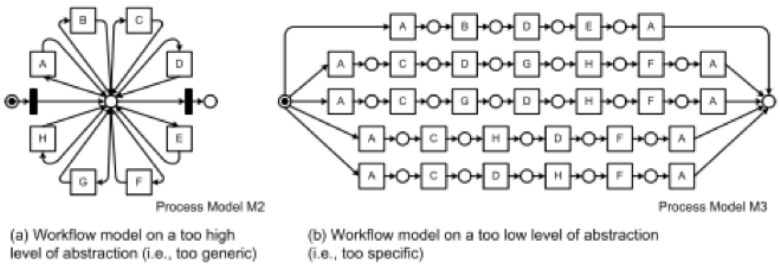


Figure 2 - Fitting models do not need to be a good representation of the observed behaviour (from [Rozinat 2008])

Therefore, a second dimension was introduced: appropriateness. Appropriateness tries to capture the idea of Occam's razor, i.e., "one should not increase, beyond what is necessary, the number of entities required to explain anything" [Rozinat 2008]. So, if one wants to analyse conformance these two dimensions must be considered. First, the fitness between the log and the model needs to be ensured and second, the appropriateness of the model can be analysed with respect to the log [Rozinat 2008]. Together they allow for the quantification of conformance.

2.2.2. Social Network Discovery

Process Mining uses sociometry methods [Wasserman 1994], [Scott 1994] e [Burt 1983] to discover the interpersonal relationships of the resources executing the process. The end result is a Social Network (SN) of all the process resources and their respective relations.

Many enterprise information systems store relevant events in a more structured form such as the enabling, start and completion of activities [Van Der Aalst 2005], [Leymann 2000], [Jablonski 1996] [Fischer 2001] e [Van Der Aalst 2002] . When people are involved in events, logs will typically contain information on the person executing or initiating the event.

Aalst [Van Der Aalst 2005] states that "An event may be denoted by (c, a, p) where c is the case, a is the activity, and p is the person. Events are ordered in time allowing the inference of causal relations between activities and the corresponding social interaction. For example, if (c, a1, p1) is directly followed by (c, a2, p2), there is some handover of work from p1 to p2 (note that both events refer to the same case). If this pattern occurs frequently but there is

never a handover of work from p1 to p3 although p2 and p3 have identical roles in the organization, then this may indicate that the relation between p1 and p2 is stronger than the relation between p1 and p3 [Van Der Aalst 2005]."

By considering the above one can build a social network expressed in terms of a graph as illustrated in Figure 3.

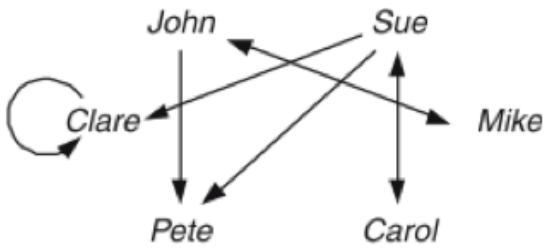


Figure 3 - Sociogram based on an event log (from [Van Der Aalst 2005])

To sum up, Social Network Analysis (SNA) extracts information from event logs and constructs a sociogram that can be used to analyse interpersonal relationships in an organization. These sociograms are based on the observed behaviour from the log and uses events like the delegation of work from one individual to another. Figure 4 shows a discovered SN using a SNA tool.

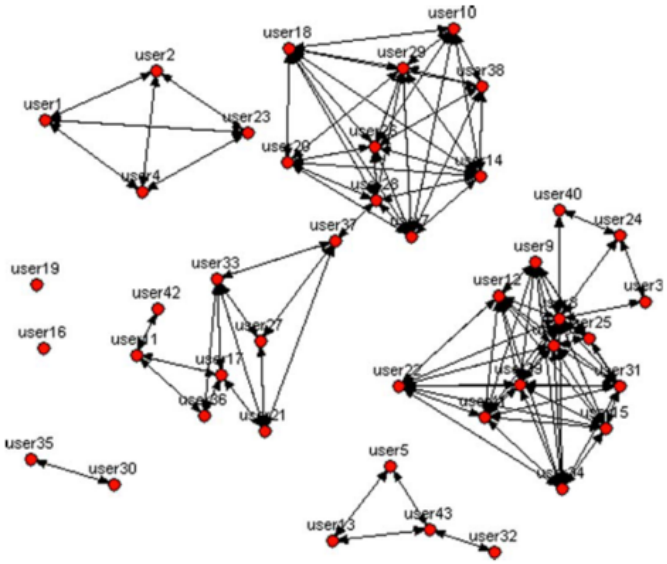


Figure 4 - A discovered Social Network (from [Van Der Aalst 2005])

2.2.4 Process Mining applications to career progress

We did not found direct applications of Process Mining techniques to career progression data. This absence does not undermines our motivation in taking this research. Indeed, the

application of Process Mining to unusual contexts is endorsed by the community and there are several researches in that case [Edgington 2010], [Nakatumba 2010] e [Van Der Aalst et al. 2010].

The most common analysis applied to career progression data is data mining, such as in [De Bruin 2006]. Data mining applied to these kind of contexts is suitable for finding correlations between career paths, however, by applying Process Mining techniques one can derive more clear career paths, as we show in this paper.

2.3. Summary

In this section we provided a brief related work on the subject of Process Mining. Process Mining has the goal of extracting information from transaction logs and with it discovering process models that reflect the reality of the operating process. Furthermore, besides this *play-in* approach, Process Mining uses the discovered process and the event log to pinpoint deviations regarding a normative process and from there, propose extensions to the running process or refinements to the normative one based on what is performed in reality (if those deviations prove to be beneficial to the outcome of the process). Finally, by combining Process Mining and SNA approaches, it is possible to discover the processes Social Network and perceive how the processes workload is being handled across all the process resources.

3. Field Study - Applying Process Mining techniques to career development data

The field study described in this paper took place in the context of the First Sight Model (FSM) project[1]. FSM project, Quidgest's individual QREN project number 30283, has the objective to build a framework based on multiple abstraction layers that facilitates the software development and the Information Systems evolution. Process Mining was used to identify the processes behind the applications developed with Genio (the rapid development solution from Quidgest).

Process Mining is typically applied to "process-based" cases, i.e, field studies where the software or reality observed are following well-defined processes. In those cases, the objective of the Process Mining analysis is to take conclusions about the performance and sequence of activities, always taking into account the standard processes defined previously.

[1] http://www.quidgest.pt/q_ProjetosQRENPT.asp?LT=PTG

In this paper, the objective is to apply Process Mining to unexplored field studies where this type of analysis brings added value to the organizations, providing conclusions that would be difficult to obtain without this.

Within Quidgest and Genio software applications, we found some cases where the application of Process Mining would be innovative, one of them, the Career Progression of a Portuguese Public Organization. This organization is accountable for promotion, enhancement and sustainability of Tourism activities and in 2014 had 585 employees.

The Career Progression was chosen because it is not a process "as-is", but it can be seen as a process if we consider that each employee is an instance/case of the process and each professional category is an activity. The date when an employee passed from a category to another is the timestamp.

3.1 Collect Data

As mentioned before, we used data from Quidgest, more specifically, from Genio software. After the decision of using Career Progression as field study to this paper, the next step was to collect the right data to apply the right Process Mining algorithms. Quidgest has many clients with his Human Resources software deployed. For this research we used the data of one client that uses the HR software since 35 years ago, i.e, has data from 35 years of careers. So, to collect the data about the careers of the employees of that organization, we met with the head of Human Resources from Quidgest because the experience of someone inside of business is important to collect and understand the right data. With him we extracted the data from Genio to an excel spreadsheet.

After exporting, data formatting was done, because data was not in the correct format. One of the problems was the accuracy of the names of categories, because there was not a coherence between the names of all categories. After that, we had to normalize the career progressions, since the rules of the progression have been changed over the years. Lastly, some errors in the data (loops in categories) were fixed in the final export excel spreadsheet. All these corrections were done with the help of Human Resources chief from Quidgest and these were found by importing the excel spreadsheet to Disco [Gunther 2012] and analyzing quickly the overview Map that Disco produced.

3.2 Apply Process Mining using Disco

The purpose of this paper is to apply Process Mining to unexplored field studies, as mentioned before. So, we focused on the quality of the data collected and on a deep analysis

to take useful conclusions to the Process Mining research. We decided to use a software which already has the algorithms implemented and is simple and intuitive to use - Disco [Gunther 2012].

Disco uses an intuitively understandable and 100% truthful process map visualization. The thickness of paths and coloring of activities show the main paths of the process flows, and wasteful rework loops are quickly discovered [Gunther 2012].

The Disco miner is based on the Fuzzy miner, but has been further developed in many ways. The Fuzzy Miner was the first mining algorithm to introduce the "map metaphor" to Process Mining, including advanced features like seamless process simplification and highlighting of frequent activities and paths [Gunther 2012].

Disco matched the expectations and with it, we have made all the analysis we needed.

3.3 Results

In this section we present the results obtained from the application of Process Mining in the career progression of a public organization. First we describe a set of statistics that provided initial insight of the reality of the organization and what could be expected from the discovered process map. Then, we analyse the discovered process map in terms of frequency, performance and the relationship between both. In the end, we summarize the results and state some conclusions.

3.3.1 Event Log Statistics

As a first approach, after importing the event log into the Disco tool, we analysed the respective log from a statistics perspective. The event log was composed of 714 events, a total of 51 different career categories, and 585 cases (people associated to the organization). The log timeline goes from 17-06-1980 up to 01-04-2015.

By analysing the number of events per case we came to the conclusion that most people (cases) only have registered a single change in category. To be more precise, 482 people only have one category registration, 86 people have two registrations, 12 people with three registrations, 2 people with 4 and 5 category registrations and finally, only 1 person had six registrations. Figure 5 Illustrates these findings.

Figure 5 - Events per case chart

We can conclude from these statistics that, although the records are from a period of 35 years, there are indeed few records of the career progression of the organization's employees, meaning that, concerning this particular organization, people do not change their category as frequently as it would be expected. Probably in the private sector we could observe a completely different scenario since private organizations tend to be more dynamic in terms of career progression than in the public sector. Would be interesting to compare both perspectives in future work. Furthermore, by analysing the active cases over time, we observed that most active cases (employees) are distributed in the last period of the log timeline. Figure 6 illustrates this.

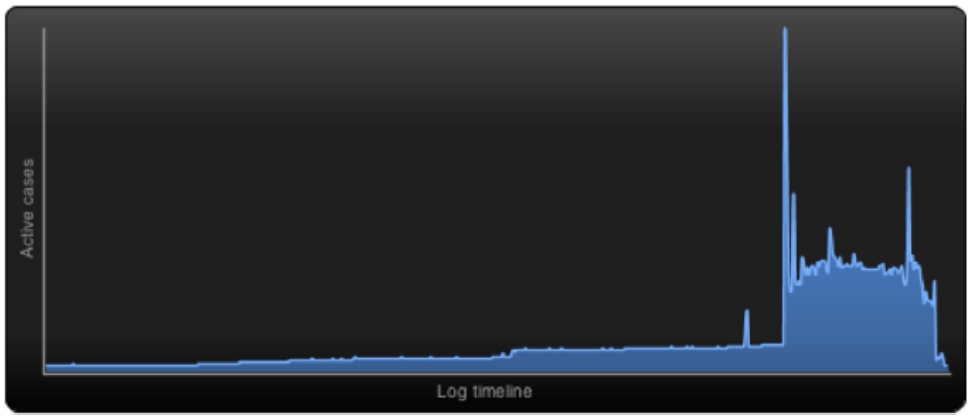


Figure 6 - Active cases over time chart

This might have a direct relation with the production year of the career progression system. That would explain the drastic change in the chart regarding the number of active cases after the year of 2008. Next we analyse the discovered process map.

3.3.2. Discovered Process Map

In order to obtain an understandable and well-defined process map that clearly showed the category progressions of the organization, a detail filter was applied to the initial process map shown in Figure 7.

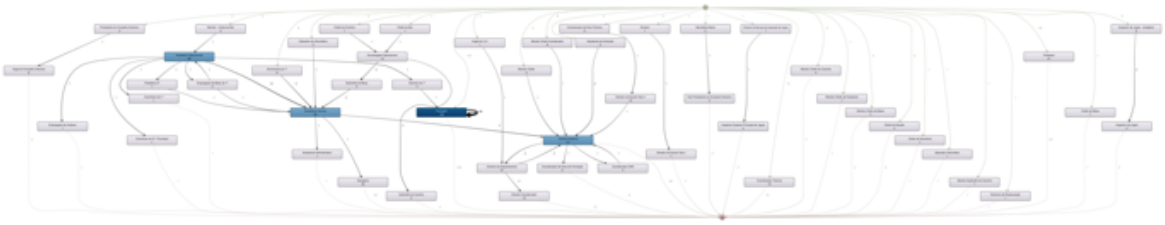


Figure 7 - Discover Process Map (Overview)

As it is clearly observed in the figure above the discovered process in its "as-is" state is hard to understand from a workflow perspective. Therefore, we applied a detail filter that would filter the discovered process map by its most frequent activities (categories) and most frequent paths between those activities. We used a percentage of 20% for both activities and paths resulting in the process map illustrated in Figure 8.

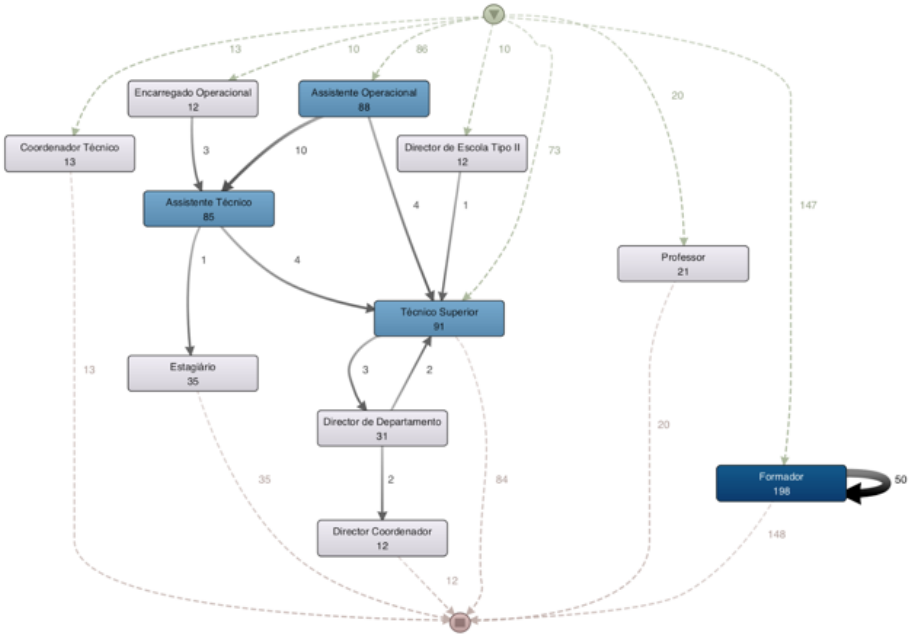


Figure 8 - Discovered Process Map (Filtered)

The process is mainly composed of 11 different categories and 10 paths connecting those categories. Some categories do not possess any kind of relation with the remaining categories, which is why they are only connected to the source and sink nodes (e.g. *Coordenador Técnico* – Technical Coordinator, *Professor* - Professor, etc.). The more bluish and darker nodes correspond to the categories with a higher number of registrations. Each number below the name of the category is the number of people registered in that category. The number next to each path represents the number of people that progressed from a category A to a category B. In this scope of frequency we are only considering the absolute frequency. The categories with the highest number of employees, in descendent order, are as follows:

- *Formador* - Trainer (198 employees)
- *Técnico Superior* - Senior Technician (91 employees)
- *Assistente Operacional* - Operational Assistant (88 employees)
- *Assistente Técnico* - Technician Assistant (85 employees)

Concerning performance, we took into consideration the mean duration of each category progression, i.e., the average of all path durations between every two categories. We observed that there were two main paths that would stand out from the remaining ones as Figure 9 illustrates, those being:

- *Assistente Técnico -> Técnico Superior* (10,4 years)
- *Técnico Superior -> Diretor de Departamento* – Department Director (114,6 months ~ 4.5 years)

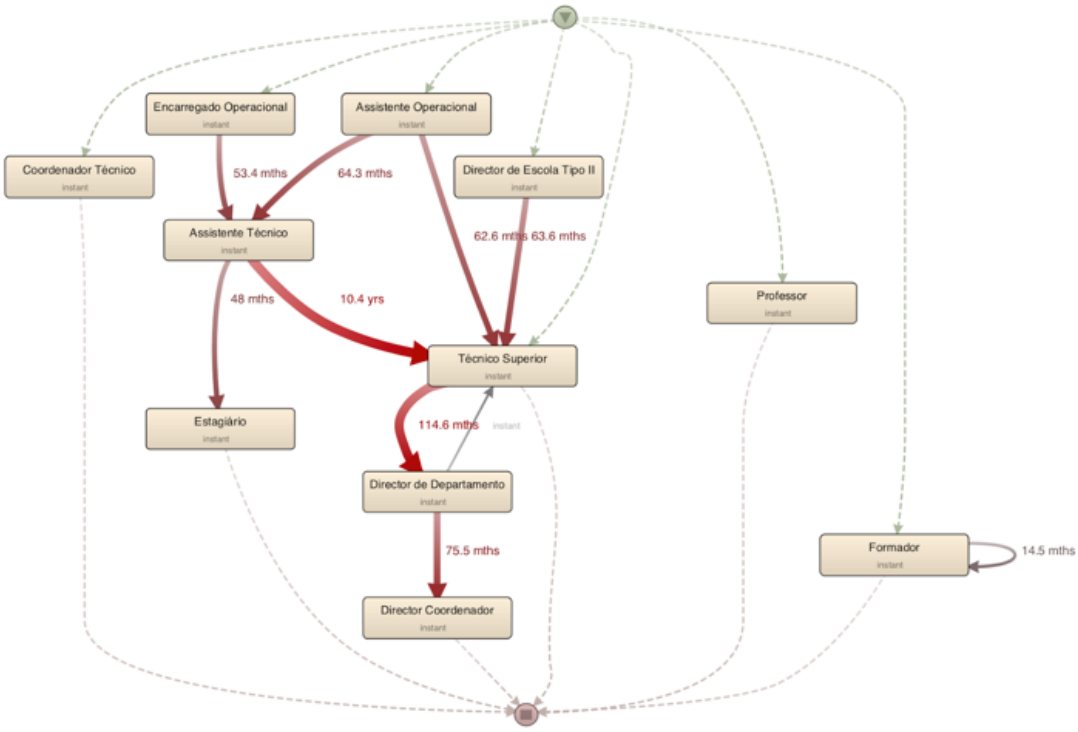


Figure 9 - Discovered Process Model (Performance View)

By considering the mean duration between the category progressions of an employee's career based on the process map shown above, one can verify category progression duration of approximately 5 years between the first set of categories. The duration of these progressions becomes considerably higher in the medium set of categories now being of approximately 10 years (two times higher than the initial set of categories). Although, there are exceptions such as the progression from *Assistente Técnico* to *Estagiário* – Intern with an estimated duration of 4 years. At the end phase of the progression process, the progression duration becomes again smaller as observed in Figure 9 from *Director de Departamento* to *Director Coordenador* – Coordinator Director (approximately 6 years). Being a public organization and considering that most of the records were made in the last few years, it is somewhat expected that the mean duration of progression between mainly the middle categories is considerably high, mostly due to the economical crisis of the country.

We also performed a velocity analysis to one career to see if the career progression of one employee follows a regular velocity when passes through all categories. The career analyzed

was ***Assistente Operacional -> Assistente Técnico -> Técnico Superior -> Director de Departamento -> Director Coordenador***. The conclusion is illustrated in Figure 10 and as we can see the velocity is not constant. In the beginning of the career passing from *Assistente Operacional* to *Assistente Técnico* took 5.4 years, then the velocity decreased passing to *Técnico Superior*, which took almost twice, 10.4 years. From *Técnico Superior* to *Director de Departamento* the velocity remained almost constant, decreased 0.8 years, to 9.6 years. Finally the last passage from *Director de Departamento* to *Director Coordenador* again increased the velocity to 6.3 years.

We can conclude that career progressions start with a high velocity, then the velocity decreases in the middle categories of the career and in the final phase of the career, the velocity increases again until the end.

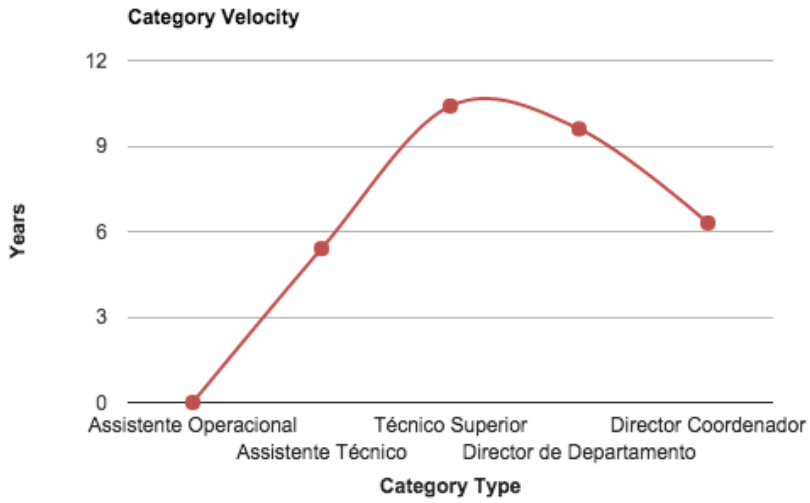


Figure 10 - Velocity Graph of Career Progression

In the end we tried to correlate both Absolute Frequency and Mean Duration to understand if they had some sort of relation with each other. We noticed that in most paths, regardless of the mean duration being high or low, the path frequency was low. This has to do with the few case variants that were obtained from the long, i.e. the diversity of cases was low. Furthermore, we only collected 585 cases (employee registrations) and from those 585 cases, only 359 actually relate to the discovered process model. Not only that, the fact that this log is from a public organization, it is known that the progression hierarchy is relatively slow and bureaucratic, which is why the 35 years of collected data are possibly not enough to clearly understand a complete career progression of a public employee in this specific context.

Despite all the above, one can see that the paths *Assistente Operacional -> Assistente Técnico* and *Formador -> Formador* are the ones that happen more frequently and with less duration (10 times – 64,3 months and 50 times – 14,5 months respectively). Figure 11 shows the relation between the frequency of the categories and the mean duration of each category progression

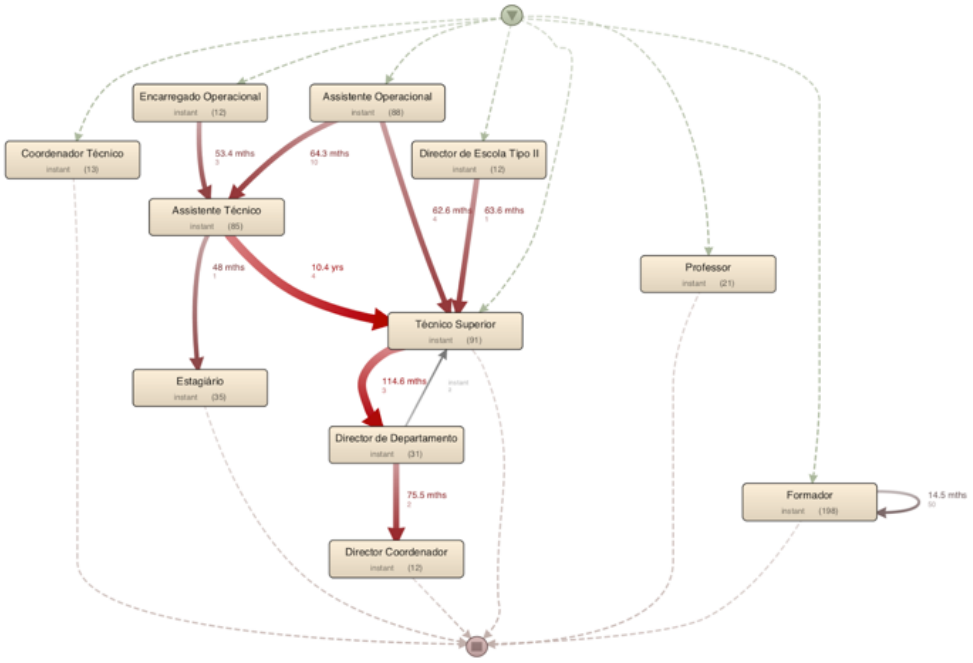


Figure 11 - Discovered Process Map (Relationship Between Frequency and Performance Views)

To assess whether the discovered career process represents the reality in an accurate and trustworthy manner, the quality of the event log data must be taken into account. For that reason, we used the *standard deviation* statistical metric to analyse the data quality of the log. This is a proper metric to quantify the amount of dispersion of a set of data values which is precisely what we want to assess. We considered the career progression ***Assistente Operacional -> Assistente Técnico -> Técnico Superior -> Director de Departamento -> Director Coordenador*** to assess data quality since it represents the career path with most flow of data according to the discovered process frequency-wise and performance-wise. We sampled this workflow into four samples representing each of the four category transitions that compose that same workflow.

- *Assistente Operacional -> Assistente Técnico*
- *Assistente Técnico -> Técnico Superior*

- *Técnico Superior -> Director de Departamento*
- *Director de Departamento -> Director Coordenador*

For each sample we considered all cases concerning them. After computing the standard deviation for each sample we calculated the average deviation in order to have a single value form where conclusions concerning the log quality could be made. The approach we used for computing the standard deviation was to calculate the total transition years of each sample's instance. The Tables 1, 2, 3 and 4, below represent the total years of each transition sample for each of its instances and its respective standard deviation.

Assistente Operacional -> Assistente Técnico										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Total Years	15.68	5.00	4.42	4.75	0.00	4.92	4.92	4.59	4.59	4.75
Std. Deviation	3.92									

Table 1 – Total Years and Standard Deviation to transition: *Assistente Operacional -> Assistente Técnico*

Assistente Técnico -> Técnico Superior				
	1	**2**	**3**	**4**
Total Years	30.39	4.59	4.34	2.2
Std. Deviation	13.39			

Table 2 – Total Years and Standard Deviation to transition: *Assistente Técnico -> Técnico Superior*

Técnico Superior -> Director de Departamento			
	1	**2**	**3**
Total Years	9.09	19.55	0.0028
Std. Deviation	9.78		

Table 3 – Total Years and Standard Deviation to transition: *Técnico Superior -> Director de Departamento*

Director de Departamento -> Director Coordenador		
	1	2
Total Years	5.24	7.34
Std. Deviation	1.49	

Table 4 – Total Years and Standard Deviation to transition: *Director e Departamento -> Director Coordenador*

Average Std. Deviation: 7.14

The average standard deviation value of 7.14 years (being the highest deviation of 13.39 years and lowest deviation of 1.49) reveals a considerable dispersion of data according to the mean of the data set. Knowing that a standard deviation of approximately 0 would classify the data as being accurate, this value shows that the given data set is somewhat inaccurate and must not be completely reliable. A reason for this (discussed below) has to do with the incorrect insertion of data in the IS responsible for managing the career progression data. Next we state the main conclusions of our analysis.

4. Discussion

In the previous section we covered an analysis of the Process Mining approach applied in a public organization. The goal was to analyse a set of factors concerning career progression. We analysed the discovered process map from two scopes: absolute frequency and performance.

In the frequency scope we could observe the number of employees related to each category and how many progressed from a specific category to another. In the performance view, we used as main criteria of analysis the mean time a person would take from progressing from a category to another. Here we saw two main paths where the duration of progression was considerably higher than the remaining progression paths; those were *Assistente Técnico -> Técnico Superior* (10,4 years) and *Técnico Superior -> Diretor de Departamento* (114,6 months ~ 4.5 years). We also tried to find a correlation between both scopes. However, regardless of the mean duration value, the number of times each path was crossed is low (3-4 passages), therefore we could not clearly distinguish any kind of relation between absolute frequency and mean duration.

We gathered a set of main conclusions concerning this field study based on the previous analysis, the discovered process map and considering Quidgest's HR management system used by the public organization, after applying a variation of replay (a Conformance

Checking technique) to the model (i.e. reproducing the event log in the discovered model but also taking into account the system's business rules) and the system's business rules. The replay analysis was made with the contribution of Quidgest system's manager. Apart from delivering the required event log to do process discovery, he also provided us with insight concerning the discovered map transitions between categories and the business rules that compose the system. In this analysis we used the entire discovered process (i.e. both the Activities and Paths detail filters were removed). The main conclusions were as follows:

- We observed 7 deviations between the process model and the business rules all of the same type (cyclical transitions to the same category);
- A possible reason for 5 of these deviations has to do with an employee returning to the organization. Take the *Formador* category for example as illustrated in Figure 12. Here we have 50 re-entries of the 'Formador' category in the system. One can assume that from all the 198 employees of this category type, 50 were re-employed to do more training events;
- One of the 6 cyclical deviations was found to be a false positive. The reason for this particular deviation (*Secretária Geral* - Secretary) had to do mostly with incorrect register insertion. Despite on a business level not being possible for an employee to have two registers of a same category in the same date, the software does not verify cases when data is introduced incorrectly. This issue makes the discovered process incoherent with the organization business rules. Figure 13 illustrates this scenario;
- The above-mentioned situation does not occur only when an employee gets two registers of a same category in the same day. It also occurs when an employee gets two registers for different category types in the same day, which also violates, in a more serious manner, the organization business rules. This deviation happened in the same category type as the false positive deviation (possibly being the same employee). Figure 13 illustrates the last two mentioned deviations.

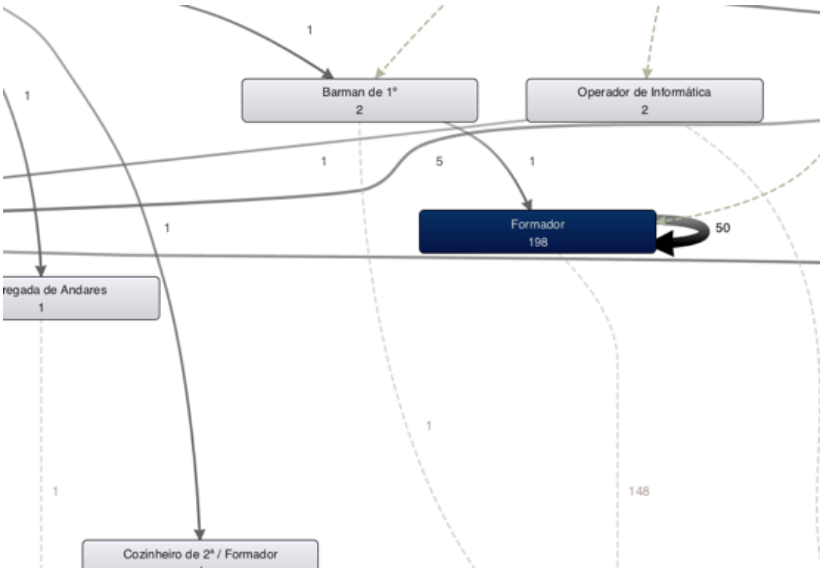


Figure 12 - Discovered Process Map Overview (Zoomed View)

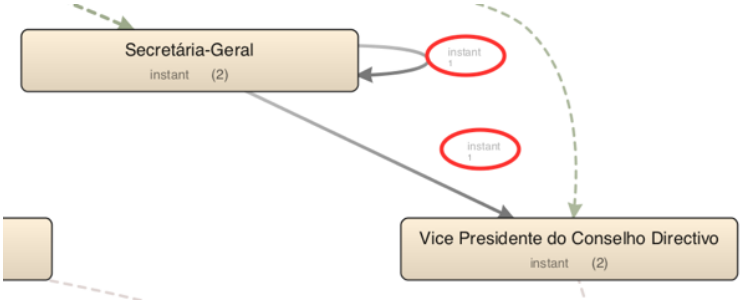


Figure 13 - Discovered Process Map Overview (Zoomed View)

After this analysis we can identify some limitations about applying Process Mining to this type of cases - Career Progression. The first limitation is concerned with the amount of data available. The limitation occurs if the number of events, activities or instances is small. For example, in this case, we had 585 instances (employees) and only 714 events, what does mean that almost all employees only have one event associated to their career. Actually, there are 482 instances with just 1 event and only 1 instance with 6 events. This can be a limitation because there are few instances with a full career registered and, because of it, the results may be less adapted to the reality.

Another limitation is related with changes in the number of employees. If an organization increases the number of employees in the last years, even if there are many employees already working, it will be difficult to capture the usual career progression, because there are insufficient data to the algorithm find the usual progression. As is a public organization, the progression is slow and there is insufficient data, i.e, the majority of employees only changed one time of category. To discover a real career progression through this type of analysis, a large number of years of data is necessary.

The last limitation is data quality and is transversal to all field studies in Process Mining. But, in Career Progression systems, this becomes even worse because in the majority of cases the categories and change dates (timestamps) are manually placed in the information system. This situation creates many problems, for example: three or more categories that represents only one major category and an employee that belongs to more than one category at the same time, violating the business rules.

5. Conclusion

In this paper we applied Process Mining techniques to career progression data of a Portuguese public organization. By performing deep analysis of career progression, we found evidence that Process Mining is suitable to analyse career development and provide insights about employees that would be difficult to achieve only with the raw data about the passages of each employee between different categories. Conclusions like the mean duration of each passage of category or how the development of career along time is, are two of many conclusions that Process Mining analysis provides. However, limitations related to the accuracy and the amounts of data needed were also identified and these are important to alert analysts that, before start analyzing results, the normalization of data is needed.

As future work it will be important to apply Process Mining analysis to different organizational contexts like private organizations, where the career progression is usually faster and there are more subcategories. Furthermore, it will be interesting to compare the analysis results of two different organizations with the same context, to see if a pattern can be assumed or if the results are different between organizations of the same industry. Finally, another comparison that can be done is to see if exists any relation between career progression and performance evaluation. It would be interesting to analyse if performance evaluation has any influence in the career progression of employees.

6. Acknowledgements

We would like to thank to Filipe Santos from Quidgest who helped us to collect and analyse the data. The authors also wish to express their appreciation to Rodrigo Serafim, whose comments and suggestions positively contributed to this paper.

7. References

Burt, R. e Minor M., *Applied Network Analysis: A Methodological Introduction*, Newbury Park CA: Sage http://www.cbpweb.nl/index.htm, 1983.

De Bruin, J. S. Cocx T. K. Kosters W. A. Laros J. F. J. e Kok J. N., "Data Mining Approaches to Criminal Career Analysis," *ICDM '06. Sixth International Conference on Data Mining*, 2006, 171-177.

Edgington, T. M. Raghu T. S. e Vinze A. S., *Using Process Mining to identify coordination patterns in IT service management*, Decision Support Systems, Volume 49, Issue 2, 2010, 175-186.

Fischer, L. (Eds.), *Workflow Handbook 2001, Workflow Management Coalition*, Lighthouse Point, FL: Future Strategies, 2001.

Gunther, C. W. e Rozinat A., BPM 2012 Demonstration Track of the 10th International Conference on Business Process Management (BPM 2012) Tallinn, Estonia, 2012, 40-44.

Jablonski, S. e Bussler C., *Workflow Management: Modeling Concepts, Architecture, and Implementation*, London, UK: International Thomson Computer Press, 1996.

Leymann, F. e Roller D., *Production Workflow: Concepts and Techniques*, vol. 26. 2000.

Nakatumba, J. e Van Der Aalst W., *Analyzing Resource Behavior Using Process Mining*, Lecture Notes in Business Information Processing, Springer, 2010.

Rozinat, A. e Van Der Aalst W., "Conformance checking of processes based on monitoring real behavior," Inf. Syst., vol. 33, no. 1, 2008, 64-95.

Scott, J., *Social Network Analysis*, vol. 22, no. 1, 1994.

Van Der Aalst, W. e Van Hee K. M., *Workflow Management: Models, Methods, and Systems*, Cambridge, MA: MIT press, 2002.

Van Der Aalst, W. Reijers H. A. e Song M., "Discovering social networks from event logs," Comput. Support. Coop. Work, vol. 14, no. 6, 2005, 549-593.

Van Der Aalst, W. Van Hee K. M. Van Werf J. M. e Verdonk M., "Auditing 2.0: Using Process Mining to Support Tomorrow's Auditor," IEEE *Computer* , vol.43, no.3, 2010, 90-93.

Van Der Aalst, W. Weijters T. e Maruster L., "Workflow mining: Discovering process models from event logs", in *IEEE Trans. Knowl. Data Eng.,* vol. 16, no. 9, 2004, 1128-1142.

Wasserman, S. e Faust K., *Social Network Analysis: Methods and Applications*, vol. 8. 1994.

ADOÇÃO DE TECNOLOGIAS DA INFORMAÇÃO: UMA PROPOSTA DE MODELO DE ANÁLISE EXPLORANDO O MODELO UTAUT PARA AMBIENTES VIRTUAIS DE APRENDIZAGEM PARA SUPORTE AO ENSINO PRESENCIAL

Ivo Pedro Gonzalez Junior
Universidade Federal da Bahia - Núcleo de Pós-Graduação em Administração – NPGA –Brasil,
ivojunior@gmail.com

Ernani Marques dos Santos
Universidade Federal da Bahia - Núcleo de Pós-Graduação em Administração – NPGA –Brasil,
ernanims@gmail.com

RESUMO

Com o surgimento dos Ambientes Virtuais de Aprendizagem (AVA) no contexto educacional, as formas de ensino estão passando por transformações, tanto para Ensino a Distância, como também para suporte ao ensino presencial. O crescimento do uso do AVA e a aceitação dessa tecnologia, seu comportamento e o seu sucesso está diretamente relacionado com sua adoção pelos docentes e discentes. A adoção da tecnologia tem sido estudada durante as últimas décadas, e muitos modelos foram criados. Este estudo tem como base a Teoria Unificada de Aceitação e Uso da Tecnologia (*Unified Theory of Acceptance and Use of Technology* - UTAUT), que é um modelo de tentativa de uma reunião de constructos de outras teorias e modelos, porém, aponta lacunas não abordadas e aspectos não analisados em alguns contextos. Este artigo objetiva propor uma extensão do modelo UTAUT para análise da adoção AVA como suporte ao ensino presencial, incluindo elementos como a interatividade, acompanhamento e continuidade de uso.

Palavras-chave: Ambiente Virtual de Aprendizagem, UTAUT, Adoção de Tecnologia, Modelo de adoção, Ensino.

1. INTRODUÇÃO

Com os avanços das Tecnologias da Informação e Comunicação (TIC), os benefícios também se estenderam aos processos inovadores de ensino-aprendizagem, mais especificamente ao computador e a internet, aliados a métodos didático-pedagógicos. Este incremento possibilitou a utilização destes recursos/ferramentas tanto na educação à distância (EAD), e como suporte ao ensino presencial, cuja plataforma tecnológica denomina-se Ambiente Virtual de Aprendizagem (AVA), e suportam o processo de ensino em que professor e aluno estão separados fisicamente, mas ligados através das tecnologias de informação e comunicação.

A introdução de AVAs e outras formas de *e-learning* trouxeram vantagens para o ensino universitário e ajudam atravessar as fronteiras do tempo e do espaço. No entanto, o sucesso

dessas novas tecnologias depende em grande medida da aceitação e uso do professor e do estudante desses ambientes. (RAMOS; OLIVEIRA, 2012).

Segundo Pereira (2007), os AVAs estão sendo cada vez mais utilizados no âmbito acadêmico e corporativo como uma opção tecnológica para atender uma demanda educacional. A partir disso, verifica-se a importância de um entendimento mais crítico sobre o conceito que orienta o desenvolvimento ou o uso desses ambientes, assim como, o tipo de estrutura humana e tecnológica que oferece suporte ao processo ensino-aprendizagem.

A aceitação do indivíduo e à utilização de novas tecnologias tem sido estudadas durante as últimas três décadas, e muitos modelos de aceitação da tecnologia têm sido criados, tais como o Modelo de Aceitação de Tecnologia (*Technology Acceptance Model* - TAM), a Teoria de Difusão de Inovação (*Innovation Diffusion Theory* - IDT) e, mais recentemente, a Teoria Unificada de Aceitação e Uso da Tecnologia (Unified Theory of Acceptance and Use of Technology - UTAUT), entre outros.

Este estudo tem como base a Teoria Unificada de Aceitação e Uso da Tecnologia (*Unified Theory of Acceptance and Use of Technology* - UTAUT), uma teoria de integração metodológica que apresenta capacidade preditiva superior aos modelos anteriores, de acordo com Leal *et al* (2011). De acordo com Li e Kishore (2006), o UTAUT apresenta um dos modelos mais completos sobre aceitação da tecnologia, devendo ser utilizado extensivamente para tal fim.

Bobsin, Visentini e Rech (2009), afirmam que existe uma unanimidade entre os estudos sobre a referência a esse modelo como um avaliador da aceitação da tecnologia da informação. Segundo os autores, os objetos de análise desses estudos variam consideravelmente, com certa tendência de estudos relacionados a tecnologias móveis e ambientes virtuais, mas não se pode fazer generalizações.

Nogueira (2014) identificou uma baixa taxa de adoção de ambientes virtuais de aprendizagem e que grande parte das pesquisas mais significativas de adoção de tecnologia é oriunda da área de Sistemas de Informação, sendo elaboradas por Davis (1989) e Venkatesh (2003), e partindo dos modelos como *Theory of Reasoned Action* (FISHBEIN E AJZEN, 1975), da Psicologia Social, *Technology Acceptance Model* – TAM (DAVIS, 1989) e *Unified of Acceptance and Use of Technology* – UTAUT (VENKATESH *et al*., 2003).

Considerando as diferentes abordagens, buscou-se apropriar-se das contribuições destas teorias para este estudo que trouxe parte do modelo UTAUT (VENKATESH *et al*., 2003) como base para a pesquisa e incorporou os elementos interatividade, acompanhamento e continuidade de

uso. A opção por complementar o UTAUT se justifica, uma vez que este foi criado de uma forma genérica, mais especificamente relacionada à aplicação em empresas e, de acordo com Ibrahim, Khalil, e Jaafar (2011), em função do contexto, é necessário adicionar novas variáveis. Assim, buscou-se nesta pesquisa a adequação do modelo para o processo de ensino, com a utilização de AVA para suporte ao ensino presencial.

2. REVISÃO DE LITERATURA

2.1 AMBIENTES VIRTUAIS DE APRENDIZAGEM

Segundo Dias *et al*. (2011), genericamente, Ambiente Virtual de Aprendizagem (AVA), é entendido como um ambiente em rede utilizado para apoiar o processo de ensino e aprendizagem na educação presencial ou a distância. Na literatura, encontramos algumas definições para ambientes virtuais de aprendizagem que vão desde um entendimento exclusivamente técnico (quando os AVAs são associados simplesmente ao conceito de um "ambiente computacional" ou de uma "plataforma de ensino") a uma concepção ampliada no sentido do potencial da virtualização.

Um ambiente de aprendizagem pode ser entendido como um espaço onde se constrói conhecimento, como convencionalmente conhecemos a sala de aula presencial. Um AVA, considerando a vertente do virtual relacionada à informatização, é um ambiente que se estabelece a partir das tecnologias em rede. Mas compreendendo o virtual também como um estado de problematização, de potência, os AVAs podem ser vistos como espaços de aprendizagem nos quais se ampliam as possibilidades de construção de conhecimento, pois, em potência, são ambientes que promovem e fomentam o questionamento e a problematização, contribuindo, assim, para o processo contínuo de virtualização/atualização inerentes à reflexão. Ademais, podemos compreendê-los como espaços multirreferenciais de aprendizagem (RICCIO, 2010).

2.2. MODELOS DE ADOÇÃO DE TECNOLOGIA

Os estudos sobre adoção de tecnologia ganharam força a partir do final da década de 1980, quando as empresas passaram a investir quantias significativas em recursos tecnológicos e, por muitas vezes, estas inovações não eram utilizadas pelos funcionários. (DAVIS, 1989).

A adoção de tecnologias é uma das mais focadas áreas nas pesquisas de Sistema de Informação, porém estas são fragmentadas e existem vários modelos diferentes que tentam explicar o mesmo fenômeno (VENKATESH *et al*., 2003). Apesar das diferenças nas formulações teóricas, nas teorias e modelos construídos para entender o processo de aceitação e nos construtos utilizados, Agarwal e Prasad (1997) afirmam que todas essas pesquisas, na verdade, buscam entender e explicar o que leva o sucesso de TI, levando em consideração que o sucesso é equivalente ao uso de tecnologia.

Inúmeros fatores influenciam a adoção da TI, podendo inclusive mudar em função do tipo de tecnologia e do contexto onde ela está inserida (KIM; CHAN; GUPTA, 2007). Dessa forma, a adoção de tecnologia tem despertado o interesse tanto em nível de adoção individual, como também organizacional.

Oliveira e Martins (2011) afirmam que dentre as muitas teorias usadas para análise da adoção de TI as mais usadas são: *Technology Acceptance Model (TAM), Theory of Planned Behaviour (TPB), Unified Theory of Acceptance and Use of Technology (UTAUT), Diffusion of Innovation (DOI) e a Technology*, *Organization and Enviroment (TOE)* (quadro 1).

Teoria	Referência	Nível de análise
Teoria da Difusão da Inovação (DOI)	Rogers (1995)	Organizacional
Modelo de Aceitação de Tecnologia (TAM)	Davis (1989)	Individual
Teoria do Comportamento Planejado (TPB)	Ajzen (1991)	Individual
Tecnologia, Organização e Ambiente (TOE)	Tornatzky e Fleischer (1990)	Organizacional
Teoria Unificada de Aceitação e Uso de Tecnologia (UTAUT)	Venkatesh et AL (2003)	Individual

Quadro 1**:** Teorias usadas na análise de adoção de TI.
Fonte: Adaptado de Oliveira e Martins (2011)

Com o intuito de unificar esses modelos e gerar um ainda mais completo, que abrangesse os principais construtos relacionados à aceitação da TI, Venkatesh *et al*. (2003) desenvolveram a Teoria Unificada de Aceitação e Uso da Tecnologia (UTAUT) (quadro 2).

Modelos	Referências
Teoria da Ação Raciona - TRA [*Theory of Reasoned Action*]	FISCHBEIN; AZJEN, 1975
Modelo de Aceitação de Tecnologias – *TAM [Technology Acceptance Model]*–	DAVIS, 1989
MM [Motivational Model – Modelo Motivacional]	DAVIS et al. 1992
Teoria do Comportamento Planejado - TPB [*Theory of Planned Behavior*]	AJZEN, 1991
Modelos TAM e TPB Combinados C-TAM-TPB	TAYLOR; TODD, 1995

[*Combined TAM and TPB*]	
Modelo de Utilização do PC - MPCU [*Model of PC Utilization]*	THOMPSON et al. 1991
Teoria da Difusão da Inovação – TDI - [*Innovation Diffusion Theory*]	ROGERS, 1995
Teoria Social Cognitiva - TSC [*Social Cognitive Theory*]	COMPEAU; HIGGINS, 1995

Quadro 2 - Modelos integrantes do UTAUT
Fonte: Venkatesh et al., 2003

O quadro 3 contém os modelos e construtos usados na definição do modelo UTAUT. Embora algumas teorias não estejam diretamente ligadas com a área da TI, contribuíram significativamente para a construção do modelo.

Modelos	**Variáveis Independentes**
Teoria da Ação Racional	Atitude
	Norma Subjetiva
Modelo de Aceitação da Tecnologia	Utilidade Percebida
	Facilidade de uso percebida
	Norma Subjetiva
Modelo Motivacional	Motivação Extrínseca
	Motivação Intrínseca
Teoria do Comportamento planejado	Atitude para usar tecnologia
	Norma Subjetiva
	Controle Comportamental percebido
Modelo de Utilização do PC	Ajuste ao trabalho
	Complexidade
	Consequências de longo prazo
	Afeito ao uso
	Fatores sociais
	Condições facilitadoras
Teoria da Difusão da Inovação	Vantagem relativa Facilidade de uso percebida Demonstrativo de resultado Julgamento Visibilidade Imagem Compatibilidade Voluntariedade
Teoria Social Cognitiva	Expectativas de resultado
	Auto eficácia
	Efeito
	Ansiedade

Quadro 3 – Teorias/Modelos de Adoção de TI e variáveis formadores do UTAUT
Fonte: Venkatesh et al. (2003)

O modelo proposto por Venkatesh et al. (2003) possui os seguintes elementos: expectativa de performance, expectativa de esforço, influência social e condições facilitadoras. Há também os quatro construtos moderadores da intenção e uso da TI: o gênero, a idade, a experiência do indivíduo e a voluntariedade do uso (o grau pelo qual o uso da tecnologia é voluntário ou livre, ou seja, não obrigatório). O UTAUT é apresentado na figura 1.

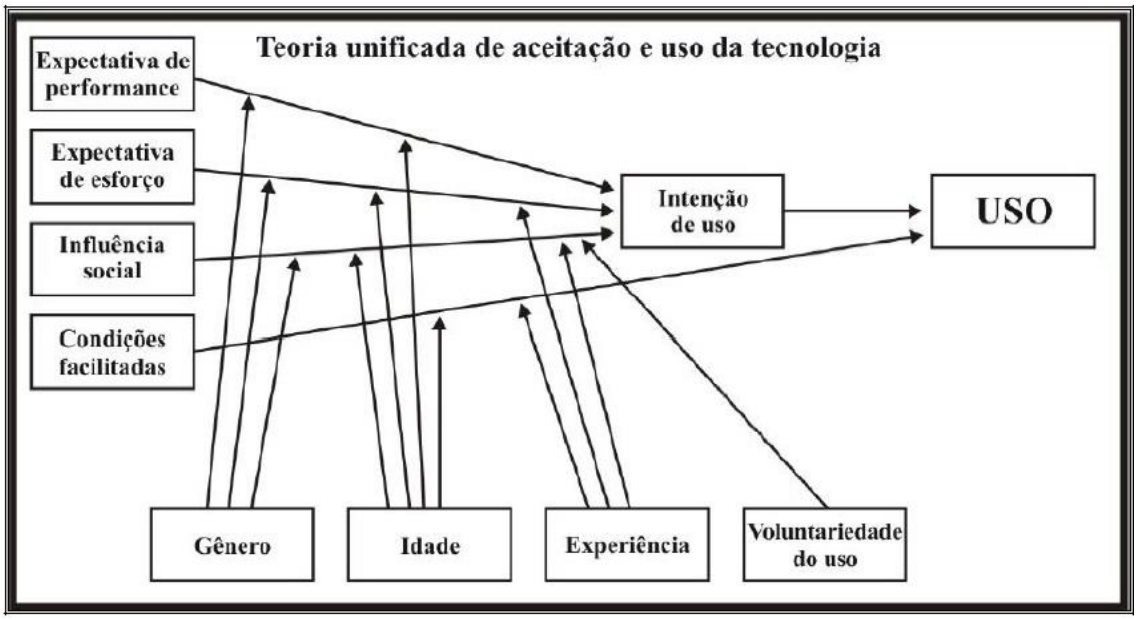


Figura 1 - Modelo Unificado de Aceitação e uso da Tecnologia -UTAUT
Fonte: Venkatesh *et al*. (2003)

2.3 PESQUISAS UTILIZANDO O UTAUT APLICADOS A ÁREA DE ENSINO

O UTAUT tem sido utilizado em várias pesquisas na área de ensino. Anderson *et al* (2006) utilizaram o modelo para verificar a adoção de tablets no ensino universitário norte-americano. Os autores empregaram análise estatística do Partial Least Squares (PLS), e inferiram que o modelo atingiu um poder explicativo de 69% da variância do uso do tablet. A expectativa do desempenho e voluntariedade foram os fatores mais significantes.

Especificamente em relação a AVA, o UTAUT foi utilizado na China por Raij (2008). Os resultados indicaram que a utilidade percebida tem um efeito direto sobre o uso de AVA. Facilidade percebida de uso e norma subjetiva têm apenas efeitos indiretos através da utilidade percebida. Ambos referem à inovação pessoal e ansiedade para ter efeitos diretos sobre a facilidade percebida de uso.

Rezaei, Mohammadi, Asadi e Kalantary (2008) investigaram os fatores externos que podem influenciar a intenção de uso do e-learning em estudantes de agricultura da Universidade de Teerã. Os resultados do artigo apontam que a intenção do aluno em usar o e-learning diminui com o aumento da idade, bem como a utilidade percebida, as experiências em internet, a auto eficácia computacional e a emoção apresentam relações positivas em torno da intenção do estudante em usar o referido sistema.

Brauer e Albertin (2010) baseando-se, principalmente, nos construtos do UTAUT, desenvolveram um modelo para explicar as causas de resistência à EAD na Educação Corporativa. Os resultados apontaram que tal resistência é causada por oito fatores: auto eficácia, competência em TI, expectativa de desempenho, expectativa de esforço, influência social, condições facilitadoras, interatividade e comunicação interna.

Islam (2011) investigou a intenção de continuar a utilização do MOODLE de 175 professores universitários em uma universidade da Finlândia. A pesquisa utilizou o modelo UTAUT fazendo uma adaptação no fator Condições Facilitadoras e dividindo-o em três fatores: Compatibilidade, Acesso e Controle do Comportamento Percebido. Os resultados indicaram que os principais fatores que influenciam a intenção de continuar o uso foram expectativa de desempenho e acesso. Expectativa de esforço e influencia social não demonstraram influência direta sobre a intenção de continuar o uso do MOODLE.

Oye *et al*. (2011) investigaram professores universitários na Nigéria sobre o uso de TIC utilizando o modelo UTAUT. Os resultados demonstraram que todos os fatores tiveram relação com a intenção de uso, e o mais influente foi a expectativa de desempenho.

2.4 CRÍTICAS E LIMITAÇÕES DO UTAUT

Embora o UTAUT seja um modelo robusto e validado em diferentes contextos, existem algumas limitações. Com base nas teorias adotadas em sistemas de informação e em psicologia social, a intenção comportamental intercede à influência de várias crenças e variáveis externas sobre o comportamento, e é um importante preditor de comportamento. No entanto, ele tem três limitações:

1. a intenção comportamental é um reflexo de uma esquema interno do indivíduo de crenças (VENKATESH *et al*., 2008), e não representa os fatores externos que podem afetar o desempenho de um comportamento; assim, o papel da externo das variáveis que podem potencialmente impedir ou facilitar o desempenho de um comportamento não é

totalmente capturado por intenção comportamental (que não considera plenamente todos os possíveis fatores externos);

2. a intenção comportamental tem uma capacidade preditiva e explicativa fraca para lidar com a incerteza e imprevistos de eventos entre o momento em que a intenção é formada e o comportamento é realizado. (SEDIGHEH; SALLEH; MASOUD, 2013). No âmbito de novas informações, uma crença individual e intenção comportamental pode e deve se alterar. De acordo com Venkatesh *et al*. (2008), vários estímulos internos e externos podem mudar drasticamente a intenção provisória ao longo do tempo, tornando intenção comportamental imprecisa, instável e menos previsível do comportamento;
3. a intenção comportamental tem uma fraca capacidade de prever comportamentos que não estão completamente dentro do controle do indivíduo (VENKATESH *et al*., 2008).

Críticos postulam que o modelo UTAUT é uma reencarnação da Teoria da Ação Racional e Teoria Comportamento Planejado (BENBASAT; BARKI, 2007). Outros indicam o modelo UTAUT, mesmo com quarenta e uma variáveis independentes para predizer as intenções e pelo menos oito variáveis independentes para prever o comportamento, omita variáveis independentes importantes (BAGOZZI, 2007). Esses pesquisadores sugerem uma modelo que conceitue melhor a utilização do sistema, olhando para uma perspectiva mais ampla do usuário, que integre estudos longitudinais, e identifique as causas das crenças inerentes à adoção. Os autores complementam afirmando que o modelo UTAUT, tem fornecido resultados inconsistentes como um preditor de intenção comportamental.

A limitação da Teoria Unificada de aceitação e utilização do modelo de tecnologia é também a sua inflexibilidade de se adaptar a diferentes contextos. Neste contexto, Al-Gahtani, Hubona e Wang (2007) relataram em sua pesquisa sobre a aceitação da tecnologia da informação na Arábia Saudita, que é um país do Oriente Médio, a diferença cultural da Arábia Saudita a partir de um país ocidental típico tornou-se um obstáculo para que eles que utilizam o UTAUT para analisar a adoção do trabalhador de computadores na Arábia Saudita.

A auto eficácia e atitude são outras limitações do modelo UTAUT. Venkatesh *et al*. (2003) não incluem como um determinante direto da intenção comportamental no modelo UTAUT, e considera-se como uma construção indireta, medindo uma auto eficácia específica, não em um computador a auto eficácia global em direção a uma tecnologia específica (STRAUB, 2009). Os usuários inexperientes visualizam novas tecnologias (inovação de TI) como complexas, e a confiança na própria capacidade de lidar com elas tem uma influência significativa sobre a sua aceitação (YUEN *et al*., 2008).

Uma limitação apontada por Brauer (2008) é a interatividade. A interatividade entre professores e alunos é importante, visto que sem essa interação a qualidade do aprendizado pode ficar comprometida. Cursos sem interatividade ou com interatividade fraca têm risco maior de os alunos ficarem desestimulados, o que pode influenciar bastante nos resultados do aprendizado.

3. MODELO PROPOSTO A PARTIR DA EXTENSÃO DO UTAUT

Após a abordagem das limitações das teorias e do modelo escolhido – UTAUT, percebeu-se que o modelo não aborda alguns conceitos importantes para verificar a aceitação e uso da tecnologia AVA. Como complemento a este modelo, foi proposta a inserção de três elementos: interatividade, acompanhamento e continuidade de uso, descritos a seguir.

3.1 INTERATIVIDADE

Antes de compreendermos a importância da interatividade é preciso definir alguns conceitos recorrentes na maioria dos estudos que abordam essa temática, ou seja, o significado dos termos interação e interatividade. Segundo Pibreran (2011) o termo interação refere-se ao "fenômeno que permite a certo número de indivíduos constituir-se em grupo, e que consiste no fato de que o comportamento de cada indivíduo se torna estímulo para o outro". Já o termo interatividade está relacionado a componentes tecnológicos que promove a "[...] permuta entre o usuário de um sistema informático e a máquina, por meio de um terminal dotado de um 'ecrã' de visualização" (PIBRERAN, 2011).

Deste modo, a utilização de salas virtuais ou AVA como um recurso tecnológico interativo possui o objetivo de promover o diálogo entre professores e alunos, bem como fomentar a transmissão do conhecimento por meio de ferramentas apropriadas.

A interação possibilitada por um AVA é muito importante, devido ao fato que sem ela a qualidade do aprendizado pode ser comprometida. A interação entre professor e aluno e entre alunos pode fazer a diferença nos resultados do aprendizado, e a interatividade em níveis altos podem influenciar também o estímulo (KENSKI, 2012).

De acordo dom Kenski (2012), as tecnologias ampliam as possibilidades de ensino para além do curto e delimitado espaço de presença física de professores e alunos na mesma sala de aula. A possibilidade de interação entre os professores e alunos, e entre alunos, objetos e informações que estejam envolvidos no processo de ensino redefine toda a dinâmica da aula e cria novos vínculos entre os participantes (KENSKI, 2012).

Um conceito para medir o grau de interação em atividades educativas é apresentado por Moore (2004), chamado de "distância transacional", ou seja, a distância física e comunicativa seja ela em sala ou no ambiente virtual. No seu estudo este autor afirma que a distância transacional será maior ou menor, dependendo da forma como os alunos são tratados e interagidos. Assim a distância transacional atinge seu auge quando os docentes e discentes não tem qualquer interação e interatividade. Assim os conceitos de interação e interatividade estão inter-relacionados, sendo que um influencia o outro. Neste contexto para este estudo, iremos abordar o termo interatividade, por representar mais a tecnologia e o modelo de adoção proposto.

A lógica do vetor interação estabelece que a uso ou resistência ao uso da tecnologia decorre do relacionamento pessoas e grupos e sistemas, ou seja, a forma como as pessoas e grupos e suas características peculiares interagem com as especificidades dos softwares que devem ser implantados para uso dessas mesmas pessoas ou grupos (JOIA; MAGALHÃES, 2009).

Segundo o Cegoc (2007), ao saberem que alguém está atento ao que fazem, os alunos se sentem mais estimulados a envolverem-se com uso do AVA. A interação entre os participantes e o formador é algo fundamental para manter elevados os níveis de atenção e estímulos ao longo do curso.

Galusha (1997) afirma que para todo tipo de instrução é importante que o aluno tenha um feedback rápido, e ambientes virtuais possibilitam esta interação rápida. A ausência do professor virtual é um fator de insatisfação e desestímulo. É provável que uma interatividade fraca entre os alunos virtuais aumentam a possibilidade de resistências aos AVAs.

A interatividade pode ter influência na intenção de uso e o uso da tecnologia AVA como suporte ao ensino presencial. Este elemento é influenciado pela experiência anterior com tecnologia. Indivíduos com experiência anterior tendem a ter mais interatividade. Níveis maiores de interatividade pode influenciar positivamente a adoção de AVA.

3.2 ACOMPANHAMENTO

Após o efetivo uso e aceitação da tecnologia, outro elemento pode ser agregado, o acompanhamento. Perez (2008) afirma que a forma pela qual um sistema é adotado depende diretamente dos atributos percebidos seus usuários. Estes atributos necessitam ser acompanhados e essas percepções decorrem do seu envolvimento com o processo de adoção.

Moore e Benbasat (1991) introduziram novos atributos ao modelo de difusão da inovação, porém não inseridos no modelo UTAUT. A junção dos construtos observabilidade com a

compatibilidade, identificada inicialmente, por Rogers (1983), transformou-se na categoria Demonstração do Resultado, que indica o grau em que os resultados do uso de uma inovação são tangíveis e acompanhados. Outro atributo é visibilidade, definido como o grau em que uma inovação se torna visível para os indivíduos ou grupos de uma organização (MOORE; BENBASAT, 1991).

Segundo Perez (2010), a tecnologia de informação para apoio ao ensino superior, necessita avaliar alguns atributos de inovação, alguns já previstos no UTAUT, e outros não, que são: Visibilidade, Resultados. Assim os efeitos proporcionados com o acompanhamento do seu uso teriam consequências para as pessoas (mudanças na forma de procura e obtenção de informação, aprendizado/qualificação), na flexibilidade para o futuro (continuidade do uso).

Venkatesh e Davis (2000) observaram uma relação direta entre demonstração de resultado e utilidade percebida. Os autores argumentaram que os indivíduos formam percepções mais positivas do sistema, se a relação entre o uso e desempenho é facilmente discernível.

O Acompanhamento reúne conceitos de demonstração de resultados, visibilidade como suporte ao ensino presencial. Dessa forma, é definido como o grau em que o acompanhamento do uso de uma tecnologia são tangíveis. Um ponto importante a se destacar, que para se considerar isto é necessário ter as mesmas condições dos constructos e moderadores para a intenção e uso da tecnologia.

3.3 CONTINUIDADE DE USO

Segundo Hertas (2007), na perspectiva educacional, o sucesso das tecnologias no ensino depende tanto do seu estado de adoção inicial quanto do seu uso continuado. Tal situação pode gerar uma série de oportunidades e desafios que transformam o ambiente educacional a partir da modificação de fatores tecnológicos, sociais e organizacionais, que pode acontecer através da aquisição ou criação de inovações.

Apesar do reconhecimento generalizado do valor AVA, muitos usuários interrompem o uso destes ambientes após a aceitação inicial de seu uso (LEE, 2010). Assim, os benefícios esperados com a utilização destes ambientes, que complementam o ensino presencial ou até mesmo substituem a forma tradicional face-a-face de aprendizagem, não se refletem de forma adequada na sua efetiva adoção (WU; TSAI; CHEN; WU, 2006).

Davis (1989) em seu modelo define utilidade percebida definido como "o grau em que uma pessoa acredita que a utilização de um sistema particular aumentaria a sua ou seu desempenho no trabalho". Seddon (1997) sugere que, no uso da palavra "aumentaria" na definição de Davis

possa ser substituído por "Melhoraria". O argumento de Seddon (1997) que "não importa quão bom um sistema tenha sido no passado, o benefício não é uma condição suficiente para uso futuro, assim, o uso futuro deve basear-se em expectativas de benefício futuro ".

Kim e Malhotra (2005) afirmam que é plausível que a aceitação inicial afeta mais tarde a continuação. Além disso, tanto a satisfação e utilidade percebida são motivações extrínsecas, e as pessoas têm tanto motivações intrínsecas e extrínsecas associadas à sua aceitação e uso continuado de TI (MALHOTRA *et al*., 2008). Sorebo et al. (2009) constataram que é necessário explicar a continuidade de uso de ambientes virtuais.

Segundo o estudo de Hung *et a*l (2011), a medição de continuidade de uso pode determinar o sucesso de um sistema *e-learning*, e seu acompanhamento do pós-uso tem um efeito significativo sobre a continuidade de uso de um AVA.

Além disso, Hayashi *et al*. (2004) sugeriu que a continuidade de uso de AVA depende da diferença entre expectativas e experiência de uso. Este estudo também descobriu que a utilidade percebida, considerado como expectativas pós-uso e atribuições, tem um efeito significativo sobre continuidade de uso.

Assim, nesta pesquisa, a continuidade de uso é definida como: diferença entre expetativa e experiência do uso. Ela tem relação bidirecional com o acompanhamento e o uso. Isto mantendo-se as mesmas condições dos construtos e variáveis moderadoras já definidas.

O modelo proposto, estendido do UTAUT para a adoção de AVA como suporte ao ensino presencial pode ser visualizado na figura 2. As marcações em cor destacada demonstram a extensão realizada no UTAUT.

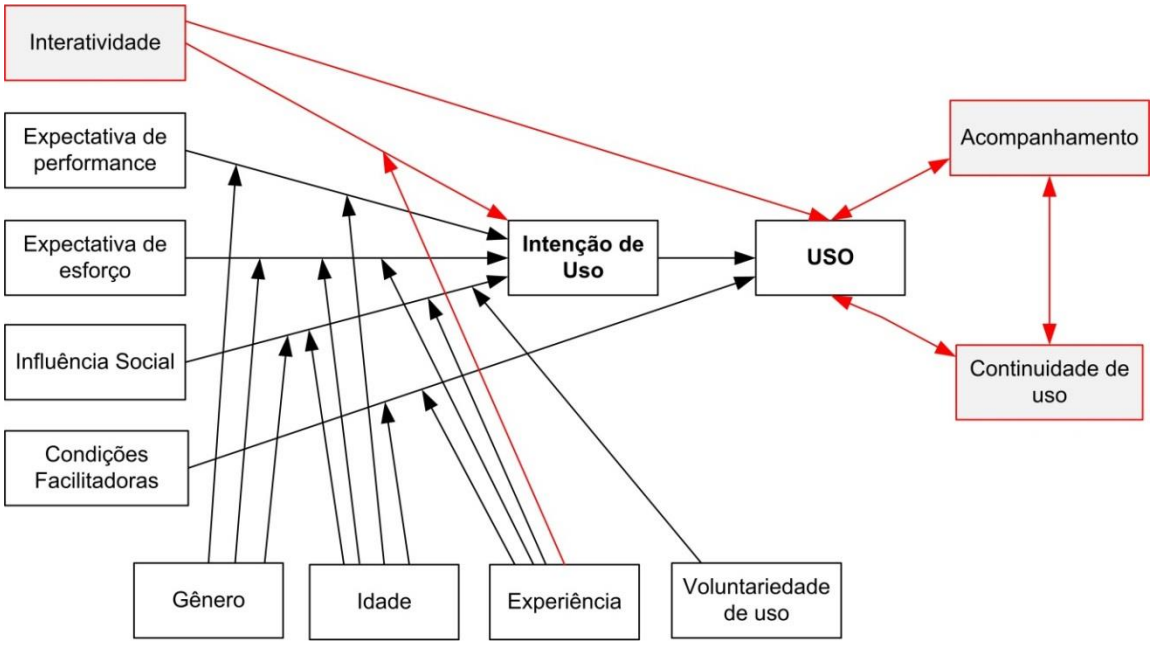


Figura 2 - Modelo estendido do UTAUT para AVA
Fonte: Elaborado pelos autores a partir de Venkatesh *et al*. (2003).

Após o modelo proposto são definidos no Quadro 4 os Fatores e definições teóricas da pesquisa, com os indicadores.

Fator	Referência	Definição	Indicadores
Expectativa de Desempenho	Venkatesh et al (2003)	O grau em que o indivíduo acredita que utilizando a tecnologia ela ajudará a melhorar o desempenho no trabalho.	1- O uso do AVA permite realizar as tarefas mais rapidamente 2- O uso do AVA torna mais fácil a realização trabalho/estudo 3- O uso do AVA melhora a qualidade do trabalho/estudo 4- O AVA é útil no trabalho/estudo
Expectativa de Esforço	Venkatesh et al (2003)	O grau de facilidade associado à utilização do sistema.	1- Facilidade em utilizar o AVA 2 – Facilidade em Aprender a usar o AVA 3- Dificuldade no uso do AVA 4- O AVA se encaixa bem na forma como trabalho/aprendo
Influência Social	Venkatesh et al (2003)	È o grau em que um indivíduo percebe que pessoas importantes acreditam que ele deve utilizar o novo sistema.	1- Pessoas que são importantes para mim pensam que eu deveria utilizar o AVA 2-Pessoas que influenciam meu comportamento pensam que eu deveria utilizar o AVA 3- Pessoas acima pensam que eu deveria utilizar o AVA 4- Em geral, a universidade tem apoiado o uso do AVA
Condições	Venkatesh et al	É o grau em que o	1- Existe bom suporte ao uso do AVA ;

Fator	Referência	Definição	Indicadores
Facilitadoras	(2003)	indivíduo acredita que existe uma estrutura técnica organizacional para dar suporte na utilização do sistema.	2- Fácil acesso a internet e ao sistema AVA 3- Uma pessoa específica (ou grupo) está disponível para dar assistência nas dificuldades; 4- Eu tenho os recursos necessários para usar o sistema AVA
Uso	Venkatesh et al (2003)	Utilização Efetiva do Uso	1-Utilização do AVA 2-Início de utilização do AVA 3- Utilização em todas as disciplinas 4-Recursos utilizados 5-Periocidade de uso
Interatividade	Galusha, (1997); Kenski (2012)	O grau de interatividade entre os professores-alunos, alunos-alunos e o uso do AVA.	1-Minha interação como o AVA é clara e de fácil compreensão 2-O AVA possibilita entrosamento entre Professor e aluno 3-O AVA possibilita o entrosamento entre aluno e aluno 4-No AVA, o feedback (retorno) é rápido
Acompanhamento	Moore; Benbasat (1991)	o grau em que o acompanhamento do uso de uma tecnologia são tangíveis.	1-Minhas dúvidas com são sanadas de forma satisfatória; 2-No AVA o professor monitora bastante meu aprendizado 3-No AVA a universidade monitora bastante meu aprendizado 4- A demonstração de resultados aponta para melhoria e continuidade do uso 5-A visibilidade uso do AVA possibilita melhoria no meu trabalho/estudo
Continuidade de uso	Hertas (2007); Hung et al (2011)	A diferença entre expectativas e experiências de uso.	1- Eu pretendo continuar utilizando o AVA no futuro; 2- Pretendo utilizar mais o AVA no futuro; 3-Pretendo fazer um curso ou buscar informações de utilização do AVA no futuro; 4-Os benefícios com a utilização do uso aumentam as expectativas 5- Há preocupação da faculdade com a atualização do AVA para uma continuidade?

Quadro 4 - Fatores e definições teóricas da pesquisa
Fonte: elaborado pelos autores

4. CONSIDERAÇÕES FINAIS

A partir deste modelo proposto, objetiva-se compreender a adoção de AVA como suporte ao ensino presencial em instituições de ensino superior. Sua utilização poderá ajudar a identificar

os fatores que condicionam os agentes envolvidos no ensino e aprendizado na adoção e uso destas tecnologias.

Este modelo poderá ajudar também a compreender questões relacionadas a diferenças entre o comportamento de intenção de uso e o uso propriamente dito ao adotar esta tecnologia. Compreendendo questões como essas, vislumbra-se ser possível melhorar a utilização dos ambientes virtuais de aprendizagem pelos professores e alunos.

Como próximos passos, é necessária a definição dos indicadores de cada uma das dimensões propostas, de modo a operacionalizar a aplicação do modelo para sua validação e/ou ajustes necessários.

Assim, o presente trabalho buscou evidenciar a relevância de revisitar o tema da Adoção de TI e adicionalmente propor um modelo de adoção de AVA. A partir da proposição, sugere-se que pesquisas futuras venham a aplicar o modelo proposto para que se verifique a riqueza exploratória de análise em instituições de ensino superior, em virtude de sua relevância, conforme dados apresentados neste trabalho.

REFERÊNCIAS

Agarwal, R.; Prasad, J. The role innovation characteristics and perceived voluntariness in the acceptance of information technology. *Decision Sciente*, v. 28, n.3, p. 557-582, 1997.

Ajzen, I. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, 50 (2), 179-211, 1991.

Al-Gahtani, S.; Hubona, G. S.; Wang, J. Information technology (IT) in Saudi Arabia: Culture and the Acceptance and use of IT. *Information & Management*. 2007.

Anderson, B. Comparison of student performance in paper-based versus computer based testing. *Journal of education for business*, 84 (1), 13-17

BAGOZZI, R. P. The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift. Journal of the Association for Information Systems, 4, 244-254, 2007.

Baker, J. *The Technology-organization-environment Framework*. In: Information Systems Theory: explaining and predicting our digital society. Vol. 1. Springer Science Business Media, 2012.

Benbasat, I.; Barki, H. Quo Vadis, TAM? Journal of the Association for Information Systems, 4, 211-218, 2007

Bobsin, D. Em busca do estado da arte do UTAUT: ampliando as considerações sobre o uso da tecnologia. RAI - *Revista de Administração e Inovação*, São Paulo, v. 6, n. 2, p. 99-118, 2009.

Brauer, M. *Resistência à Educação a Distância na Educação Corporativa*. 188f. Tese (Doutorado em Administração de Empresas) – Programa de Mestrado Acadêmico e Doutorado em Administração de Empresas da Fundação Getúlio Vargas, São Paulo, 2008.

Brauer, M.; Albertin, A. L.. Educação corporativa a distância: Por que tanta resistência? *Revista de design, inovação e gestão estratégica*, v. 1, n. 1, p. 141-157, 2010.

Cegoc. *Estratégias para motivar online*. Disponível em: <http://www.cegoc.pt/elearning/artigos.htm>. Acesso em: 20 mar. 2007.

Chiu, C. M.. Wang, E. T.G. Understanding Web-based learning continuance intention: The role of subjective task value. *Information and Management*. 2008.

Compeau, D. R.; higgings, C. A. Computer self-efficacy: development of a measure and initial test. *MIS Quaterly*, 19(2), 189-211, 1995.

Davis, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, Minneapolis, v. 13, n. 3, p. 318-339, Sept. 1989.

Davis, F. D., bagozzi, R. P., & Warshaw, P.R. Extrinsic and intrinsic motivaion to use computers in the workplace. *Journal of Applied Social Psychology*, 22(14), 1111-1132, 1992.

DIAS et al. Ambiente virtual de aprendizagem como ferramenta para o estudo extra-classe e educação continuada. *Revista Cogitare*. Jul/Set; 8-16. 2011.

Fishbein, M., Ajzen, I. *Belief, attitude, intention, and behavior*: an introduction to theory and research. Reading: Addison-Wesley: 1975.

Galusha, J. M. *Barriers to Learning in Distance Education*. Interpersonal Computing and Technology: an electronic journal for the 21st century. 1997. Disponível em:

Hayasashi, A., Chen, C., Ryan, T. Wu, J. The role of social presence and moderating role of computer self-efficacy in predicting the continuance usage of e-learning systems. *Journal of Information Systems Education*, 15(2), 139–154. 2004

Hertas, A. *Teaching and learning logic in a virtual learning environment*. Oxford University Press, v. 15, n. 4, p. 321-331, 2007.

Hu, P. J.; Chau, P. Y. K.; Sheng, O. R. L.; Tam, K. Y. Examining the Technology Acceptance Model using physician acceptance of telemedicine technology. *Journal of Management Information Systems*, v. 16, n. 2, p. 91-112, 1999.

Hung, M. C. et al. Exploring academic teachers' continuance toward the web-based learning system: The role of causal attributions. *Computers & Education* 57 p. 1530–1543. 2011.

Ibrahim, R; Khalil, K; Jaafar A. Towards educational games acceptance model (EGAM): a revised unified theory of acceptance and use of technology (UTAUT). International *Journal of Research and Reviews in Computer Sciente*, 2(3), 839-846

Islam, A. K. M.N Understanding the continued usage intention of educators toward an e-learning system. *International Journal of E-Adotion*, 3 (2) 54-69. 2011

Joia, L. A., Magalhães, C.. Evidências Empíricas da Resistência à Implantação de Prescrição Eletrônica: Uma Análise Explano-Exploratória. *RAC - Eletrônica*, v. 3, n. 1, art.5, p. 81-104, 2009.

Kenski, V. M. *Educação e tecnologias: o novo ritmo da informação*. 8ºed. Papirus, Campinas-SP: 2012.

Kim, H., Chan, H.; Grupta, S. Value-based adoption of mobile internet: an empirical investigation. *Decision Support Systems*, 43(1), 111-126. 2007.

Leal, E. A. ; Albertin, A. L. ; Pereira,Jj. M. ; Nomelini, Q. S. S. . *Utilização da Análise Fatorial para Identificação dos Fatores Determinantes da Aceitação do Uso de Tecnologias de Informação na Educação a Distância*. In: XXXV EnANPAD, 2011, Rio de Janeiro. Anais do XXXV EnANPAD 2011.

Lee, M. C.. Explaining and predicting users´continuance intention toward e-learning: An extension of the expectation-confirmation model. *Computers & Education*, v. 54, p. 506-516, 2010.

Li J. P.; Kishore, R. How robust is the UTAUT instrument? a multigroup invariance analysis in the context of acceptance and use of online community weblog systems. In: 2006. Acm sigmis cpr conference on computer personnel research: forty four years of computer personnel research: achievements, challenges & the future, 2006, Claremont. Proceedings… New York: ACM, 2006. p. 183-189.

Malhotra, Y., Galletta, D. F., Kirsch, L. J.: How endogenous motivations influence user intentions: beyond the dichotomy of extrinsic and intrinsic user motivations. Journal of Management Information Systems, 25(1), 267–299, 2008.

Moore, G.; Benbasat, I. Development of an insrument to measure the perceptions of adopting and information technology innovation. *Information Systems Research*, v.2, n. 3, p. 192-22, 1991.

Moore, M.; Kearsley, G. *Distance education*: a systems view. Belmont: Wadsworth Publishing Co.,1996.

Nogueira, D. R. *Vento da mudança: estudo de caso sobre a adoção de ambientes virtuais no ensino presencial em Contabilidade* / Tese (Doutorado) – Universidade de São Paulo. São Paulo, 2014.

Oliveira, T.; Martins, M. Literature Review of Information Technology Adoption Models at Firm Level. *The Electronic Journal Information Systems Evaluation*, v. 14, n. 1, p. 110-121, Jan. 2011.

Oye, N, et al A model of ICT acceptance and use ofr teachers in higher education instituitions. *International Journal of Computer Sciente e Communication Networks*, 1 (1) 22-40. 2011

Pereira, A. T. Cybis. (org.). AVA - *Ambientes Virtuais de Aprendizagem em Diferentes Contextos*. Rio de Janeiro: Editora Ciência Moderna, 2007.

Perez, G.; Zwicker, R. *Fatores determinantes da adoção de sistemas de informação na área de saúde: um estudo sobre o prontuário médico eletrônico*. 2012

Pibreran, *Dicionário de Língua Portuguesa*. (online) Disponível em: < http://www.priberam.pt/dlpo/ > Acesso em: 13 de julho de 2014.

Raaij,. M. V; Schepers, J. J.L. The acceptance and use of a virtual learning environment in China. *Computers & Education* 50 p.838–852. Elsevier, 2008.

Ramos, F. et al. Perspectivas e práticas em e-learning no ensino superior e no ensino ao longo da vida em Portugal na Irlanda e no Reino Unido. In: SILVA, A. C. da (Org.). *Aprendizagem em Ambientes Virtuais e educação a distância*. Porto Alegre: Editora Mediação, 2009. p.18-52.

Ramos, S.M; Oliveira, B.M.K. *Fatores que influenciam a aceitação e uso de um ambiente virtual de aprendizagem no contexto de um curso de capacitação para servidores públicos*. XV SEMEAD. Out 2012

Rezaei, M.; Mohammadi, h. M.; Asadi, A.; Kalantary, K. Predicting e-learning application in agricultural higher education using technology acceptance model. *Turkish Online Journal of Distance Education*, v. 98, n. 1, p. 85-94, 2008.

Riccio, N. C. R. *Ambientes Virtuais de Aprendizagem na UFBA:* a autonomia como possibilidade. 348f. 2010. Tese (Doutorado) – Faculdade de educação, Universidade Federal da Bahia, Salvador, 2010.

Rogers, E.M. *Diffusion of innovations*. 4th Edition. The Free Press, New York:1995.

Sedigheh; Salleh; Masoud,. Determinants of it-related innovation acceptance and use behavior: theoretical integration of unified theory of acceptance and use of technology and entrepreneurial potential model. Social Technologies. 2013, 3(2): 243–260, 2013

Seddon, P. B. A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research*, 8(3), 240–253, 1997.

Silva, A.C.C A implantação das dependências o-lin da PUCPR – campus Londrina. Dissertação de Mestrado, PUC-PR Curitba 2006

Sorebo, O.; Halvari, H.; Gulli V. F.; Kristiansen, Roar: The role of self-determination theory in explaining teachers' motivation. Kongsberg, Noruega, 2009

Straub, D. Deil, M; Brenner. W Testting the technology acceptance model across cultures: a three country study. *Information & Management*, 33(1) 1-11, 2009

Tornatzky, L. G., Klein, K. J. *Innovation Characteristics and Innovation Adoption Implementation:* A Meta-Analysis of Findings. IEEE Transactions on Engineering Management EM 29(1), 28-45. 1982

Venkatesh, V. et al. User acceptance of information technology: toward a unified view. *MIS Quarterly*, Minneapolis, v. 27, n. 3, p. 425–478, Sept. 2003.

Venkatesh, V., Brown, S.A., Maruping, L.; Bala, H., 2008. Predicting different conceptualizations of system use: The competing roles of behavioral intention, facilitating conditions, and behavioral expectation, *Mis Quarterly*, vol. 32, no. 3, pp. 483-502.

Venkatesh, V.; Sykes, T. A.; Zhang, x. '*Just what the doctor ordered': a revised UTAUT for EMR system adoption and use by doctors*. Proceedings of the 44th Hawaii Internacional Conference on System Sciences, Hawaii, 2011.

Yuen, A.; Ma, W. *Exploring teacher acceptance of e-learning technology*. *AsiaPacific*, *Journal of Teacher Education*, v. 36, n.3, p. 229-243. 2008.

Análise da relação entre processos de gestão do conhecimento e de gestão de projetos

Fernanda Gomes [1], Mírian Oliveira [2]

1) Pontifícia Universidade Católica do Rio Grande do Sul, Brasil
fernanda.gomes@live.com

2) Pontifícia Universidade Católica do Rio Grande do Sul, Brasil
miriano@pucrs.br

Resumo

O presente trabalho pretende analisar as relações entre os Processos de Gestão do Conhecimento e os Processos de Gestão de Projetos sob três perspetivas: tecnologias que suportam os processos de Gestão do Conhecimento, atores envolvidos e tipo de conhecimento originado dessa interação. Trata-se de um trabalho teórico que, à luz da literatura, consolida os processos de Gestão de Conhecimento e os Processos de Gestão de Projetos considerando perspetivas abordadas por diferentes autores e, por fim, é mapeada a relação entre esses processos.

Palavras chave: Gestão do Conhecimento, Gestão de Projetos, Processos

1 Introdução

O rompimento das fronteiras geográficas introduziu nas organizações maiores níveis de complexidade em quesitos como a competição por expansão de mercado e a gestão das expetativas dos clientes [Dey et al. 2010]. Aliados aos fatores competitivos provenientes do mercado consumidor, os administradores precisam lidar também com fatores de complexidade internos relacionados aos seus processos de produção e de gestão e, consequentemente, com os seus colaboradores. Um fator agravante está no fato da força de trabalho estar "envelhecendo", fator este que já impacta indústrias ao redor do mundo [Pollack 2012], e que pode ser controlado e minimizado através da retenção do conhecimento destes profissionais por parte das organizações. A Visão Baseada em Conhecimento considera que este é o ativo mais importante das organizações, uma vez que é essencial para que os produtos e serviços possam ser disponibilizados ao mercado de acordo com as estratégias por estas adotadas [Grant 1996].

O início do século XXI marca a adoção, por parte de muitas empresas, da Gestão do Conhecimento de forma multidisciplinar e como agente integrador dos processos, recursos humanos e tecnologia [Mitrevski e Aceski 2015], apesar do valor do conhecimento dos colaboradores já ser um conceito mais antigo [Davenport e Prusak 1998]. O conhecimento é considerado um bem intangível e contribui para a sustentação da vantagem competitiva ao longo do tempo [Lee e Kim 2001] e, portanto, a sua gestão eficaz torna-se fundamental. A Gestão do Conhecimento (GC) pode ser vista como uma coleção de processos com o objetivo de criar, utilizar e disseminar o conhecimento entre os colaboradores da organização [Teixeira et al. 2004], seja este de natureza tácita ou explícita [Nonaka 1994], pois ambos requerem formas específicas de gestão e são importantes no que tange a sustentação da vantagem competitiva e, portanto, sua gestão deve estar alinhada à estratégia da organização.

A estratégia organizacional, muitas vezes, é implementada através de projetos, que são esforços únicos e que servem como meio para mudança e evolução [Pereira 2011], e permite gerir os recursos com foco na minimização dos riscos e custos [Juliani et al. 2012]. Assim como a GC, os projetos também envolvem a gestão de múltiplos processos, inclusive a forma como os colaboradores aprendem durante e depois do encerramento destes [Jugdev 2012]. O Modelo *Knowledge Management in Inter Project Learning* [Fitzek 2002] sustenta que uma organização deve basear-se em um modelo de compartilhamento e aprendizagem interprojeto, e identifica que essa transferência ou compartilhamento do conhecimento pode ocorrer de forma simultânea (quando ambos os projetos estão em andamento, e ocorre a comunicação direta entre os times de projeto) ou de forma sequencial (onde o conhecimento é passado de um projeto prévio para um projeto em execução) [Fitzek 2002].

Projetos são atividades intensas em conhecimento e estudar a forma como o conhecimento é tratado nos projetos se torna fundamental a fim de possibilitar uma gestão eficaz [Reich 2007]. Entretanto, ainda não há um consenso na literatura em relação a como realizar a GC em projetos (ou em empresas "projetizadas"). Portanto, principalmente em projetos de Tecnologia da Informação (TI), essa é uma área de pesquisa promissora a ser explorada [Lech 2014]. A gestão de projetos (GP) é complexa porque projetos envolvem múltiplos desafios como alianças, concorrência (interna e externa), recursos e questões contratuais [Jugdev 2012]. A GC, em contrapartida, pode representar a capacidade das empresas em adaptar-se e sobreviver aos diversos desafios impostos pelo mercado [Goldoni e Oliveira 2007]. A GC pode ser vista como "a coleção de processos que governam a criação, disseminação e desenvolvimento do conhecimento para atingir os objetivos organizacionais" [Lee e Yang 2000 p.784].

A GP (o que inclui projetos de Tecnologia da Informação) vem sendo estudada há décadas, porém os estudos são concentrados em uma perspetiva de ações, orçamento, cronograma, *etc*. [Lech 2014]. A perspetiva de GC em projetos também deve ser estudada a fim de alavancar os fenômenos da GP [Reich et al. 2008], uma vez que projetos lidam intensamente com o conhecimento dos membros dos seus times. A fim de contribuir com este *gap* identificado, este artigo tem o seguinte objetivo: propor a relação entre os processos de GC e de GP, no que tange às tecnologias envolvidas, tipo de conhecimento e atores envolvidos.

O presente trabalho é teórico. Inicialmente, através da revisão da literatura são definidos os processos de GC e os de GP, além de identificadas as tecnologias da informação relacionadas com a GC e os atores envolvidos no desenvolvimento de projetos de TI, assim como os tipos de conhecimento. Na sequência, é proposta a relação entre os processos de GC e de GP, considerando às tecnologias envolvidas, o tipo de conhecimento e atores envolvidos.

2 Gestão do conhecimento e gestão de projetos

Este capítulo está estruturado de forma a apresentar os conceitos e processos relacionados à GC e à GP.

2.1 Conhecimento: conceito e tipos

Embora dados e informações não sejam o foco do presente trabalho faz-se necessária a apresentação dos seus conceitos para tornar possível a distinção entre eles, e evidente o conceito de conhecimento. Dados podem ser considerados "números brutos, imagens, palavras e sons derivados de uma observação ou medição" [Nonaka e Poltokorpi 2006 p. 75]. Esses dados são classificados, sumarizados, transferidos ou corrigidos para que possam agregar valor e transformar-se em informação em um determinado contexto [Grover e Davenport 2001]. A informação é vista como uma *commodity* que causa relações de dependência com eventos distintos, e o conhecimento pode ser entendido como essa "informação analisada" [Drestke 1981]. Sendo assim, pode-se concluir que o conhecimento é "o crescimento de dados brutos e informações sistematicamente organizadas" [Hannabuss 2001 p. 357] e combinados à experiência, contexto e interpretação [Davenport et al. 1998]. O conhecimento é a maior contribuição humana, pois tem a maior relevância para decisões e ações uma vez que possui um maior nível de dependência em relação ao contexto [Grover e Davenport 2001].

Nonaka [1994] explica que o conhecimento pode ser dividido em dois tipos, sendo eles: tácito e explícito. O conhecimento tácito é "difícil de expressar e não pode ser documentado através de linguagem formal" [Tserng e Lin 2004 p. 782]. Ele consiste em conhecimento técnico e

habilidades e é conhecido como know-how [Oliveira et al. 2011]. O conhecimento explícito pode ser representado (através de documentos, por exemplo), armazenado (em base de dados, por exemplo) e efetivamente aplicado [Nonaka e Takeuchi 1995]. Ele pode ser "codificado e expresso através da linguagem formal" [Tserng e Lin LIN 2004 p. 782]. Os conhecimentos tácito e explícito não são excludentes, mas complementares [Nonaka e Takeuchi 1995].

2.2 Processos de gestão do conhecimento

A GC pode ser vista como uma série de processos que ocorrem nas organizações [Bagnoli e Roberts 2013], entretanto estes processos são denominados de diferentes formas. Para a realização do presente trabalho fez-se necessária a consolidação dos processos de GC (PGC) nas organizações. O Quadro 1 apresenta os processos identificados por diferentes autores.

	Processo de Gestão do Conhecimento nas Organizações				
Autores que definem PGC	**Criação**	**Armazenamento**	**Comaprtilhamento**	**Uso**	**Controle**
ALBINO; CLAUDIO; SCHIUMA, 1999	Desenvolvimento	Armazenar	Transferir	Uso	Controle
ALAVI; LEIDNER, 1999	Criação, Construção	Armazenar, Reter	Distribuir, Transferir	Aplicação	Manutenção, Gerenciamento
LEE; YANG, 2000	Adquirir	Armazenar	Disseminar	Integração	Gerenciamento; Inovação; Proteção
TSERNG; LIN, 2004	Criar	Capturar, Reter	Distribuir	Combinar	Coordenar, Proteger
COAKES, 2006	Criação	Armazenar, Capturar	Compartilhar, Acesso, Distribuir	Organizar, Aplicar	Gerenciamento
MASSEY; MONTOIA-WEISS, 2006	Criando	Armazenar, Capturar	Trocar	Aplicação	Gerenciamento
SHARMA; FOO; MORALES-ARROYO, 2008	Criar	Armazenar, Capturar, Buscar	Transferir	Re-uso	Gerenciamento, Organização
CHOO; NETO, 2010	Criação	Armazenar	Compartilhar, Acesso, Transferir	Uso	Gerenciamento
GOLDONI, V.; OLIVEIRA, M., 2010	Criação	Armazenar	Disseminar	Utilização	Mensuração
SHANG, *et al.* 2011	Criando; Adquirir	Armazenar	Trocar, Comaprtilhar	NA	Controle
KRAAIJENBRINK, 2012	Criação	Reter	Integrar	Aplicação	Manutenção
JANET; WEE; CHUA, 2013	Criação	Armazenar	Comaprtilhamento	Re-uso, Ulitilização	Gerenciamento
FERENHOF; FORCELLINI; VARAKIS, 2013	Criação, Identificar	Capturar	Compartilhar	Aplicação, Uso	Mensuração
Karkoulian et. al., 2013	Adquirir, Criar, Obter	Capturar	Compartilhar	Aplicação	Gerenciamento
NAICKER, 2013	Criar	Capturar	Transformar	Disseminated	Mensuração
NEJATIAN et al., 2013	Criar, Desenvolver	Armazenar	Compartilhar, Socializar, Transferir	Uso	Proteção, Sistema de GC
BURNETT; WILLIAMS; GRINNALL, 2013	Criar, Identificar, Capturar	Armazenar, Reter	Compartilhar, Transferir	Desenvolvimento	Auditar, Mapear, Organizar
MUELLER, 2014	Criação	Recuperar	Compartilhar	NA	Gerenciamento, Integração

Quadro 1 - Processos da GC (PGC).

Após analisar os processos apresentados no Quadro 1, foram adotadas neste trabalho as seguintes nomenclaturas: criação, armazenamento, compartilhamento, uso e controle.

2.2.1 Criação do conhecimento

A criação do conhecimento tem por objetivo criar valor ao negócio [Coakes 2006], pois conduz a soluções inovadoras para atender às necessidades dos clientes [Janet et al. 2013]. A criação do conhecimento é um processo contínuo e influenciado por algumas variáveis atuantes no ambiente organizacional, que podem ampliar ou inibir a criação do conhecimento [Nejatian et al., 2013]: Cultura Organizacional - sistema onde os valores são compartilhados e que guia o comportamento dos funcionários; Tecnologia da Informação - funcionalidades suportadas pela tecnologia adotada pela empresa; Pessoas - colaboradores e habilidades de cada um; Estrutura organizacional - refere-se ao grau de centralização e formalização adotado por cada organização.

A inovação é fortemente influenciada pela capacidade de criação do conhecimento, criatividade e habilidade de aprender dos colaboradores, a partir da indústria [Wee e Chua 2013]. Este é o primeiro processo do GC e possibilita que todos os outros ocorram (independentemente da fonte da criação), assim como todas as atividades relacionadas à GC, como adquirir, documentar, transferir e aplicar o conhecimento [Coakes 2006]. Fundamentalmente este é o processo que cria e adquire novos conhecimentos [Kraaijenbrink 2012], ou atualiza o conhecimento já existente [Alavi e Leidner 2001] perante um determinado contexto. Esse conhecimento pode ser novo para a organização, mas não necessariamente para toda a indústria, porém o que pode diferi-lo em um contexto específico são as relações que os indivíduos fazem com conhecimentos já existentes e a forma como este pode ser usado.

2.2.2 Armazenamento do conhecimento

A economia direcionada ao conhecimento demanda, cada vez mais, o armazenamento de forma organizada para possibilitar sua utilização. "*Knowing what we know"* é um "cliché" que resume a necessidade de um armazenamento eficaz, pois apenas é possível fazer o uso do conhecimento explícito no momento adequado quando este é armazenado e organizado de forma apropriada e acessível [Sharma et al. 2008], diferentemente do conhecimento tácito que já está arraigado no colaborador. Por este motivo, o processo de armazenamento e extração do conhecimento devem ser compreendidos para que o conhecimento não fique restrito a alguns grupos de profissionais, principalmente em multinacionais devido a dificuldades temporais e espaciais – os bancos de

dados são uma forma de armazenamento e a sua padronização é uma alternativa para facilitar a recuperação do conhecimento neste armazenado [Coakes 2006].

Capturar e armazenar o conhecimento resume-se ao processo de "codificar, armazenar e prover acesso eficaz a este", sendo as ferramentas de TI facilitadoras para tal, como ferramenta (não o fim em si) [Choo e Neto 2010 p. 602]. As ferramentas de TI devem complementar e alavancar as atividades relacionadas à GC disseminadas nas organizações de forma individual e coletiva, tendo papel facilitador que só é possível alcançar uma vez que estas ferramentas estejam alinhadas a natureza do conhecimento e de forma estratégica às práticas de GC da organização. O conhecimento armazenado pode ser considerado a memória organizacional uma vez que consiste no conhecimento humano codificado e armazenado [Alavi e Leidner 2001].

2.2.3 Compartilhamento do conhecimento

O compartilhamento do conhecimento é uma atividade que ocorre de forma corriqueira nas organizações [McDermott e Odell 2001], e é fundamental para possibilitar inovação e aprendizado organizacional, assim como para o desenvolvimento de novas habilidades que, consequentemente, podem resultar no aumento de produtividade [Mueller 2014]. O compartilhamento do conhecimento pode ser visto como um elemento fundamental para que as organizações possam manter-se competitivas no mercado [Jasimuddin 2007], pois possibilita que o conhecimento seja compartilhado para onde se faz necessário e deve ser usado [Alavi e Leidner 2001].

No presente trabalho o compartilhamento do conhecimento é considerado sinônimo de transferência e disseminação, pois pode ser entendido como um processo de comunicação envolvendo atores que usam os mecanismos apropriados para tal [Albino et al. 1999]. O compartilhamento ocorre de forma independente aos papéis, sobrepõem funções e é facilitado quando há proximidade física no ambiente organizacional [Janet et al. 2013]. Quando não há proximidade física, a tecnologia pode auxiliar no compartilhamento do conhecimento, através da criação de um ambiente "virtualizado" [Coakes 2006].

O compartilhamento do conhecimento permite que o conhecimento tácito seja reutilizado em outros projetos (ou por outros profissionais), possibilitando a otimização das operações [Tserng e Lin 2004]. Além disso, o compartilhamento do conhecimento e sua assimilação por parte de outros colaboradores pode contribuir para aumentar o valor percebido pelos clientes [Shang et al. 2011] uma vez que possibilita o aprendizado organizacional e os serviços / produtos prestados pelas organizações tendem a seguir um padrão.

2.2.4 Uso do conhecimento

O conhecimento possibilita o aprendizado organizacional, agregando valor ao negócio a partir do seu uso por parte de seus colaboradores [Mueller 2014], impulsionando o aprendizado organizacional e, consequentemente, a performance das organizações [Karkoulian et al. 2013]. O uso do conhecimento previamente adquirido pela organização traz benefícios, sendo alguns citados abaixo [Tserng e Lin 2004]:

- Reduz o tempo da solução dos problemas;
- Reduz o custo em relação à resolução de problemas;
- Melhora a qualidade da solução de problemas;
- Reduz a possibilidade de recorrência de problemas;
- Alavanca o conhecimento interno e externo, incluindo os *stakeholders*.

Sem o uso / re-uso do conhecimento, armazená-lo, por exemplo, teria pouco valor [Sharma et al. 2008], pois o uso do conhecimento é o que possibilita as vantagens competitivas da organização - não apenas o conhecimento em si [Alavi e Leidner 2001], sendo o uso o que justifica os processos de GC, apesar de este ser um processo difícil de isolar, pois as pessoas, muitas vezes, não percebem que estão efetivamente usando o conhecimento aprendido em algum momento anterior.

2.2.5 Controle do conhecimento

As empresas estão compreendendo a importância de uma gestão estruturada do conhecimento de forma gradativa e vêm introduzindo este conceito na cultura [Naicker 2013]. O controle do conhecimento assume diferentes perspetivas, que variam de acordo com o foco da pesquisa de cada autor. Porém, foi identificado que ele proporciona a GC das organizações perante aos processos nessas disseminados (como exemplo, pode-se citar o controle em relação às atividades de armazenamento do conhecimento para garantir que a sua aquisição, transformação, compartilhamento e aplicação foram realizados de forma efetiva [Shang et al., 2011]. Uma das formas de realizá-lo é através da criação de medições para os processos de GC, apesar de ser considerado um desafio medir ativos intangíveis, nos quais o conhecimento está incluso [Goldoni e Oliveira 2010]. Como exemplo de tais métricas pode-se citar o número de grupos de discussão relacionados a inovação de processos / produtos, o custo da disseminação do conhecimento e o Índice de Performance de GC (KMPI). Sendo assim, as autoras relacionam as métricas de GC à avaliação do resultado obtido em cada um de seus processos.

O controle do conhecimento organizacional deve ser a base para aprimorar os processos de trabalho, o que provavelmente vai refletir na *performance* da organização [Chan et al. 2012], além de observar as tecnologias disponíveis para suportá-la, que podem contribuir para manter a organização globalmente integrada [Burnett et al. 2013].

O processo de Controle garante que todos os outros processos de GC são realizados de forma eficaz.

2.3 Tecnologias que suportam a gestão do conhecimento

Os maiores níveis de complexidade introduzidos às organizações provenientes de um ambiente global e descentralizado [Dey et al. 2010] salientam a necessidade da criação de mecanismos para suportar os processos de GC. Estes mecanismos vão além de práticas de gestão, mas também abrangem as tecnologias necessárias para que os processos ocorram no ambiente organizacional [Oliveira et al. 2011].

Alguns autores classificam os mecanismos de GC como tecnológicos e não tecnológicos, pois se considera uma visão sistêmica onde os processos são geridos de forma integral e não apenas como foco nas ferramentas que os suportam [Bollinger et al. 2001]. Como ferramentas não tecnológicas abordam-se tópicos como iniciativas que promovam o conhecimento (ex: comunidades de práticas, reuniões informais e programas com mentores). Como ferramentas tecnológicas são mencionados os fóruns de discussões eletrônicas, e-mails, entre outros.

Para o presente trabalho abordar-se-á as ferramentas tecnológicas mais utilizadas para suportar a GC, conforme identificado na literatura [Oliveira et al. 2011]. A seguir serão listadas as ferramentas com as respectivas descrições:

1. Blog: é abordado na literatura como um diário eletrônico disponível na internet para acessos de terceiros.
2. *E-mails*: Ferramenta usada para enviar e receber mensagens eletrônicas.
3. Fóruns de discussões eletrônicas: considerado uma forma de comunicação onde às pessoas postam as mensagens desejadas e têm a possibilidade de comentar postagem de outros autores do fórum.
4. Sistemas inteligentes: softwares com a capacidade de prover respostas a problemas complexos.
5. Mensagens instantâneas: troca de mensagens por pessoas conectadas à internet.
6. Intranet: rede privada (geralmente adotada por empresas) que usa protocolos da internet.
7. Repositórios: sistemas de informação que possibilitam o armazenamento e o acesso ao conhecimento organizacional nestes documentados.

8. Programas de simulação: softwares destinados a simular situações reais
9. Telefone: ferramenta de telecomunicação que possibilita que pessoas de diversas localidades comuniquem-se.
10. Vídeo-conferências: ferramenta de telecomunicação que possibilita que pessoas de diversas localidades comuniquem-se através de áudio e vídeo.
11. Mensagens de voz: tecnologia que permite que uma mensagem gravada seja recebida, possibilitando um retorno posterior;
12. *Wiki:* tem a funcionalidade de armazenar o conhecimento construído por diversos indivíduos (através do registro das suas contribuições).
13. *Yellow Pages* / Diretórios: repositórios que possibilitam armazenar o conhecimento dos colaboradores de determinada organização.

Essas tecnologias são usadas de forma corriqueiras em um ambiente de projetos uma vez que estes o uso intenso do conhecimento [Reich 2007].

2.4 Gestão de projetos

Projeto é "um esforço temporário empreendido para criar um produto, serviço ou resultado exclusivo" [PMBOK, 2008]. Por suas características, os projetos têm sido parte das estratégias organizacionais na busca pela competitividade, pois eles permitem a gestão eficaz dos recursos organizacionais e, consequentemente, minimização dos custos, assim como gestão de riscos, o que aumenta as chances de sucesso [Juliani et al., 2012]. Eles podem ser considerados empreendimentos complexos, que envolvem múltiplos desafios para as organizações [Jugdev 2012] como, por exemplo, lidar com a diversidade dos múltiplos atores envolvidos, sendo este um fator crítico de sucesso [Lampel 2001].

Os projetos têm uma duração pré-determinada e geralmente possuem alguns marcos a serem alcançados e, na maioria das vezes, estão estruturados em diferentes fases, que requerem que o conhecimento e aprendizado gerado durante essas fases sejam capturados por parte do time de projeto e difundido de forma a atender as necessidades de cada fase [Wu e Passerini 2013]. Existem algumas ferramentas utilizadas pelas organizações que possibilitam a gestão da base de conhecimentos (como por exemplo: blogs, wikis e redes sociais). Essas proporcionam o armazenamento e distribuição do conhecimento de forma efetiva [Juliani et al. 2012], contribuindo para o aprendizado do time de projeto, que implica em melhorias nos grupos de planejamento, execução, monitoramento e controle [Ferenhof et al. 2013], ou seja, contribuem desde a iniciação até o seu fechamento. Nas seções seguintes, serão abordados os processos.

2.5 Processos de gestão de projetos

Os autores divergem em relação à nomenclatura e processos de GP. Para a realização do presente trabalho, fez-se necessária a consolidação desses conceitos. Este modelo sintetiza os processos de GP nas organizações. Optou-se por manter os grupos de processos descritos pelo PMI (*Project Management Institute*) uma vez que este é um instituto mundialmente reconhecido pela padronização das melhores práticas em GP e certifica profissionais na área [PMI, 2015]. Cada grupo possui um conjunto de processos, os quais são sintetizados no Quadro 2. Outras referências foram consultadas, através de uma busca estruturada, de forma a possibilitar uma visão mais ampla, e permitir uma análise crítica em relação as opiniões diferentes sobre o mesmo assunto. Como resultado de tal estudo, constituiu-se o Quadro 2 para que as nomenclaturas suportadas por cada autor se tornem explícitas de forma a facilitar a identificação das descrições utilizadas em cada fase e processo.

	Processo de Gestão de Projetos				
Autores que definem PGC	**Iniciação**	**Planejamento**	**Monitoramento e Controle**	**Execução**	**Fechamento**
LAMPEL, 2001	*Design*	Construção	Monitoramento	Execução	NA
MOTAWA et al., 2006 et. al., 2006	Preparação	Plano	Planejamento e Controle	Execução	NA
PMBOK, 2008	Iniciação	Planejamento	Monitoramento e Controle	Execução	Fechamento
DEY; CLEGG; BENNETT, 2010	NA	Plano	NA	Execução, Implementação	Pós *go-live* / operação
PEREIRA, 2011	Iniciação	Planejamento	Monitoramento e Controle	Execução	Fechamento
JUGDEV, 2012	Iniciação	Planejamento	Monitoramento e controle	Executar	Fechamento
TURSKIS; GAJZLER; DZIADOSZ, 2012 et. al., 2012	NA	NA	NA (foco em gestão de riscos apenas)	NA	NA
FERENHOF; FORCELLINI; VARAKIS, 2013	NA	Planejamento	Monitoramento e controle	Execução	NA
LECK, 2014	Preparação	*Blueprint*	NA	Realização; Preparação para a fase de *go-live*	Go Live, Fase de Suporte

Quadro 2 - Processos de GP (PGP).

2.5.1 Iniciação

Este grupo contem diversos processos, os quais objetivam definir o novo projeto assim como obter autorização formal para a sua execução. Os processos que suportam esse grupo resultam na definição inicial do escopo assim como no comprometimento do recurso financeiro necessário para suportar todas as atividades do projeto, e identificar todas as partes interessadas,

incluindo o gerente do projeto, que é nomeado nesta fase de forma oficial [PMBOK, 2008]. Este grupo resulta na definição do contrato do projeto, que contem a confirmação em relação ao escopo, orçamento e cronograma, servindo como base para um detalhamento posterior (durante o planejamento) [Leck 2014].

Os processos que estão estruturados como parte deste grupo são descritos a seguir, no Quadro 3.

Iniciação	
Processos	**Descrições**
Desenvolvimento do termo de Abertura	Consiste no desenvolvimento de um documento formal que autoriza a iniciação de um novo projeto ou de uma fase, além de registrar o escopo inicial e as expectativas das partes interessadas em relação ao projeto.
Identificar as partes interessadas	Consiste em identificar todas as pessoas que, de algum forma, possam ter alguma relação (seja ela positiva ou negativa) com o projeto, além de documentar as informações relevantes em relação ao seu interesse pelo projeto.

Quadro 3 - Processos do Grupo de Iniciação

2.5.2 Planejamento

O grupo de processos de planejamento tem como objetivo estabelecer o escopo do projeto, definir e detalhar os objetivos e desenvolver o plano de ação necessário para atingi-los através do desenvolvimento do plano de gerenciamento do projeto, assim como todos os documentos auxiliares necessários para a execução [PMBOK, 2008]. Durante o planejamento é realizada uma análise detalhada dos processos de negócio e dos requisitos para atendê-los durante o projeto. Em projetos de software, por exemplo, é a fase em que é determinado como os processos e requisitos serão refletidos no sistema [Leck 2014]. Existe uma crescente evidência de que problemas de planejamento são os principais causadores de falhas na implementação de sistemas como, por exemplo, o *Enterprise Resource Planning* [Dey et al. 2010]. Os processos do grupo de planejamento têm por objetivo definir de forma detalhada o escopo final a ser entregue pelo projeto assim como todos os seus objetivos e a forma como se pretende alcançá-los. Para tal, faz-se necessário desenvolver o plano de gestão do projeto em questão assim como a documentação necessária para a sua execução.

Os processos que estão estruturados como parte deste grupo são descritos a seguir, no Quadro 4.

Planejamento	
Processos	**Descrições**
Desenvolver o Plano de Gestão do Projeto	Consiste em dicumentar todas as "ações necessárias para definir, preparar, integrar e coordenar todos os planos auxiliares".
Coletar os Requisitos	Consiste na definição e na documentação adequada das solicitações provenientes das partes interessadas, ou seja, no produto, serviço ou resultado do projeto.
Definir o Escopo	Consiste em descrever, de forma detalhada, o produto e o projeto, de forma a possibilitar o entendimento comum entre o time do projeto e todas as outras partes interessadas.
Criar a estrutura Analítica do projeto (EAP)	Consiste em dividir as atividades em menores pacotes de trabalho a fim de facilitar o seu gerenciamento.
Definir as Atividades	Consiste em identicar as ações necessárias para realizar o trabalho do projeto.
Sequenciar as atividades	Consiste em organizar as dependências entre as atividades e ordená-las de acordo.
Estimar os Recursos das Atividades	Consiste em estimar todos os recursos necessários para completar as atividades.
Estimar as Durações das Atividades	Consiste em estimar o tempo necessário para completar as atividades.
Desenvolver o Cronograma	Consiste em criar o cronograma do projeto com base nas atividades identificadas, duração e dependências dessas.
Estimar os Custos	Consiste em estimar os recursos monetários necessários para completar o trabalho do projeto.
Determinar o Orçamento	Consiste na consolidação de todos os riscos associados ao projeto de forma a criar o *baseline* do projeto.
Planejar a Qualidade	Consiste em identificar o nível de qualidade necessário para o trabalho do projeto (sendo ele produto ou serviço).
Desenvolver o plano de Recursos Humanos	Consiste em identificar e documentar as necessidades em relação ao pessoal do projeto.
Planejar as Comunicações	Consiste em identificar as necessidades de informação e na criação de um plano de comunicações que atenda a essas necessidades.
Planejar o Gerenciamento dos Riscos	Consiste em definir como os riscos serão conduzidos durante o projeto.
Identificar os Riscos	Consiste em identificar os riscos que podem afetar o projeto.
Realizar a Análise Qualitativa dos Riscos	Consiste na priorização dos riscos para análise adicional.
Realizar a Análise Quantitativa dos Riscos	Consiste em analisar numericamente os riscos em relação ao seu efeito nos objetivos do projeto.
Planejar as Respostas aos Riscos	Consiste na definição de ações para serem implementadas a fim de minimizar os riscos do projeto.
Planejar as Aquisições	Consiste na documentação das decisões relacionadas a compras realizadas durane o projeto.

Quadro 4 - Processos do Grupo de Planejamento.

2.5.3 Execução

"A Execução consiste nos processos realizados para concluir o trabalho definido no plano de gerenciamento do projeto" a fim de garantir que o escopo seja entregue da forma especificada [PMBOK, 2008]. Neste grupo de processos o produto do projeto é realizado de acordo com os requisitos estabelecidos durante a fase de planejamento, além de ser preparado de forma a possibilitar seu uso (por exemplo, em projetos de software devem ser realizadas as atividades relacionadas com a preparação para o software entrar em produção) [Leck 2014]. Existem lacunas durante a execução de alguns projetos, muitas vezes relacionadas a recorrência de problemas já identificados em projetos similares executados anteriormente, o que resulta em perda de recursos para a organização [Juliani et al. 2012].

Os processos que estão estruturados como parte deste grupo são descritos a seguir, no Quadro 5.

Execução	
Processos	**Descrições**
Orientar e Gerenciar a Execução do Projeto	Consiste em executar o trabalho planejado para o projeto.
Realizar a Garantia da Qualidade	Consiste em auditar o trabalho do projeto para verificar se está de acordo com as especificções iniciais.
Mobiliar a Equipe do Projeto	Consiste em garantir a alocação dos recursos humanos necesários para realizar o trabalho do projeto.
Desenvolver a Equipe do Projeto	Consiste na melhoria da interação entre os membros da equipe a fim de aumentar a coesão entre eles e criar um ambiente de trabalho que possibilite aprimorar o desempenho do projeto.
Gerenciar a Equipe do Projeto	Consiste no acompanhamento e desenvolvimento profissional da equipe do projeto.
Distribuir as Informações	Consiste em disponibilizarda forma planejada as informações do projeto.
Gerenciar as expectativas das Partes Interessadas	Consiste na interação e comunicação junto às partes interessadas a fim de atendê-las conforme necesário.
Conduzir as Aquisições	Consiste na conceção do contrato.

Quadro 5 - Processos do Grupo de Execução.

2.5.4 Monitoramento e controle

O grupo de processos de monitorando e controle "consiste nos processos necessários para acompanhar, revisar e regular o progresso e o desempenho do projeto, identificar todas as áreas nas quais serão necessárias mudanças no plano e iniciar as mudanças correspondentes" [PMBOK, 2008]. As mudanças podem ser demandadas de diferentes fontes e devem ser analisadas a fim de que seu impacto seja adequadamente incorporado às demandas do projeto e as ações pertinentes sejam realizadas (sejam estas relacionadas à alteração do orçamento, cronograma, escopo, etc.) [Motawa et al. 2006]. Além das atividades inerentes ao controle do projeto em si, há também a necessidade da gestão dos riscos que o permeiam, que podem estar relacionados, por exemplo, aos prazos acordados e ao orçamento disponibilizado, entre outros fatores importantes para as partes interessadas do projeto [Turkis et al. 2012]. Existem diversas formas de realizar a gestão de riscos, porém uma delas é o *framework* abaixo, que inclui cinco etapas [Dey et al. 2010]:

1. **Identificar o risco:** o risco pode ser identificado como um problema que possa vir a ocorrer no futuro, com potencial impacto no trabalho do projeto e com uma probabilidade de ocorrência;
2. **Registrar o risco:** o risco deve ser devidamente registrado, e as informações relativas a este devem ser apropriadamente compartilhadas nos diferentes níveis organizacionais;
3. **Revisar o risco:** Os riscos devem ser revisados pelas partes interessadas apropriadas a fim de aceitar-se a análise realizada;

4. **Gestão do risco:** O impacto dos riscos devem ser determinados juntamente às partes interessadas e às ações de mitigação adotadas para tal;

5. **Fechamento do risco:** Os riscos devem ser acompanhados, fechados quando apropriado e comunicado às partes interessadas.

Os processos que estão estruturados como parte deste grupo são descritos a seguir, no Quadro 6.

Monitoramento e Controle	
Processos	**Descrições**
Monitorar e Controlar o Trabalho do Projeto	Consiste na monitoração do progresso do trabalho que está sendo realizado a fim de garantir que este está de acordo com os requisitos pré-estabelecidos.
Realizar o Controle Integrado de Mudanças	Consiste em avaliar e implementar as solicitações de mudanças, quando aprovadas.
Verficar o Escopo	Consiste na validação do trabalho do projeto e sua formal aceitação.
Controlar o Escopo	Consiste na gestão do escopo do projeto.
Controlar o Cronograma	Consiste em monitorar o andamento do projeto.
Controlar os Custos	Consite em monitoraro andamento do projeto para que os custos sejam atualizados de acordo com o trabalho planejado.
Realizar o Controle da Qualidade	Consiste no monitoramento e documentação do resultado e desempenho das atividades realizadas.
Reportar o Desempenho	Consiste em comunicar informações referentes ao desempenho do projeto.
Monitorar e Controlar os Riscos	Consiste ma monitoração dos riscos identificaos e em prover respostas a esses.
Administrar as Aquisições	Consiste em monitorar as aquisições realizada o projeto.

Quadro 6 - Processos do Grupo de Monitoramento e Controle.

2.5.5 Fechamento

O grupo de processos de fechamento consistem no encerramento de todas as atividades de todos os grupos de processos que foram realizadas durante o projeto [PMBOK, 2008]. Em projetos de software, por exemplo, é neste grupo que o sistema entra em produção e o suporte pós-implementação é realizado [Leck 2014].

É importante que sejam documentadas as lições aprendidas durante o projeto, pois elas são frutos da experiência vivida e provavelmente irão conter um conhecimento relevante para projetos futuros [Schindler e Eppler 2003 apud Ferenhof et al. 2013]. Elas são uma forma de registrar e compartilhar o conhecimento formal e informal aprendido durante o projeto [Jugdev 2012].

Os processos que estão estruturados como parte deste grupo são descritos a seguir, no Quadro 7.

Fechamento	
Processos	**Descrições**
Encerrar as atividades do projeto	Consiste na finalização de todas as atividades do projeto, inclusive de suas documentações.
Encerrar as Aquisições	Consiste em finalizar as aquisições realizadas em função do projeto.

Quadro 7 - Processos do Grupo de Fechamento.

3 Relação entre os processos de gestão do conhecimento e de gestão de projetos

Inicialmente, nesta seção, apresenta-se um quadro que resume os conceitos dos processos de GC (Quadro 8).

Processo	Definição
Criação	Cria valor ao negócio pois possibilita soluções invadoras. É um processo contínuo e influenciado por diversas variáveis do ambiente organizacional. Este processo cria conhecimento novo ou atualiza conhecimento existente.
Armazenamento	Consiste em guardar o conhecimento e armazená-lo de forma organizada, que possibilite que o conhecimento seja recuperado conforme necessário. O conhecimento é codificado, armazenado e disponibilizado de forma eficaz.
Compartilhamento	Atividade que ocorre de forma corriqueira nas organizações. O compartilhamento possibilita inovação e aprendizagem organizacional e é fundamental para que as organizações possam manter-se competitivas no mercado. O compartilhamento do conhecimento é considerado sinônimo de trabsferência e disseminação.
Uso	O uso do conhecimento consiste na sua aplicação - ele agrega valor ao negócio e impulsiona o aprendizado organizacional.
Controle	Proporciona a gestão do conhecimento das organizações perante os processos disseminados. Está relacionado a indicadores, métrcas e auditorias.

Quadro 8 - Consolidação dos Processos de GC.

Na sequência, apresenta-se um quadro que resume os processos de GP (Quadro 9).

Processo	Definição
Iniciação	Definições iniciais do projeto: é realizado um contrato, que contém o escopo, orçamento e cronograma.
Planejamento	Estabelecimento do escopo do projeto, definição e detalhamento dos objetvos e desenvolvimento do plano de ação necessário para atingí-los. Além disso são desenvolvidos os planos que auxiliam a execução do projeto.
Monitoramento e Controle	Processos necessários para acompanhar, revisar e regular o progresso e o desempenho do projeto, identificar todas as áreas nas quais serão necessárias mudanças no plano e iniciar as mudanças correspondentes.
Execução	Processos realizados para concluir o trabalho definido no plano de gerenciamento do projeto a fim de que o escopo seja entregue da forma especificada.
Fechamento	Encerramento de todas as atividades de todos os gurupos de processos que foram realizadas durante o projeto

Quadro 9 - Consolidação dos Processos de GP.

A relação entre os processos de GC e de GP serão investigados nesta pesquisa sob três perspetivas: tipo de conhecimento gerado através dessa interação, tecnologias e atores

envolvidos (ver Quadros 10, 11 e 12). A partir da revisão da literatura e do conhecimento tácito da autora identifiou-se possíveis relações entre os processos sob as 3 perspetivas propostas. Pretende-se, em uma pesquisa futura, testar-se empiricamente essas relações a fim de validá-las e enriquecê-las com base em estudos de casos.

Relação entre os processos em termos de Tipo de Conhecimento		Processo de Gestão do Conhecimento nas Organizações				
		Criação	Armazenamento	Compartilhamento	Uso	Controle
Processo de Gestão de Projetos	Iniciação	Criação de conhecimento em torno de soluções inovadoras e que agregam valor ao negócio podem suportar a criação de novos contratos e soluções que necessitem de projetos para a implementação.	O conhecimento gerado a partir das interações iniciais deve ser devidamente armazenado de acordo com as políticas da organização e necessidades legais.	O conhecimento gerado a partir das interações iniciais deve ser compartilhado de acordo com as necessidades do time de projetos e obrigações contratuais.	O conhecimento gerado a partir das interações iniciais deve ser usado pelo time de projeto de forma a possibilitar o planejamento posterior.	O conhecimento inicial (explícito em forma de documentações) deve ser controlado no nível organizacional de forma a garantir que o projeto esteja em conformidade com as normas.
	Planejamento	Potencialmente como parte do exercício de planejamento do projeto, novos conhecimentos serão criados no time de projetos.	Conhecimento criado a partir do planejamento do projeto deve ser armazenado de forma a possibilitar acesso futuro.	Conhecimento criado a partir do planejamento do projeto deve ser compartilhado com as partes interessadas.	O conhecimento gerado a partir das interações iniciais deve ser usado pelo time de projeto de forma a combiná-lo com o conhecimento prévio dos indivíduos a fim de que o planejamento eficaz das atividades seja realizado.	Conhecimento em relação ao planejamento deve ser devidamente controlado, seguindo as normas da organização. O cronograma, por exemplo, pode ser um dos itens a ser controlado.
	Monitoramento e Controle	O processo de monitramento e controle pode gerar novos conhecimentos em relação a tendências de projetos, riscos, etc.	O conhecimento proveniente das atividades de monitoramento e controle devem ser adequadamente armazenados pelos times e/ou gestores.	O conhecimento proveniente dos processos de monitoramento e controle devem ser comapritlhados com os stakeholders de acordo com a sua necessidade e normas organizacionais (assim como obrigações legais).	1) Usa-se o conhecimento gerado através dos processos de monitoramento e controle para a tomada de decisão em relação as ações necessárias para a gestão do projeto. 2) Não foi identificada alguma tecnologia específica 3) Todo o time de projeto, partes interessadas	O conhecimento proveniente das atividades de monitoramente e controle podem relatar os status dos projetos, os quais devem ser controlados pela organização (essa prática inclui múltiplos projetos e pode ser controlada no nível do portfólio).
	Execução	Durante a execução do projeto ocorre a criação de novos conhecimentos provenientes das interações dos times de projetos frente às situações adversas.	O conhecimento produzido durante a execução do projeto deve ser armazenado de forma a suportar as necessidades dos *stakeholders*.	Durante a execução o conhecimento deve ser compartilhado entre o time de projeto de forma a garantir alinhamento dos membros do time de projeto (conhecimento em relação ao produto) mas também ao time de gestão para possibilitar acompanhamento da evolução.	O conhecimento adquirido deve ser usado a fim de completar o trabalho do projeto.	O conhecimento gerado durante o trabalho do projeto deve ser controlado de acordo com as normas e padrões (inclui controle de acesso ao conhecimento, restrições legais do projeto, etc).
	Fechamento	Durante o fechamento do projeto torna-se explícito o conhecimento criado durante o projeto proveniente do trabalho neste executado.	O conhecimento gerado durante o projeto deve ser armazenado, seguindo as normas da organização.	Durante o fechamento do projeto ocorre o compartilhamento do conhecimento adquirido durante sua execução pelo time de projeto.	O conhecimento adquirido durante o projeto pode ser usado para treinar os clientes a forma como o produto do projeto deve ser usado / gerido.	Durante o fechamento do projeto são realizadas auditorias para garantir que toda a documentação do conhecimentos foi realizada, conforme previsto pelas normas da organização.

Quadro 10 - Relação entre os processos

A seguir, o quadro que representa a relação entre os processos de GC e de de GP em relação às tecnologias envolvidas.

Relação entre os processos em termos de Tecnologia		Processo de Gestão do Conhecimento nas Organizações				
		Criação	Armazenamento	Comaprtilhamento	Uso	Controle
Processo de Gestão de Projetos	Iniciação	Sistemas Inteligentes; Mensagens Instantâneas, Telefone, Video-conferências	Repositórios, Intranet, Yellow Pages	E-mail, Mensagens Instantâneas, Intranet, Repositórios, Telefone, Video-conferência, Mensagem de voz, Wikis	Blog, Fórum de discussão eltrônica	Repositórios, Intranet, Yellow Pages
	Planejamento	Sistemas Inteligentes; Mensagens Instantâneas, Telefone, Video-conferências, Wikis, Yellow Pages	Intranet, Repositórios, Wiki, Yellow Pages	E-mail, Mensagens Instantâneas, Repositórios, Telefone, Video-conferência	Blog, Fórum de discussão eltrônica	Repositórios, Intranet, Yellow Pages
	Execução	Blog, E-mail, Mensagens Instantâneas, Telefone, Video-conferência, Mensagens de Voz	Intranet, Repositórios, Wiki, Yellow Pages	E-mail, Mensagens Instantâneas, Repositórios, Telefone, Video-conferência	Blog, Fórum de discussão eltrônica	Intranet, Repositórios
	Monitoramento e Controle	E-mail, Mensagens Instantâneas, Telefone, Video-conferência, Mensagens de Voz	Intranet, Repositórios, Wiki, Yellow Pages	Emails, intranet, repositórios, telefone, video-conferências	Blog, Fórum de discussão eltrônica	Intranet, Repositórios
	Fechamento	Telefone, Video-conferência	Intranet, Repositórios, Wiki	Emails, intranet, repositórios, telefone, video-conferências, wikis	Blog, Fórum de discussão eltrônica	Intranet, Repositórios, Wikis

Quadro 11 - Relação entre os processos de GC e de GP em relação às tecnologias envolvidas.

A seguir, o quadro que representa a relação entre os processos de GC e de de GP em relação aos atores envolvidos.

Relação entre os processos em termos de Atores		Processo de Gestão do Conhecimento nas Organizações				
		Criação	Armazenamento	Comaprtilhamento	Uso	Controle
Processo de Gestão de Projetos	Iniciação	Sponsor, GP, Cliente	Sponsor, GP, Cliente	Sponsor, Cliente	Sponsor, GP, Cliente	Sponsor, Membros de Auditorias, Gerente de Projetos, Clientes, Outros stakeholders
	Planejamento	Time de Projeto, Cliente, Outros Stakeholders	Time de Projeto, Cliente, Outros Stakeholders	Time de Projeto, Cliente, Outros Stakeholders	Time de Projeto, Cliente, Outros Stakeholders	Sponsor, Membros de Auditorias, Gerente de Projetos, Clientes, Outros stakeholders
	Execução	Time de Projeto, Outros Stakeholders	Time de Projeto, Cliente, Outros Stakeholders	Time de Projeto	Time de Projeto, Cliente, Outros Stakeholders	Sponsor, Membros de Auditorias, Gerente de Projetos, Clientes, Outros stakeholders
	Monitoramento e Controle	Sponsor, Gerente de Projetos, Clientes, Outros stakeholders	Sponsor, Gerente de Projetos, Clientes, Outros stakeholders	Sponsor, Gerente de Projetos, Clientes, Outros stakeholders	Sponsor, Gerente de Projetos, Clientes, Outros stakeholders	Sponsor, Membros de Auditorias, Gerente de Projetos, Clientes, Outros stakeholders
	Fechamento	Time de Projeto, Cliente, Outros Stakeholders	Time de Projeto, Cliente, Outros Stakeholders	Sponsor, Gerente de Projetos, Clientes, Outros stakeholders	Sponsor, Gerente de Projetos, Clientes, Outros stakeholders	Sponsor, Membros de Auditorias, Gerente de Projetos, Clientes, Outros stakeholders

Quadro 12 - Relação entre os processos de GC e de GP em relação aos atores envolvidos.

5 Conclusões

A partir do presente estudo concluiu-se que os projetos são atividades intensas em conhecimento e que este tem o potencial de alavancar os projetos de diferentes formas: suportando os processos que estes englobam e também em função de o conhecimento representar a capacidade de adaptação dos colaboradores das organizações a fim que eles possam adequar-se às constantes demandas dos projetos. Além disso, o trabalho contribui teoricamente para que se pudesse consolidar as diferentes perspetivas abordadas na literatura em relação aos processos de GC e de GP: percebeu-se que existem diversas nomenclaturas referentes aos mesmos conceitos e essa consolidação foi fundamental a fim de possibilitar a conceitualização dos processos de forma eficiente.

As relações entre os Processos de GC e de GP possibilitam um planejamento dos gestores para garantir que os processos de GC possam, de fato, alavancar os processos de GP nas organizações. No que tange aos atores envolvidos, o mapeamento possibilita que os gestores planejem a alocação dos seus recursos de forma otimizada pois é possível identificar os recursos que precisam estar disponíveis em cada interseção dos processos o que permite melhorar a alocação e evitar desperdícios. No que tange as tecnologias envolvidas, é possível identificar quais são as ferramentas necessárias aos colaboradores em cada uma das etapas a fim de que eles estejam munidos das tecnologias necessárias para otimizar o seu trabalho. Por fim, no que tange o tipo de conhecimento gerado a partir dessas interações é possível mapear as oportunidades que as empresas têm de criar novos conhecimentos ou atualizar conhecimento existente, formas de armazenamento organizado que permitam o resgate do conhecimento quando este se fizer necessário, identificar como o conhecimento é compartilhado no nível organizacional a fim de facilitar este processo e restringí-lo quando necessário, identificar a forma como os colaboradores de fato usam o conhecimento e, por fim, criar mecanismos de controle que permitam uma gestão eficaz do conhecimento durtante todos os processos de GP.

Agradecimentos

Os autores agradecem o suporte recebido do CNPq e da CAPES.

6 Referências

Abtahi, M. S.; Edrisi, F.; Mahmoudi, A.: *The Influence of Processes of Knowledge Management on Efficiency Improvement. International* Journal of Academic Research in Business and Social Sciences October, Vol. 2, No. 10, 2012.

Alavi, M.; Leidner, D. E.: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, v.25, n.1, p. 107-137, 2001.

Albino, V.; Claudio, G. A.; Schiuma, G.: Knowledge transfer and inter-firm relationships in industrial districts: The role of the leader firm. *Technovation,* 19, 53–63, 1999.

Bagnoli, C.; Roberts, H.: Governing strategy and knowledge: tools and methodologies Published online: 13 December 2011.

Bollinger, A. S.; Smith, R. D.: Managing organizational knowledge as a strategic asset. *Journal of Knowledge Management,* 5, 1 .8 - 18.2001

Burnetts, S.; Williams, D.; Grinnali, A.: The Strategic Role of Knowledge Auditing and Mapping: An Organisational Case Study; Knowledge and Process Management Volume 20 Number 3 pp 161–176 (2013)..

Chan, K. H.; Chu, S. K.W.; Wu, W. W. Y.: Exploring the Correlation Between Knowledge Management Maturity and Intellectual Capital Eficiency in Mainland Chinese Listed Companies *Journal of Information & Knowledge Management*, Vol. 11, No. 3,. 2012.

Choo, C. W.; Alvarenga, R. C. D. N..; Beyond the ba: managing enabling contexts in knowledge organizations. *Journal of Knowledge Management*. Vol. 14 no. 4 2010, pp. 592-610.

Coakes, E.: Storing and sharing knowledge: supporting the management of knowledge made explicit in transnational organisations. *The Learning Organization* Vol. 13 No. 6, 2006 pp. 579-593 q Emerald Group Publishing Limited. 0969-6474 DOI 10.1108/09696470610705460. 2006.

Davenport, T.H.; Long, D. W. W.; Beers, M. C.; Successful Knowledge Management Projects Sloan Management Review; Winter 1998; 39, 2; *Abi/Inform Global* .1998.

Davenport, T.H.; Prusak, L.: *Conhecimento empresarial: como as organizações gerenciam o seu capital intelectual,* Campus, Rio de Janeiro. 1998.

Dey, P. K.; Clegg, B. T.; Bennett, D.J.: Managing enterprise resource planning projects. *Business Process Management Journal* Vol. 16 No. 2, 2010 pp. 282-296.

Drestke, F: Knowledge and the Flow of Information. *MIT Press, 1981*. http://philpapers.org/rec/DREKAT, (10de maio de 2014), 1981.

Ferenhof, H.A., Forcellini, F. A., Varakis, G.: Lessons learned: adding value to the project management. *Revista de Gestão e Projetos - GeP*, São Paulo, v. 4, n. 3, p 197-209, set./dez. 2013.

Fitzek, D.: *Knowledge Management in Inter Project Learning – A systematic attempt of Integration***.** Alemanha, 2002.

Goldoni, V.; Oliveira, M.: *Indicadores para Gestão do Conhecimento na visão de Especialistas.* REAd, v. 13, n. 3. 2007.

Goldoni, V.; Oliveira, M. Knowledge management metrics in software development companies in Brazil. *Journal of Knowledge Management* v. 14 N. 2. pp. 301-313.2010.

Grant, R. M. Grant toward a knowledge-based theory of the firm. *Strategic Management Journal,* Vol. 17(Winter Special Issue), 109-122.1996.

Grover, V.; Davenport, T. H.: General perspectives on knowledge management: Fostering a research agenda. *Journal of Management Information Systems*.18, 1; ProQuest Central pg. 5. 2001.

Hannabuss, S.: A wider view of knowledge. *Library Management*; 22, 8/9. 2001.

Hansen, M. T.; Nohria, N., Tierney, T.: What is your strategy for managing knowledge. *Harvard Business review*, 77, 2 . 106 – 121.1999.

Jasimuddin, S.: Exploring knowledge transfer mechanisms: The case of a UK-based group within a high-tech global corporation. *International Journal of Information Management,* 27, 294–300.2007.

Jugdev, K.: Learning from Lessons Learned: Project Management Research Program. *American Journal of Economics and Business Administration* 4 (1): 13-22, 2012.

Juliani, D. P.; Juliani. J. P.; Bello, J. S. A.; Souza, J. A.: Modelo para construção de base de conhecimento sobre projetos suportado por ferramentas colaborativas. *Revista de Gestão e Projetos - GeP*, São Paulo, v. 3, n. 3, p 277-290, set./dez. 2012.

Karkoulian, S.; Messarra, L. C.; McCarthy, R.: The intriguing art of knowledge management and its relation to learning organizations. *Journal of Knowledge Managemet.* VOL. 17 NO. 4, pp. 511-526, 2013.

Kraaijenbrink, J.: Integrating knowledge and knowledge processes: A critical incident study of product development projects. *Product Development & Management Association*, 2012.

Lampel, J.: *The core competencies of effective project execution*: the challenge of diversity. University of Nottingham Business School, Jubilee Campus, Wollaton Road, Nottingham, 2001.

Lech P.: Managing knowledge in IT projects: a framework for enterprise system implementation. *Journal of Knowledge Managemrnt*. VOL. 18 NO. 3, pp. 551-573, 2014.

Lee, C. C.; Yang, J.: Knowledge value chain. *Journal of Management Development*, v. 19, n. 9, p. 783-793, 2000.

Lee, J-H.; Kim, Y-G.: *A stage model of organizational knowledge management*: a latent content analysis. Expert Systems with Applications 20. 299 – 311.2001.

McDermott, R., e O'Dell, C. Overcoming cultural barriers to sharing knowledge. *Journal of Knowledge Management,* 5. 76–85.2001.

Mitrevski, V.; Aceski, A.: Creation and transfer of knowledge management os sports institution organizations. *Indian Journal of Commerce & Management Studies* ISSN: 2240-0310 EISSN: 2229-5674. Volume VI Issue 1, Jan. 2015.

Motawa, I. A.; Anumba, C., J.; Lee, S., Peña-Mora, F.: An integrated system for change management in construction. *Automation in Construction* 16 (2007) 368–377, 2007.

Mueller, J. A specific knowledge culture: Cultural antecedents for knowledge: sharing between project teams. *European Management Journal* 32 . 190–202, 2014.

Naicker, V.: Uncovering Knowledge Management Practices In Organizations. *The Journal of Applied Business Research* – November/December 2013 Volume 29, Number 6 2013.

Nejatian, M.; Nejati, M.; Zarei, M. H.; Somaye, S.: Critical Enablers for Knowledge Creation Process: Synthesizing the Literature. *Global Business and Management Research: An International Journal V*ol. 5, Nos. 2 & 3 .2013.

Nonaka, I. A Dynamic Theory of Organizational Knowledge Creation. *The Institute of Management Sciences***.** v. 5, n. 1, p. 14-37, 1994.

Nonaka, I.; Takeuchi, H. *The Knowledge-Creating Company: how Japanese Companies Create the Dynamics of Inovation*. Nova York, Oxford University Press, 1995.

Nonaka, I.; Poltokorpi, V. *Objectivity and Subjectivity in Knowledge Management:* A Review of 20 Top Articles, Knowledge and Process Management. Volume 13 Number 2 pp 73–82, 2006.

Oliveira, M., Maçada, A., Curado, C. Knowledge management mechanisms within the SECI model, *CAPSI'2011*, , Lisboa, Portugal.2011.

Pereira, L.: *Gestão de Conhecimento em Projetos*. Escola Superior de Actividades Imobiliárias – FCA Editora de Informática . 2011.

PMBOK: Um guia em gerenciamento de projetos. *Project Management Institute*. 4a edição, 2008.

Pollack, J.: Transferring knowledge about knowledge management: Implementation of a complex organisational change programme. University of Technology, Sydney, Australia. *International Journal of Project Management* 30 . 877–886, 2012.

Reich, B.H.: Managing knowledge and learning in IT projects: a conceptual framework and guidelines for practice*", Project Management Journal*, Vol. 38 No. 2, pp. 5-17, 2007.

Reich, B.H.; Gemino, A.; Sauer, C.: Modelling the knowledge perspective of IT projects. *Project Management Journal*. Vol. 39 No. S1, pp. S4-S14, 2008.

Shang, S.S.C.; Li, E. Y.; Wu, Y.; Hou, O. C. L.: Understanding web. 2.0 service models: a Knowledge-creating perspective. *Information & Management*, 48. 2011.

Sharma, R. S.; Foo, S. M-A. M.: Developing Corporate Taxonomies for Knowledge Auditability – A framework for good practices *Knowledge Organisation*, 35(1), 30-46. 2008.

Teixeira, J.; Silva, R.; Lapa, E.: Os prjectos de implantação. In: Silva, R.V.; Neves, A. (Org.) *Gestão de empresas na era do conhecimento*. Serinews, São Paulo, 2004.

Tserng, H. P.; Lin, Y-C.: Developing an activity-based knowledge management system for contractors. *Automation in Construction* 13 (2004) 781– 802, 2004.

Turskis, Z.; Gajzler, M.; Dziadosz, A: Reliability, risk management and contingency of construction processes and projects. *Journal of Civil Engineering and Management,* Volume 18(2). 2012.

Wu, D.; Passerini, K.: Uncovering knowledge-based time management practices: Implications for project management. *International Journal of Managing Projects in Business* Vol. 6 No. 2, 2013 pp. 332-348, 2013.

Analysis of Enterprise Architecture Models with Description Logics Reasoning and SPARQL

Gonçalo Antunes [1], Artur Caetano [1, 2], José Borbinha [1, 2].

1) Instituto de Engenharia de Sistemas e Computadores – Investigação e Desenvolvimento em Lisboa, Portugal

goncalo.antunes@inesc-id.pt

2) Instituto Superior Técnico, Universidade de Lisboa, Portugal

{artur.caetano, jlb}@tecnico.ulisboa.pt

Abstract

Enterprise architecture (EA) model analysis can be defined as the application of property assessment criteria to EA models. Ontologies can be used to represent conceptual models, providing means for their integration, at the semantic and syntactic level, and allowing the application of computational inference to derive logical conclusions from the facts present in the models. As the commonly used EA modelling languages are conceptual, advantage can be taken of representing such conceptual models using ontologies. This paper explores and demonstrates the use of ontologies and associated techniques in the analysis of enterprise architecture models. Two techniques are used to this end: computational inference and the use of SPARQL. The aim is to demonstrate the possibilities brought by the use of these techniques in EA model analysis.

Keywords: Enterprise architecture, model analysis, ontologies, reasoning, SPARQL

1. Introduction

Enterprise architecture (EA) model analysis has a focus on the application of techniques that process the artefacts, properties and dependencies of a model, generating information that can be used to assess, transform or redesign organizational systems [Bucher et al. 2006], [Johnson, Lagerstrom et al. 2007]. Model analysis can be seen as the "application of property assessment to models" with the goal of observing the system's functional and non-functional qualities [Bertolino et al. 2013].

Ontologies enable representing the aspects of a conceptual model, which is composed of a conceptual schema and an information base [Olivé 2007]. Ontologies also provide the mechanisms to integrate different schemas through the definition of rules that relate the concepts at structural or semantic level. Moreover, with logic-based ontologies, inference can be used to derive logical conclusions from the ontological facts and support model analysis.

Enterprise architecture models expressed in the actual common EA languages can be expressed using ontologies. Hence, advantage can be taken of the features of ontologies, namely the integration and the computational inference mechanisms. Moreover, different techniques to apply to this end are widely available as part of the semantic web standards and frameworks. OWL [W3C 2012] can be used to represent ontologies, and the particular sublanguage of OWL-DL is in fact an implementation of description logics, which can be reasoned upon using one of the several reasoners available. Available semantic web frameworks such as the Apache Jena allow the storing and exposing of ontologies in a high-performance triple-store, with APIs for interfacing with other systems and reasoners, and SPARQL [W3C 2013] query facilities for accessing and manipulating the data contained in the ontologies.

This work explores the use of ontologies in the analysis of enterprise architecture models. In particular, two techniques are targeted to this end: the use of computational inference to reason upon the models and deduce implicit knowledge, and the use of SPARQL to retrieve and manipulate the data stored in the models. The use of these techniques is made on top of a federated model for the integration and analysis of EA models previously published in [removed for blind review]. In order to demonstrate the application of the techniques, a fictional scenario modelled using the ArchiMate language will be used.

This paper is organized as follows. Section 2 briefly describes related work in EA analysis, while Section 3 provides an overview in Ontologies, including description logics and OWL. Then, in Section 4, the referred ontology related techniques are demonstrated in performing different types of EA model analysis. Finally, the paper concludes in Section 5.

2. Enterprise Architecture Analysis

Enterprise architecture analysis assists the continuous enterprise architecture program by providing information to support the planning, improvement, and management processes [Hamalainen 2008]. It also supports the governance, optimization, re-engineering and decision-making processes of an organization [Hamalainen 2008], [Lankhorst 2005], [Johnson and Ekstedt 2007]. According to [Niemann 2005], EA analysis can be classified according to the following categories:

- **Dependency analysis** relates enterprise architecture artefacts to derive direct and indirect dependencies between them;
- **Coverage analysis** detects redundancies and gaps existing in an EA description, such as missing artefacts or missing relationships between them;

- **Interface analysis** assesses how the interfaces of artefacts relate usually with the goal of determining the degree of coupling and cohesion;
- **Heterogeneity analysis** is used for determining the identifies elements that need re-factoring as means to homogenize the overall architecture description;
- **Complexity analysis** is typically used for determining the number of elements existing in an EA description and the number of relationships they have;
- **Compliance analysis** determines if the artefacts or the overall architecture description meet policies, rules and requirements;
- **Cost analysis** calculates the costs of an artefact (e.g. creation, maintenance) or costs pertaining to the architecture description;
- **Benefit analysis** determines the contribution of an artefact to the overall goals of the organization as described in architecture.

We can find in the literature specific proposed approaches for this purpose. In [Johnson et al. 2004], the addition of non-functional attributes to enterprise architecture models through architecture theory diagrams (ATD) is proposed. These diagrams interrelate attributes through composition, correlation, and casual relationships. The ATD is then populated with measures that support the calculation of the model's non-functional attributes. Extended influence diagrams (EID) [Johnson, Lagerstrom et al. 2007] and probabilistic relational models (PRM) [Lagerstrom et al. 2010] are other analysis techniques. EID are used to model goals and decision alternatives, thus providing support for decision making. EID use probabilistic inference to support the representation of uncertainty when computing the values of attributes. PRM extend entity relationship models, and support model analysis under uncertainty.

In [Davoudi and Sheikhvand 2012], the analytical hierarchical process (AHP) is proposed to prioritize and select architectural scenarios according to the non-functional requirements under analysis. In the work reported in [Franke, Flores and Johnson 2009], fault tree analysis (FTA), an extended Bayesian network, is used to analyze dependencies related to reliability and reusability qualities. The use model mapping and transformation to analyze the compliance of process models against a set of actor coordination patterns is reported in [Caetano, Assis and Tribolet 2012]. In [de Boer et al. 2005], an XML schema is used to encode an enterprise architecture meta-model that specifies the structure and dynamics of enterprise architecture models. This approach analyses the structure of a model in terms of its cardinality, class specialization, and concept relationships. The dynamic analysis uses scenarios that encode state-based actions with XML and RML rules to simulate the behavior of the architecture.

3. Ontologies

An ontology is defined as a "formal, explicit specification of a shared conceptualization" [Studer, Benjamins and Fensel 1998]. According to [Guarino, Oberle and Staab 2009] and [Genesereth and Nilsson 1987], "conceptualization" refers to an "abstract, simplified view of the world", containing "the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them". In [Studer, Benjamins and Fensel 1998], "explicit" is related to the definition of the "type of concepts used, and the constraints on their use", while "formal" to the fact that the conceptualization "should be machine readable". Finally, "shared" means that the ontology "captures consensual knowledge".

Description logics (DL) are "a family of logic-based knowledge representation languages suitable for the representation of ontologies", which can be seen as "a decidable fragment of first-order logic" [Vaculin 2009]. DL describe domains in terms of concepts, roles, and individuals. Roles and concepts are related using logical statements named axioms. Different varieties of DL exist with differing degrees of expressiveness. According to [Areces 2000], DL supports five different types of reasoning: subsumption, instance checking, relation checking, concept consistency, and knowledge base consistency. Subsumption organizes concepts in a hierarchy and finds the most specific super-class for each class. Instance checking verifies if a given individual is an instance of a concept. Relation checking verifies if and how two individuals relate to each other. Concept consistency verifies if there are no contradictions between the definitions or the chain of definitions of a concept. Finally, knowledge base consistency determines whether the information contained in the knowledge base contains any contradictions.

The Web Ontology Language (OWL) is described in [W3C 2012] as a "semantic web language designed to represent rich and complex knowledge about things, groups of things, and relations between things". OWL can be used in a varying level of expressiveness, with a particular variant known as OWL-DL, being an implementation of DL which retains "computational completeness" [W3C 2004]. The terminology used in OWL is different from that used in DL: a DL concept corresponds to an OWL class, a DL role corresponds to an OWL property, and a DL individual corresponds to an OWL object. An OWL-specified ontology is interpreted as a set of "objects" and a set of "properties" which relate objects with each other. Ontologies expressed in OWL consist of axioms that constrain the classes and their relationships. Axioms allow making explicit information that otherwise is implicit through the use of logical inference. Properties are used to state relations between individuals or between an individual and a data value. There are two main categories of properties: Object properties, which "link individuals to individuals", and Data properties, which "link individuals to data type values". Ontologies specified in OWL-DL follow

the “open world assumption”. This means that if a meta-model represented as an ontology does not state clearly axioms that enable to state a fact as being true, then it is considered to be undefined.

OWL ontologies can be represented using RDF (Resource Description Framework) [W3C 2014] triples. RDF is a W3C standard compliant format in which any ontology can be represented. The data represented using RDF consists of triples, each comprising of subject, predicate and object. A set of triples represents a graph database that can be queried using SPARQL [W3C 2013]. SPARQL is a query language which allows retrieval and manipulation of data stored in RDF. It can only retrieve data stored in the model and, contrary to the DL language, there is no inference in the language itself. The information retrieved from the model can be provided in three different ways: raw values in a table format (using SELECT query), RDF triples which are a subset of the queried model (using CONSTRUCT query), or a Boolean value. The selection of the kind of query mode depends on the particular needs. Additionally, SPARQL allows the use of aggregate operations, which can be used to apply expressions over groups of queried data. Count, sum, min, max, average, group concatenation and sampling are the supported operations.

4. Analysis with Description Logics Reasoning and SPARQL

Based on the federated arrangement proposed in [Antunes et al. 2014], the different types of EA model analysis will be demonstrated and described below using an example EA description and a determined EA meta-model, which will be extended when needed for demonstration purposes. The demonstration will be based on the use of the ArchiMate modelling language and will use a fictitious scenario comprising several elements belonging to the different abstraction layers depicted in ArchiMate. Figure 1 depicts a layered view of the scenario ArchiMate model.

As such, the ArchiMate meta-model, along with the Motivation and Implementation and Migration extensions were specified in OWL-DL. The method employed for creating an ArchiMate representation in OWL-DL involved three steps: (i) transform the ArchiMate meta-model; (ii) adding axioms and cardinalities; and (iii) transforming the ArchiMate models. In step (i), an analysis of ArchiMate's meta-model was performed concept-by-concept, including the relations with other concepts and the constraints existing in those relations. The result was the mapping of concepts into OWL Classes and the mapping of relations into OWL Object Properties.

In step (ii), restrictions were added to the properties, such as Inverse Object Properties and Super Object Properties axioms. Axioms were added to ensure the compliance against the ArchiMate specification, including cardinality. Finally, Step (iii) is scenario dependent and involves creating

objects of the classes existing in the ArchiMate ontology created in the two previous steps, which correspond to the elements modelled in an ArchiMate model.

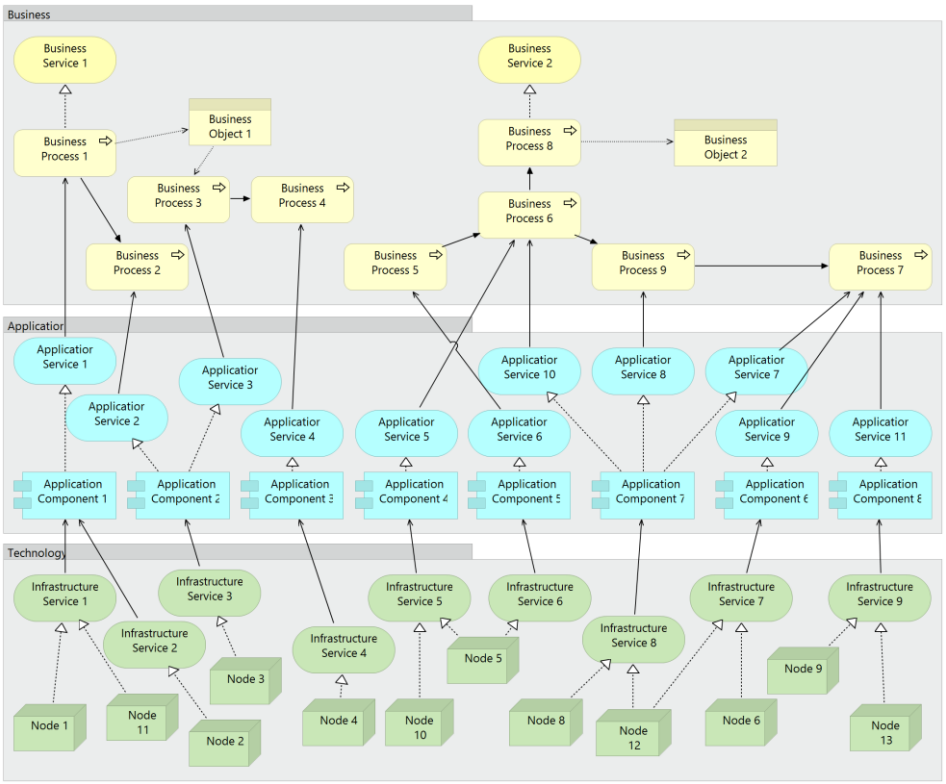


Figure 1 - Scenario ArchiMate model.

4.1. Dependency Analysis

Taking advantage of the possibility of introducing semantics on the models brought by the use of a federated approach, two Super Object Property chains were created for modelling dependencies between different elements. The `dependsDown` object property is thus a Super Object Property of the `uses`, `realizedBy`, `aggregates`, `composes`, and `assignedTo` Object Properties that resulted from the conversion of the ArchiMate relations, while the `dependsUp` object property fills the same purpose for the counterpart Inverse Object Properties: `aggregatedBy`, `composedOf`, `usedBy`, `realizes`, and `assignedFrom`. Moreover, these properties are transitive, which makes possible the creation of a graph of dependencies.

The definition of the dependency object properties allows retrieving answers to the following questions: what are the technological entities supporting the Business Service 1? Such a question can be formulated in DL as follows:

```
Thing and hasLayer some TechnologyLayer and dependsUp value
Business_Service_1
```

Figure 2 depicts the answer to the query using the HermiT reasoner in Protégé. As can be seen in the figure, different elements pertaining to the technology layer were identified as being dependencies of the Business Service 1.

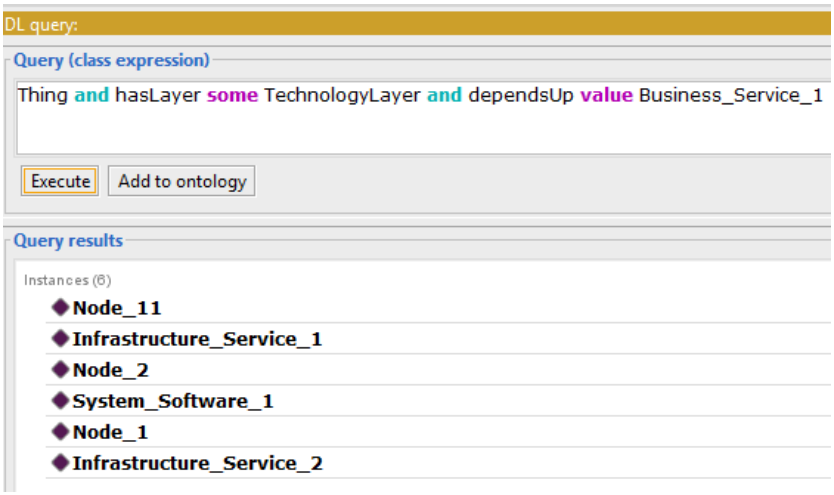


Figure 2 - Dependency analysis using DL reasoning

Figure 3 depicts the explanation given by the reasoner as why a certain node is a dependency of that service, where it is possible to see the realizes object property and the `usedBy` object property as being sub object properties of the *`dependsUp`* object property and the fact that this property is transitive.

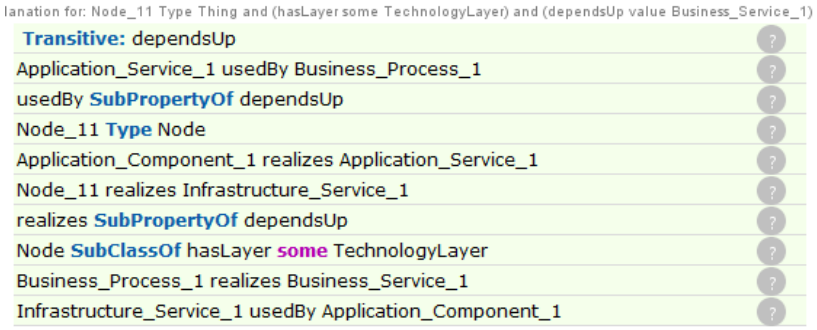


Figure 3 - Reasoner explanation for dependency.

The use of SPARQL on top of the inferred model further allows the use of aggregates. In this case, the number of technological dependencies of the aforementioned service can be computed using the below query, which returns the result "6".

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX uo: <http://sysresearch.org/ ontologies/UO.owl#>

SELECT (COUNT(?subject) as ?count)
WHERE {
   ?subject uo:dependsUp ?object .
```

```
    {?subject rdf:type uo:SystemSoftware}
    UNION {?subject rdf:type uo:Node}
    UNION {?subject rdf:type uo:Device}
    UNION {?subject rdf:type uo:InfrastructureService}.
    ?object rdfs:label "Business_Service_1" ^^xsd:string.
}
```

4.2. Coverage Analysis

The analysis of the coverage of an architecture typically tries to find gaps in the architecture, i.e. elements or relationships that are wrongly missing from the architecture description. Although the use of reasoning on top of DL allows verifying the existence of individuals of a certain class or the existence of a relationship between two individuals, the open world assumption associated with OWL does not allow us to test for the non-existence of a relationship between two individuals, unless it is explicitly specified.

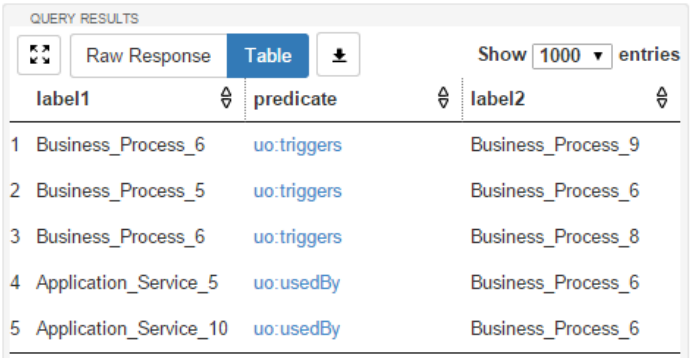
However, SPARQL can be used to this end. The SPARQL query below exemplifies the verification of the existence or not of application support provided to processes. It uses the particular query form ASK, which returns a Boolean depending on the verification of the triple patterns specified in the query. For instance, the execution of the query as exemplified below, i.e., for Business Process 7, returns precisely a Boolean value of "true", while if executed for Business Process 8, it will return the value "false".

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>

ASK {
  FILTER EXISTS {?subject uo:dependsDown ?object} .
  ?subject rdfs:label "Business_Process_7"^^xsd:string .
  ?object rdfs:label ?label2 .
  ?subject rdf:type uo:BusinessProcess .
  ?object rdf:type uo:ApplicationComponent .
}
```

4.3. Interface Analysis

The use of a DL reasoner can be a way of revealing the elements that interface, explicitly or implicitly, with a given element. However, this can also be performed with the use of SPARQL queries. Queries on top of a non-inferred ontology reveal explicit relationships, and queries on top of the inferred model make explicit otherwise implicit relationships. The query below can be used for determining the interfacing elements of Business Process 6, with Figure 4 revealing the results when executed on top of the non-inferred model.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>

SELECT ?label1 ?predicate ?label2
WHERE {
  ?subject ?predicate ?object .
  {?subject rdfs:label "Business_Process_6"^^xsd:string} UNION
  {?object rdfs:label "Business_Process_6"^^xsd:string}  .
  ?subject rdfs:label ?label1 .
  ?object rdfs:label ?label2 .
}
```


QUERY RESULTS

Raw Response | Table | Show 1000 entries

	label1	predicate	label2
1	Business_Process_6	uo:triggers	Business_Process_9
2	Business_Process_5	uo:triggers	Business_Process_6
3	Business_Process_6	uo:triggers	Business_Process_8
4	Application_Service_5	uo:usedBy	Business_Process_6
5	Application_Service_10	uo:usedBy	Business_Process_6

Figure 4 - Interface analysis SPARQL query results

4.4. Heterogeneity Analysis

As previously referred, the analysis of the heterogeneity of the architecture can take advantage of the subsumption and instance checking reasoning tasks that can be performed on top of a DL-specified ontology. In practice, the reasoner can be used to show the hierarchical arrangement of the classes existing on the architecture ontology, and it can be further enquire to show all the individuals, given a determined class.

Nonetheless, the use of SPARQL can provide interesting insights in a tabular format. The SPARQL query below can be used to return all the individuals per class of the scenario being addressed. Figure 5 depicts an excerpt of the results to the query.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>

SELECT DISTINCT ?class ?individual
WHERE {
  ?class a owl:Class.
  ?ind rdf:type ?class.
  ?ind rdfs:label ?individual.
  ?class rdfs:subClassOf uo:ArchimateConcept
}
```

QUERY RESULTS

Raw Response | Table | Show 1000 entries

	class	individual
1	uo:BusinessService	Business_Service_2
2	uo:BusinessService	Business_Service_1
3	uo:BusinessActor	Business_Actor_2
4	uo:BusinessActor	Business_Actor_1
5	uo:Node	Node_3
6	uo:Node	Node_13

Figure 5 - Heterogeneity analysis SPARQL query results excerpt.

4.5. Complexity Analysis

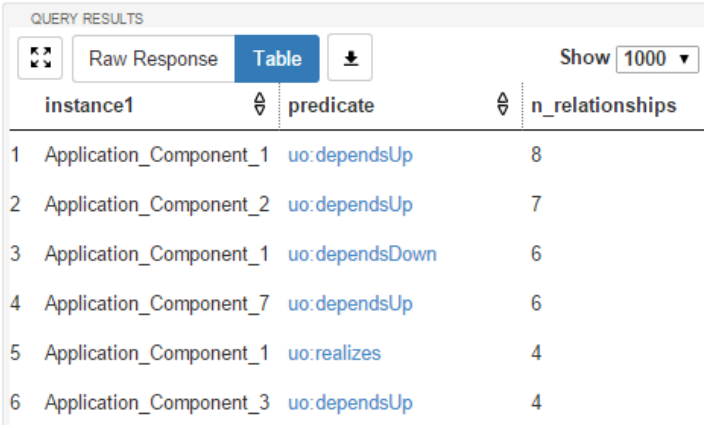
One possible measure of complexity in enterprise architecture can be given by the number of elements and relationships present in the architecture description. With the use of SPARQL, it is possible to perform simple calculations through the use of aggregates. The query below can thus be used for counting the number of individuals pertaining to different classes in the scenario ontology.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>

SELECT ?class (COUNT(?individual) as ?count)
WHERE {
  ?class a owl:Class.
  ?class rdfs:subClassOf uo:ArchimateConcept.
  ?individual rdf:type ?class.
  ?individual rdfs:label ?label.
}
GROUP BY ?class ORDER BY DESC(COUNT(*))
```

The number of relationships in which each individual of a particular class, in this case Application Component, participates in can be obtained with the query below.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>

SELECT ?instance1 ?predicate (COUNT(?object) as ?n_relationships)
WHERE {
  ?subject ?predicate ?object.
  ?subject rdfs:label ?instance1.
  ?object rdfs:label ?instance2.
  ?subject rdf:type uo:ApplicationComponent.
} GROUP BY ?instance1 ?predicate ORDER BY DESC(COUNT(*))
```

Figure 6 depicts an excerpt of the results of the execution of the query.


QUERY RESULTS

Raw Response | Table | Show 1000

	instance1	predicate	n_relationships
1	Application_Component_1	uo:dependsUp	8
2	Application_Component_2	uo:dependsUp	7
3	Application_Component_1	uo:dependsDown	6
4	Application_Component_7	uo:dependsUp	6
5	Application_Component_1	uo:realizes	4
6	Application_Component_3	uo:dependsUp	4

Figure 6 - Complexity analysis SPARQL query results excerpt.

Moreover, the query can be modified to target a specific type of relationship, in this case, the dependencies defined in Section 4.1, as can be seen below.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>

SELECT ?instance1 (COUNT(?object) as ?n_dependencies)
WHERE {
  ?subject ?predicate ?object.
  ?subject rdfs:label ?instance1.
  ?object rdfs:label ?instance2.
  ?subject rdf:type uo:ApplicationComponent.
  {?predicate rdfs:subPropertyOf uo:dependsUp} UNION {?predicate
rdfs:subPropertyOf uo:dependsDown}.
}
GROUP BY ?instance1 ORDER BY DESC(COUNT(*))
```

4.6. Compliance Analysis

In order to demonstrate this type of analysis, let us imagine that in this fictional scenario there is one requirement that concerns the availability of the infrastructure service delivered by Node 1. The requirement states that the availability of the Infrastructure Service 1 should never fall below 90%. Such indicator could be modelled as an attribute of the service. However, the ArchiMate concepts have no pre-defined attributes.

As such, using the aforementioned federated arrangement, a specialization of the concept of Infrastructure Service can be integrated with the main model, augmenting the class `InfrastructureService` with the data property `hasAvailability`. Then, through the use of automatic service monitoring tools, the value of such data property can then be regularly updated with real data concerning the availability indicator.

As the Requirement class already exists in the ontology originating from ArchiMate, the aforementioned requirement would be created as the object `AvailabilityRequirement` with the following type:

```
Requirement and (realizedBy only InfrastructureService and (has
                Availability only int[>= 90]))
```

The requirement would be associated with the Infrastructure Service 1 through an object property assertion `realizedBy Infrastructure_Service_1`. Assuming that the service monitoring kept feeding the data property value with the real availability of the service and with the reasoner continuously processing the ontology, it is possible to continuously monitor the compliance of the architecture with the requirement.

Therefore, if the value drops below 90, the reasoner will report an inconsistency in the ontology therefore indicating that compliance with that particular requirement is not met. Figure 7 depicts an inconsistency being detected by the HermiT reasoner and explained in Protégé.

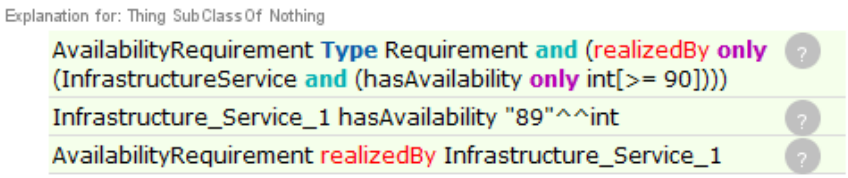


Figure 7 - Compliance analysis inconsistency explanation.

4.7. Cost Analysis

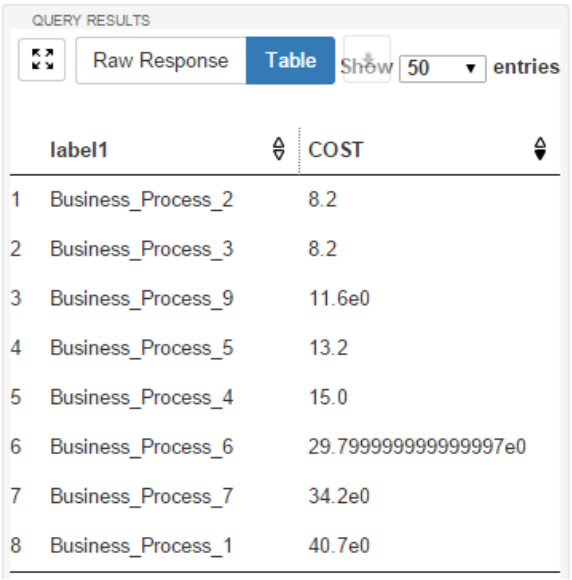
The demonstration of this type of analysis also involves augmenting the scenario's architecture description with cost related information. In this case, this can be achieved with the extension of the architecture description with cost information associated with the infrastructure nodes modeled in the scenario through a data property `cost` with a data range of the type double. Table 1 depicts the associated cost information associated with each node.

Table 1 – Cost information associated with nodes.

Element	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 8	Node 9	Node 10	Node 11	Node 12	Node 13
Cost	10.5	3.2	8.2	15.0	13.2	2.2	7.1	10.1	5.0	27.0	4.5	10.3

The addition of this information coupled with dependency information on top of the inferred ontology makes possible the execution of relevant analysis. For instance, the question "what is the cost associated with the business processes of the scenario?" can be answered with the use of the following SPARQL query, which makes use of the aggregate SUM. The results of the query can be observed in Figure 8.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>
PREFIX demo_scenario:
<http://sysresearch.org/ontologies/scenarios/demo_scenario.owl#>

SELECT ?label1 (SUM(?cost) as ?COST)
WHERE {

  ?subject uo:dependsUp ?object .
  ?object rdf:type uo:BusinessProcess.
  ?object rdfs:label ?label1.
  ?subject demo_scenario:cost ?cost.
}
GROUP BY ?label1 ORDER BY ?COST
```


QUERY RESULTS

Raw Response | Table | Show 50 entries

	label1	COST
1	Business_Process_2	8.2
2	Business_Process_3	8.2
3	Business_Process_9	11.6e0
4	Business_Process_5	13.2
5	Business_Process_4	15.0
6	Business_Process_6	29.799999999999997e0
7	Business_Process_7	34.2e0
8	Business_Process_1	40.7e0

Figure 8 - Cost analysis SPARQL query results

4.8. Benefit Analysis

Benefit analysis is another example that requires some extension to the original scenario architecture description. In this case, the extension involves the addition of business process goal information to the existing scenario. Figure 9 depicts a graph with the new added goals related to the business processes of the scenario.

Moreover, the ArchiMate classes Goal and Business Process will be impacted, being extended with data properties that can be used for defining goal metrics. For this example, the data property defined is `income`, which will have the range of double, and will be used for defining income metrics for business processes. In particular, the individual Goal 6 will have this property with the value “30”.

For the sake of this demonstration, let us imagine that the business processes in Table 2 have the following income values.

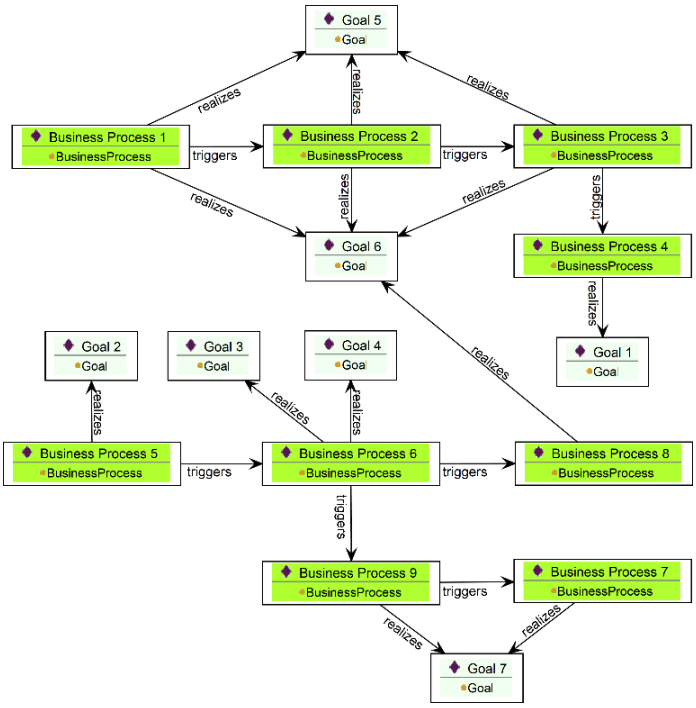


Figure 9 - Graph of goals associated with business processes.

Table 2 - Income information associated with business processes.

Element	Business Process 1	Business Process 2	Business Process 3	Business Process 8
Income	12.0	7.5	3.1	16.3

Using the inferred model, the income generated by business processes realizing Goal 6 can be verified with the SPARQL query below that takes advantage of dependency information on top of the inferred model.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>
PREFIX demo_scenario:
<http://sysresearch.org/ontologies/scenarios/demo_scenario.owl#>

SELECT ?label1 ?income
WHERE {
  ?subject uo:dependsUp demo_scenario:Goal_6 .
  ?subject rdfs:label ?label1.
  ?subject demo_scenario:income ?income.
}
```

In order to determine if Goal 6 is being met, a SPARQL query in the ASK form can be elaborated as displayed below. Note the use of the HAVING expression that filters the resulting triples to be

evaluated in the ASK query. If the sum of the income generated by processes is lower than the income defined as the goal, no triples are returned and the ASK query will be evaluated as false. In this case, the income generated by the processes is higher than the set income goal and hence the query will be evaluated as true.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>
PREFIX demo_scenario:
<http://sysresearch.org/ontologies/scenarios/demo_scenario.owl#>

ASK
WHERE {
  ?subject uo:dependsUp demo_scenario:Goal_6 .
  ?subject rdfs:label ?label1.
  ?subject demo_scenario:income ?income.
  demo_scenario:Goal_6 demo_scenario:income ?goalinc.
}
GROUP BY ?INCOME ?goalinc
HAVING (SUM(?income) >= ?goalinc)
```

Finally, in order to determine the contribution of each process to the goal as a percentage, we can use a SPARQL query in the SELECT form, as shown below. Note the use of the BIND expression that allows binding the resulting value of the calculation of the percentage of the contribution to a variable that can be then displayed in the results. Figure 10 depicts the resulting percentages for the example.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX uo: <http://sysresearch.org/ontologies/UO.owl#>
PREFIX demo_scenario:
<http://sysresearch.org/ontologies/scenarios/demo_scenario.owl#>

SELECT ?process ?percentage
WHERE {
  ?subject uo:dependsUp demo_scenario:Goal_6 .
  ?subject rdfs:label ?process.
  ?subject demo_scenario:income ?income.
  demo_scenario:Goal_6 demo_scenario:income ?goalinc.
  bind((?income * 100)/?goalinc as ?percentage)
}
```

QUERY RESULTS

Raw Response | Table | Show 1000 entries

	process	percentage
1	Business_Process_1	40.0e0
2	Business_Process_2	25.0e0
3	Business_Process_3	10.333333333333334e0
4	Business_Process_8	54.333333333333336e0

Figure 10 - Benefit analysis SPARQL query results.

5. Conclusion

This paper demonstrated the application of ontologies and associated techniques in the analysis of EA models, namely the use of computational inference to reason upon the models and deduce implicit knowledge, and the use of SPARQL to retrieve and manipulate the data stored in the models. The demonstration involved the use of practical examples of the application of the mentioned techniques for performing different types of analysis, in the context of a fictional scenario.

We believe that through this work some insights were given at the possibilities brought by the use of the techniques in scope. By representing the models using logic-based ontologies, it is possible to perform analysis just by using the syntactic and semantic information contents of the models and reason upon it. The use of SPARQL widens the possibilities, since besides allowing for the retrieval of information, it allows performing simple calculations on the retrieved data. Moreover, it can be used in conjunction with computational inference, as the queries can be applied on top of an inferred model.

6. Acknowledgement

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with references UID/CEC/50021/2013 and EXCL/EEI- ESS/0257/2012 (DataStorm).

7. References

Antunes, G., M. Bahkshandeh, R. Mayer, J. Borbinha, and A. Caetano. "Using Ontologies for Enterprise Architecture Integration and Analysis." *Complex System Informatics and Modeling Quarterly* 1 (April 2014): 1-23.

Areces, C. "Logic Engineering. The Case of Description and Hybrid Logics." Ph.D. dissertation, Institute for Logic, Language and Computation, University of Amsterdam, 2000.

Bertolino, A., P. Inverardi, and H. Muccini. "Software architecture-based analysis and testing: a look into achievements and future challenges." *Computing* 95 (2013): 633-648.

Bucher, T., R. Fisher, S. Kurpjuweit, and R. Winter. "Enterprise architecture analysis and application. An exploratory study." *Proceedings of the 1st Workshop on Trends in Enterprise Architecture Research (TEAR 2006)*. 2006.

Caetano, A., A. Assis, and J. Tribolet. "Using business transactions to analyze the consistency of business process models." *Proceedings of the 45th Hawaii International Conference on system Science (HICSS)*. 2012.

Davoudi, M. R., and K. Sheikhvand. "An Approach towards Enterprise Architecture Analysis using AHP and Fuzzy AHP." *International Journal of Machine Learning and Computing* 2 (2012): 46-51.

de Boer, F.S., M.M. Bonsangue, J. Jacob, A. Stam, and L.van der Torre. "Enterprise Architecture Analysis with XML." *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS)*. 2005.

Franke, U., W. R. Flores, and P. Johnson. "Enterprise Architecture Dependency Analysis using Fault Trees and Bayesian Networks." *Proceedings of the 2009 Spring Simulation Multiconference SpringSim '09*. 2009.

Genesereth, M. R., and N. J. Nilsson. *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann, Los Altos, CA, 1987.

Guarino, N., D. Oberle, and S. Staab. "Handbook on Ontologies." Chap. What Is an Ontology?, 1-17. Springer Berlin Heidelberg, 2009.

Hamalainen, N. "Evaluation and Measurement in Enterprise and Software Architecture Management." Master's thesis, University of Jyvaskyla, 2008.

Johnson, P., and M. Ekstedt. *Enterprise Architecture: Models and Analyses for Information Systems Decision Making*. Lightning Source Incorporated, 2007.

Johnson, P., M. Ekstedt, E. Silva, and L. Plazaola. "Using Enterprise Architecture for CIO Decision-Making:On the Importance of Theory." *Proceedings of the Second Annual Conference on Systems Engineering Research*. 2004.

Johnson, P., R. Lagerstrom, P. Narman, and M. Simonsson. "Enterprise Architecture Analysis with Extended Influence Diagrams." *Information Systems Frontiers* 9 (2007): 163-180.

Lagerstrom, R., P. Johnson, and D. Hook. "Architecture analysis of enterprise systems modifiability - Models, analysis, and validation." *The Journal of Systems and Software* 83 (2010): 1387-1403.

Lankhorst, M. *Enterprise Architecture at Work: Modeling, Communication, and Analysis*. Springer, 2005.

Niemann, K. D. *From Enterprise Architecture to IT Governance*. Friedr. Vieweg & Sohn Verlag, 2005.

Olivé, A. *Conceptual Modeling of Information Systems*. Springer-Verlag, Berlin, 2007.

Studer, R., R. Benjamins, and D. Fensel. "Knowledge Engineering: Principles and Methods." *Data & Knowledge Engineering* 25 (1998): 161-198.

Vaculin, R. "Process Mediation Framework for Semantic Web Services." Ph.D. dissertation, Department of Theoretical Computer Science and Mathematical Logic, Faculty of Mathematics and Physics, Charles University, 2009.

W3C. "OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition)." *OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition)*. World Wide Web Consortium Recommendation, 2012.

W3C. "OWL Web Ontology Language Semantics and Abstract Syntax." *OWL Web Ontology Language Semantics and Abstract Syntax*. World Wide Web Consortium Recommendation, 2004.

W3C. "RDF 1.1 Concepts and Abstract Syntax." *RDF 1.1 Concepts and Abstract Syntax*. World Wide Web Consortium Recommendation, 2014.

W3C. "SPARQL 1.1 Query Language, World Wide Web Consortium Recommendation." *SPARQL 1.1 Query Language, World Wide Web Consortium Recommendation*. World Wide Web Consortium Recommendation, 2013.

As Tecnologias da Informação no processo de ensino-aprendizagem da música

Luís Pereira da Costa[1], Bráulio Alturas[2], Abílio Oliveira[3]

[1, 2, 3] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Portugal
lmpcost@gmail.com [1]
braulio.alturas@iscte.pt [2]
abilio.oliveira@iscte.pt [3]

Resumo

Há muito que as Tecnologias da Informação e da Comunicação (TIC) entraram no nosso quotidiano, transformando, por vezes de uma forma radical, o modo como se realizam as mais diversas actividades. Não é por isso de estranhar que também no ensino as TIC tenham vindo progressivamente a desempenhar um papel cada vez mais importante, tornando-se imperativo que sejam pensados novos métodos de ensino, de modo a que estas sejam instrumentos auxiliares para aos vários atores envolvidos. Se as TIC podem ajudar a melhorar o ensino em geral, importa propor um modelo para a utilização destas tecnologias no ensino da música aproveitando as potencialidades dos métodos tradicionais de ensino. Neste artigo pretende-se abordar a necessidade de incrementar o esforço de investigação relativo à forma como as tecnologias da informação e comunicação poderão contribuir para um melhor processo de ensino-aprendizagem da música em geral.

Palavras-chave: tecnologias da informação e da comunicação, ensino-aprendizagem, potencialidades, música

1. Introdução

O valor da música na existência humana tem implicado que, desde há muito, haja a preocupação de perceber como é que a mesma se aprende e qual a forma mais correta de ensiná-la [Vasconcelos 2003], [Krauss 1974]. O ensino/aprendizagem da Música não tem ficado imune às várias reformas educativas que se sucedem no nosso País, em espaços de tempo mais ou menos curtos.

Impulsionadas por rápidos avanços, as TIC são cada vez mais utilizadas na educação, possibilitando a criação de caminhos alternativos que, tanto ajudam os alunos na aquisição de conhecimento e no desenvolvimento de competências específicas, como suscitam a formação de novos métodos e práticas no processo de ensino-aprendizagem.

Foi considerado o interesse manifestado pela tutela em definir orientações e implementar o Programa de Generalização do Ensino do Inglês e de Outras Actividades de Enriquecimento

Curricular, através do Despacho nº 12 591/2006 (2.ª série), criando assim o conceito de escola a tempo inteiro e facultando o ensino-aprendizagem da Música aos alunos do 1.º Ciclo, por docentes especializados nesta área. Além disso, a tutela apresentou também orientações programáticas para o ensino da música no ensino básico [Ministério da Educação 2006].

A presença das Tecnologias de Informação e Comunicação (TIC) no ensino tem tido uma evolução crescente por duas grandes razões: 1) estas tecnologias proporcionam novas formas de comunicação na relação professor-aluno; 2) e possibilitam um maior e mais rápido acesso ao conhecimento e à partilha de informação. O ensino e a aprendizagem integram um mesmo processo, no qual o professor planeia, dirige e controla o processo de ensino tendo em vista estimular e suscitar a atividade autónoma dos alunos para a aprendizagem desejada. E isto evidencia que a aprendizagem escolar não é casual e espontânea, mas, antes, uma atividade planeada, intencional e dirigida.

O ensino da música é muito diverso, complexo e específico. Podendo estar integrada na formação geral do indivíduo deste os seus primeiros estudos (primário e secundário), o ensino da música tem de estar incluído na formação cultural e na preparação técnicas de profissionais (cantores, instrumentistas, compositores, regentes de agrupamentos musicais). Deste modo, para o investigador que se interesse pelo ensino da música, e pela presença das potencialidades proporcionadas pela informática, é pertinente questionar: De que modo são as TIC usadas no processo de ensino-aprendizagem da música? Esta pergunta pode derivar para uma outra questão : em que sentido podem as TIC fomentar aperfeiçoamentos qualitativos no ensino-aprendizagem da música?

2. Breve abordagem teórica

Considerando a evolução da metodologia geral de ensino de música e conhecido o papel das tecnologias informáticas na comunicação inerente ao ensino, estas terão de influenciar de algum modo o ensino da música, o que gera o interesse pela participação das formas de multimédia no processo de ensino-aprendizagem da música.

Para um melhor entendimento de como as TIC podem ser usadas no ensino da música, serão consideradas as principais perspectivas do ensino-aprendizagem.

Perspectiva comportamentalista. A perspectiva comportamentalista no ensino centra-se no desenvolvimento de técnicas de modificação do comportamento na sala de aula. Os professores são orientados no sentido de esperarem dos alunos as respostas corretas, reforçando-as de forma imediata e consistente. Segundo Gillani (2003), muitos programas de modificação do comportamento e de instrução, *software* educativo, instrução programada e ensino assistido por

computador, basearam-se num condicionamento perante a instrução programada. Numerosos sistemas de ajuda à aprendizagem foram concebidos tendo como base as premissas dos principais conceitos do comportamentalismo. As "máquinas de ensinar" foram assim concebidas segundo os princípios do ensino programado, propondo sistemas de ajuda à aquisição de conhecimentos personalizados e adaptados ao ritmo e velocidade de progressão do aluno [Legros et al. 2002].

Perspectiva cognitiva. Na perspectiva cognitiva a aprendizagem baseia-se no facto de as capacidades cognitivas do aluno progredirem através de transformações fundamentais. A investigação sobre o desenvolvimento cognitivo teve um grande impacto no movimento construtivista e nas técnicas de educação [Baviskar, et al. 2009]. Segundo esta perspectiva a aprendizagem produz-se a partir da experiência, que não é entendida como uma simples transferência mas como uma representação da mesma [Bruner 1973]. A aprendizagem é concebida como um processo de construção do conhecimento e não como um processo de aquisição do conhecimento, porém, as actividades de ensino são actividades de ajuda à construção de conhecimento e não de transmissão de conhecimento. A visão cognitivista tem possibilitado a criação de ferramentas de ajuda ao desenvolvimento do processo de tratamento da informação, nomeadamente computadores com sistemas de orientação e regulação, tutoriais e *software* inteligente. Tal como o ensino assistido por computador, estes sistemas possibilitam formas de ensino individualizado e com maiores potencialidades interactivas [Legros et al. 2002].

Perspectiva do construtivismo. A perspectiva do construtivismo relaciona-se com investigações da Psicologia sobre o processamento da informação, a organização do conhecimento, os estilos de pensamento e a tomada de decisões [Hannafin et al. 1997], mas tornou-se também numa corrente da investigação matemática e igualmente uma corrente artística do século XX, que privilegia uma construção mais ou menos geométrica das formas, de natureza espiritual e estética. Baseia-se na ideia de que a realidade do mundo é construída na mente dos indivíduos a partir de representações mentais ou modelos do mundo. O conhecimento adquire, assim, um estatuto subjectivo, uma vez que este resulta das nossas experiências sobre o mundo. No construtivismo a ênfase é colocada mais no aluno do que no professor. O aluno constrói o seu conhecimento e soluções para os problemas. Deste modo, a utilização de ferramentas tecnológicas tem de considerar a capacidade dos alunos para lidarem com essas ferramentas.

Ensino centrado no professor. Num modelo de **ensino centrado no professor** é possível usar algumas das vantagens das TIC, mas ainda é uma abordagem tradicional de ensino à distância.

Além disso, incide sobretudo no ensino, apoiando-se na transmissão de informação, adoptando as mesmas técnicas de ensino, mas com um suporte de uma ferramenta tecnológica. Possibilita a transferência das técnicas, estratégias e métodos do ensino presencial para o ensino *online*, centrando-se mais numa transmissão de conhecimentos do que num processo de aprendizagem [Grasha 1994], [Bowers e Flinders 1990].

Ensino centrado no aluno. O modelo de **ensino centrado no aluno** é uma tendência contemporânea em que se valoriza que a instituição de ensino passe a centrar-se na figura do estudante e não na do professor, o qual passa a ser um *mediador*. É um ensino com participação ativa do aluno em atividades colaborativas, voltado para o "aprender a aprender" e para a aprendizagem significativa crítica. As principais características são: a) ênfase e centralidade no aluno; b) contextos de aprendizagem com base na construção de conhecimentos e competências; c) favorecimento da autonomia e a reflexão crítica [Finkel 2008], [Moreira 2010], [Weimer 2002].

Ensino centrado nos meios. Um modelo de **ensino centrado nos meios** é resultante do desenvolvimento das TIC, e caracteriza-se por se concentrar no uso de uma ou de várias ferramentas tecnológicas. O aluno e o professor continuam a ser atores do processo ensino-aprendizagem mas têm, neste modelo, um papel secundário. O professor passa a ser um fornecedor de conteúdos e o aluno transforma-se num utilizador desses mesmos conteúdos, a partir dos quais tem de fazer a sua auto-formação. Este modelo tem como características principais: a) uma natureza visual, pois a *Internet* é um meio dessa natureza onde se utilizam as imagens, audio, vídeo e outras fontes multimédia, em vez de serem fornecidos simples textos aos alunos; b) a informação escrita que compõe o ensino centrado no conteúdo deve ser concisa e directa; c) trata-se de um modelo interactivo**,** pois os alunos podem interagir com o material didáctico através de questionários e actividades multimédia, possibilitando o desenvolvimento de competências, mostrar conhecimento, descobrir novas informações e reforçar a aprendizagem. Além disso, também concedem meios para que os alunos possam interagir uns com os outros [Duart e Sangra 2000], [Miranda et al. 2005]

Seja qual for a contribuição das TIC e o tipo de *software* a adoptar para utilização no processo de ensino-aprendizagem da música, é importante que sejam observados pressupostos pedagógicos coerentes com os objetivos educativos de cada situação e, sobretudo, que o mesmo auxilie o desenvolvimento das competências e conhecimentos musicais da forma mais abrangente possível. Alguns tipos de *software* são apenas fundamentados em métodos tradicionais de apresentação, aplicação de conceitos e avaliação de resultados, em vez de serem construídos de acordo com estudos de música e de desenvolvimento cognitivo [Krüger et al.

1999]. A abordagem teórica para algumas pesquisas curriculares no ensino da música [Hentschke 1996] tem seguido as diretivas do *Modelo (T)EC(L)A* de Swanwick (1979, 1988).

A necessidade de dispor de modelos didáticos na educação musical surge do próprio estado do ensino musical, que actualmente se encontra em grande parte nas mãos dos processos tradicionais. É verdade que a tradição corresponde a uma forma de seleção do conhecimento ao longo do tempo, embora se trate de uma seleção que responde a questões de natureza essencialmente social. A procura de uma educação musical de qualidade deve realizar-se pensando num ensino com características de sistematicidade e, portanto, fundamentado num conhecimento que não proceda exclusivamente da tradição. Os modelos didáticos proporcionam ao docente uma ferramenta que lhe possibilita analisar os sucessos da aula, para assim compreendê-los e poder introduzir mudanças que contribuam para melhorar a aprendizagem dos seus alunos.

Considerando as perspectivas e os modelos referidos, importa investigar no caso muito específico do ensino da música como podem os docentes decidir a presença dos vastos e modernos recursos tecnológicos (TIC) no processo ensino-aprendizagem no sentido de encontrarem o melhor caminho a ser construído para o desenvolvimento de competências técnicas e artísticas. Entende-se que as TIC podem ser utilizadas como ferramenta de natureza mediadora e facilitadora no processo de ensino, e os seus recursos explorados com a finalidade de extrair deles as potencialidades benéficas para docentes e alunos.

Depois de uma revisão da literatura sobre o ensino da música, métodos e técnicas utilizados é possível concluir a oportunidade de um modelo conceptual do processo ensino-aprendizagem que tenha em consideração as potencialidades introduzidas pela utilização e pela influência das TIC.

3. Contribuição para o avanço do conhecimento

Uma das razões que devem fundamentar o incremento da investigação do papel das TIC é considerar que o processo ensino-aprendizagem tem um elevado valor na evolução da cultura musical. São os resultados deste processo, em particular as competências obtidas e o seu desenvolvimento, que conduzem à formação de novos atores dos diversos cenários no mundo da música, independentemente dos géneros que comporta. Considerando esta abordagem, a questão central para esses atores tem de incluir os recursos tecnológicos a utilizar nos percursos do ensino para se estimular a aprendizagem e conseguir melhores níveis de aproveitamento.

Também de importância, uma outra razão fundamenta-se na interatividade do processo de ensino que merece uma observação mais cuidada no que se refere às suas consequências sociais

e profissionais. Participando de uma forma mais intensa e mais racional neste processo, as TIC podem prestar um valioso auxílio, fazendo com que os conhecimentos e as competências criadas a nível individual ou de turma possam partilhados entre alunos e docentes, entre várias escolas e entre produtores musicais (editoras, laboratórios e produtores de espetáculos), o que favorece o conhecimento de alunos e de artistas, dando lugar a uma mais visível competitividade que, quando positiva, pode fomentar desenvolvimento humano nas suas capacidades artísticas (composição, execução, ensino).

Neste momento de desenvolvimento pessoal, julga-se ser adequado indicar como uma outra razão a reduzida produção de trabalhos de investigação que tenham em conta a relação entre o processo ensino-aprendizagem no mundo da música e as TIC, sobretudo no que se refere a procedimentos e a técnicas que facilitem o processo.

A finalidade de um esforço de investigação é ser um contributo de relevo para a compreensão e descrição do valor das TIC no processo de ensino-aprendizagem no domínio da música. É necessário compreender melhor como é que as tecnologias informáticas estão a ser integradas neste processo e como é que poderão ser exploradas de uma forma mais orientada nas suas potencialidades em favor da interação professor-aluno e da interatividade entre alunos. Deste modo e considerando a questão inicial, são destacados os seguintes objetivos:

1. Identificar as características gerais do processo de ensino-aprendizagem da música;
2. Estudar e identificar as possibilidades de adaptar as TIC à especificidade dos cursos de música (iniciação e continuação) e a cada um dos níveis de ensino;
3. Estudar e identificar quais as características das TIC que mais se adequam às situações do ensino da música e que podem conduzir ao aumento de aproveitamento.

Importa entender como é que os diversos atores do processo de ensino encaram a utilização das TIC e como é que a sua presença tem influenciado as suas diversas fases. Considerando que não há um modelo específico para o processo ensino-aprendizagem no domínio da música que englobe a presença das TIC e as melhores práticas pedagógicas a elas devidas, qualquer esforço de investigação deve pretende também elaborar e propor um modelo, que inclua as consequências possíveis do uso das TIC e possibilite mais fundamentadas decisões nas dimensões do processo de acordo com a sua diferenciação. Deste modo, a inclusão das TIC poderá ser encarada como um fator positivo do ensino musical, contribuindo para uma melhor qualidade da cultura e do desenvolvimento socioeconómico.

4. Conclusões

O rápido crescimento das TIC provocou notáveis mudanças nas atividades humanas e tem possibilitado a sua adoção e integração no processo ensino-aprendizagem. A eficácia da sua integração nas salas de aula representa um desafio para os professores para os alunos. Embora possa haver vontade de utilizar a integração das TIC na educação, algumas barreiras ainda podem ser encontradas.

Para que esta integração tenha sucesso no domínio do ensino da música é necessário considerar os fatores que possam influenciar positivamente as atitudes e comportamentos dos professores, nomeadamente competências profissionais para utilizar estas tecnologias aplicadas e também o esforço tecnológico para adaptar as TIC à especificidade e complexidade do ensino musical.

5. Referências

Baviskar, S. N.; Hartle, R.T.e Whitney, T. (2009). Essential criteria to characterize constructivist teaching: derived from a review of the literature and applied to five constructivist-teaching method articles. *International Journal of Science Education 31* (4), 541-550.

Bowers, C. A.; Flinders, D. J., *Responsive Teaching*, Teachers College Press, New York, 1990, 5-14.

Bruner, J., *Culture and Cognitive Growth, in the relevance of education*, New York, NY: Norton, 1973.

Duart, J. M.; Sangra, A. (Compiladores), *Aprender en la virtualidad*, Barcelona: Universidad Abierta de Cataluña y GEDISA, 2000.

Finkel, D. L., *Teaching With Your Mouth Shut*, Boynton: Cook Publishers, Inc, 2000.

Gillani, B. G., *Learning Theories and the Design of E-Learning Environments*, Maryland: University Press of America, 2003.

Grasha, Anthony F., "A Matter of Style: The Teacher as Expert, Formal Authority, Personal Model, Facilitator, and Delegator", *College Teaching*, Vol. 42, No. 4 (Fall), 1994, 142-149.

Hannafin, Michael J.; Hannafin, Kathleen M.; Land, Susan M.; Oliver, Kevin, "Grounded Practice and the Design of Constructivist Learning Environments", *Educational Technology Research and Development*, Vol. 45, No. 3, 1997, 101-117.

Hentschke, Liane, "Um Estudo Longitudinal Aplicando a Teoria Espiral de Desenvolvimento Musical de Swanwick com Crianças Brasileiras da Faixa Etária de 6 a 10 Anos de Idade: Pólo Porto Alegre – 1994", *Música: Pesquisa e Conhecimento 2*, Porto Alegre: CPG em Música - Mestrado e Doutorado / UFRGS - NEA, 1996, 9-34.

Hill, Manuela M.; HILL, Andrew, *Investigação por Questionário*, Lisboa: Edições Sílabo, 2000.

Krauss, E., "O ensino da música como matéria obrigatória nas escolas primárias", *Boletim Informativo da Associação Portuguesa de Educação Musical,* 5, 1974, 1-12.

Krüger, S. E.; Fritsch, E. F.; Flores, L. V.; Grandi, R. H.; Santos, T. R.; Hentschke, L.; Viccari, R. M., "Developing a Software for Music Education: An Interdisciplinary Project", *Simpósio Brasileiro de Computação e Música*, 6., 1999, Rio de Janeiro. Anais do XIX Congresso Nacional da Sociedade Brasileira de Computação, v.3. Rio de Janeiro: EntreLugar, 1999, 251-264.

Legros, D., Pembroke, E., & Talbi, A., "Les Théories de l'apprentissage et les systèmes multimédias", D. Legros, & J. Crinon (Coords.), *Psychologie des apprentissages et multimédia*, Paris: Armand Collin/VUEF, 2002, 23-39.

Ministério da Educação – DGIDC (Ed.), *Ensino da Música 1º Ciclo do Ensino Básico – Orientações Programáticas.* Lisboa: Autor, 2006.

Miranda, L.; Morais, A.; Dias, P., *Abordagens Pedagógicas para ambientes online*, 2005, Disponível em http://bibliotecadigital.ipb.pt/bitstream/10198/1090/1/PA20_2005_Abordagens_Pedagogicas_Online.pdf ,

Moreira, Marco A., "Abandono da narrativa, ensino centrado no aluno e aprender a aprender criticamente (Disclaiming the narrative, student-centered teaching, and learning how learn critically)", *Conferência proferida no II Encontro Nacional de Ensino de Ciências da Saúde e do Ambiente*, Niterói, RJ, 12 a 15 de maio de 2010 e no VI Encontro Internacional e III Encontro Nacional de Aprendizagem Significativa, São Paulo, SP, 2010.

Pereira, A., *SPSS – Guia Prático de Utilização, Análise de Dados para Ciências Sociais e Psicologia*, 3ª ed. – Revista e Corrigida, Lisboa: Edições Sílabo, 2002.

Swanwick, K., *A Basis for Music Education*, London: Routledge, 1979.

Swanwick, K., *Music, Mind and Education*, London: Routledge, 1988.

Vasconcelos, A. A., "Políticas no ensino da música em Portugal nas últimas duas décadas do século XX: contributos para uma análise crítica", *Revista de Educação Musical*, 115, 2003, 14-28.

Weimer, Maryellen, *Learner-Centered Teaching*, San Francisco: Wiley Co, 2002.

Zikmund, W., *Business Research Methods*, 4th ed. Orlando, FL: Dryden, 1994.

AUDITORIA DO ALINHAMENTO ENTRE SISTEMAS DE INFORMAÇÃO E NEGÓCIO

Gilberto Capeça [1], Filipe de Sá-Soares [2]

1) Centro ALGORITMI da Universidade do Minho, Portugal

gilbertocapessa@hotmail.com

2) Centro ALGORITMI da Universidade do Minho, Portugal

fss@dsi.uminho.pt

RESUMO

A aplicação das tecnologias de informação nas organizações suscitou o aparecimento de preocupações com o valor real dessas tecnologias no suporte ao negócio, levando a que a problemática do alinhamento entre o sistema de informação da organização e o negócio da organização assumisse papel de relevo quer na academia quer na prática profissional.

Para que as organizações alcancem o alinhamento da aplicação das tecnologias da informação com os objectivos e a estratégia do negócio torna-se necessário que abordem e articulem aspectos como a estratégia de negócio, a estratégia de sistema de informação e as infra-estruturas tecnológicas e organizacionais.

Neste artigo argumenta-se que a governação de sistemas de informação e a gestão de sistemas de informação são condições prévias necessárias à perseguição do alinhamento. Com vista a se examinarem essas pré-condições, propõe-se neste estudo um procedimento de auditoria baseado no COBIT 5 e que foi aplicado na avaliação preliminar do suporte prestado pelas tecnologias da informação às actividades organizacionais de um banco central.

PALAVRAS-CHAVE: Alinhamento, Governação de Sistemas de Informação, Gestão de Sistemas de Informação, Auditoria, Tecnologias de Informação

1. Introdução

Os sistemas de informação (SI) são hoje instrumentos cruciais para o sucesso das organizações. À medida que a utilização das tecnologias da informação (TI) se intensifica, que as organizações se comparam entre si em termos de desempenho e que os esforços e investimentos em TI são alvo de maior escrutínio, a preocupação em garantir que a aplicação das TI se encontra alinhada com o negócio da organização assume relevância crescente.

Para que as organizações alcancem o alinhamento entre os SI e o negócio têm que articular aspectos como a estratégia de negócio, a estratégia de SI e as infra-estruturas tecnológicas e organizacionais [Croateau e Begeron 2001]. Se bem que na literatura se encontrem múltiplos trabalhos que revelam a centralidade do alinhamento para as organizações (cf. Chan et al. [2006], Kappelman et al. [2013] e Luftman e Derksen [2012]) e apontam factores críticos para o seu alcance (cf. Campbell et al. [2005] e Teo e Ang [1999]), também se observam diferentes propostas para a avaliação do alinhamento entre os SI e o negócio das organizações, com ênfases distintas e abordagens metodológicas variadas (cf. Belfo e Sousa [2012], Chan e Reich [2007], Kearns e Lederer [2003], Sabherwal e Chan [2001] e Tan e Gallupe [2006]).

Atendendo-se a que a busca do alinhamento requer um processo contínuo de ajustes entre o negócio e a aplicação das TI, consubstanciado em diálogos, negociações e decisões entre agentes das áreas do negócio e agentes da área de SI (função SI), argumenta-se neste trabalho que as actividades de governação de SI e de gestão de SI são pré-condições para a perseguição do alinhamento por parte das organizações. Consequentemente, a avaliação do grau de alinhamento implicará que, num primeiro momento, se afira a qualidade daquelas duas actividades. A formalização dessa aferição poderá assumir a forma de uma auditoria.

Neste artigo avança-se com a delineação de um procedimento de auditoria que visa avaliar as actividades de governação e gestão de SI na perspectiva do alinhamento entre a aplicação das TI e o negócio da organização.

Metodologicamente, programou-se a delineação daquele procedimento mediante a realização de um estudo de caso desenvolvido numa instituição bancária central. Estruturalmente, o trabalho organiza-se da seguinte forma: após esta introdução, revê-se o conceito de alinhamento, perspectivam-se as actividades de governação e de gestão de sistemas de informação, ao que se segue a descrição do estudo realizado. Posteriormente, apresentam-se os resultados do estudo, enumeram-se as recomendações derivadas e avança-se com um conjunto de considerações finais.

2. Alinhamento entre Sistema de Informação e Negócio

O alinhamento entre o sistema de informação e o negócio constitui um dos temas recorrentes na literatura de sistemas de informação. Desde o surgimento do interesse por esta temática, precipitado pelo reconhecimento que muitas organizações efectuaram quando perceberam que não estavam a desenvolver sistemas de informação que sustentavam as suas estratégias de negócio, variados autores se têm debruçado sobre esta problemática, acompanhados por um interesse continuado por parte dos gestores de SI que, anos após ano, apontam a busca de alinhamento entre a aplicação das TI e o negócio como uma das suas preocupações centrais (cf. Kappelman et al. [2014]).

Face às falhas na obtenção de benefícios para o negócio decorrentes da aplicação de TI e aos riscos de negócio e tecnológicos com que as organizações se debatem, a procura do alinhamento entre o sistema de informação e o negócio configura um mecanismo de defesa da própria organização, ou seja, crê-se que um maior e melhor suporte das TI ao negócio potenciará um desempenho organizacional mais robusto, o que fortalecerá a posição competitiva da organização no mercado. Tal levará a organização a alcançar sucessivos, por que temporários, equilíbrios entre o seu ambiente interno e a sua envolvente.

Com vista a alcançar este estado de convergência entre as TI e o negócio, é natural que as organizações persigam uma maior integração entre processos específicos de cada um dos dois domínios envolvidos – o sistema de informação e o negócio, por exemplo, intensificando a integração entre o processo de planeamento de SI e o processo de governação estratégica da organização [Pitassi e Leitão 2002, p. 79] ou fazendo convergir a acção da função SI com os interesses estratégicos de toda a organização [Calle Jr. e Kanter 1998, p. 1].

A busca do alinhamento entre o sistema de informação e o negócio funda-se num processo de gestão em que participam as várias unidades de negócio, norteado pelos objectivos da organização. Este processo assenta na formulação de objectivos e na concretização de acções que traduzam um maior ou melhorado apoio das TI ao negócio. Para isso, é necessário levar em conta factores externos à organização, impostos à mesma pelo seu ambiente, e factores internos à organização. Entre os factores que condicionam o sucesso da organização no alcance do alinhamento podem enumerar-se o capital humano, informacional e organizacional, fazendo-se uso da categorização proposta por Kaplan e Norton [2006]. Na categoria capital humano inserem-se aspectos relativos ao talento, qualificação, experiência e conhecimento dos funcionários da organização, enquanto agentes daquele processo de alinhamento. Na categoria capital informacional contam-se aspectos conexos às infra-estruturas de TI, às plataformas de TI, às aplicações informáticas e à informação, enquanto limitadores e potenciadores estruturais do sistema de informação no processo de

alinhamento. Por fim, na categoria capital organizacional incluem-se aspectos concernentes à cultura e liderança organizacionais, enquanto limitadores e potenciadores estruturais do negócio no processo de alinhamento.

O alinhamento entre o sistema de informação e o negócio pode entender-se como a aplicação harmoniosa das TI com a estratégia, objectivos e necessidades do negócio, de forma apropriada e em tempo oportuno. Para atingir a compaginação entre a aplicação das TI e o negócio, a organização encetará um processo contínuo de ajustes com o propósito de obter a interligação entre os objectivos e estratégias de negócio e os objectivos e estratégias de SI [Affeldt e Vanti 2009].

Deve notar-se, todavia, que não basta apenas que a aplicação das TI esteja alinhada com o negócio, é importante que o negócio entenda a importância e utilidade das TI no apoio ao alcance dos seus objectivos. Trata-se, na verdade, de uma distinção ao nível daquilo que se poderia denominar o grau de qualidade do alinhamento: uma organização pode aplicar as suas TI perfeitamente alinhadas com o seu negócio, mas simultaneamente verificar-se um subaproveitamento do potencial das TI em apoiar a estratégia e processos de negócio.

O alinhamento surge como um factor chave para o sucesso das organizações num ambiente altamente dinâmico, na medida em que as TI lhes devem permitir fazerem a coisa certa (eficácia) e da melhor maneira possível (eficiência) [Luftman 2000]. De modo a concretizar esse alinhamento, Henderson e Venkatraman [1993] propuseram um modelo de ajustamento entre quatro vertentes fundamentais: a estratégia de negócio, a estratégia de TI, a infra-estrutura e os processos organizacionais e a infra-estrutura e os processos de TI.

A estratégia de negócio envolve um momento de formulação (escolha da abordagem competitiva, produtos e mercados) e um momento de implementação (decisões sobre a estrutura e capacidades que irão executar aquelas escolhas). De igual modo, a estratégia de TI deve envolver as escolhas sobre os tipos de tecnologias de informação a empregar e os seus meios de utilização e de aquisição (formulação) e decisões de como a infra-estrutura de TI deve ser configurada e gerida (implementação).

O alinhamento entre estas quatro vertentes é realizado segundo duas dimensões, conforme se ilustra na Figura 1:

- Ajustamento Estratégico – entre o domínio interno e externo.
- Integração Funcional – entre o domínio de negócio e o de sistemas de informação.

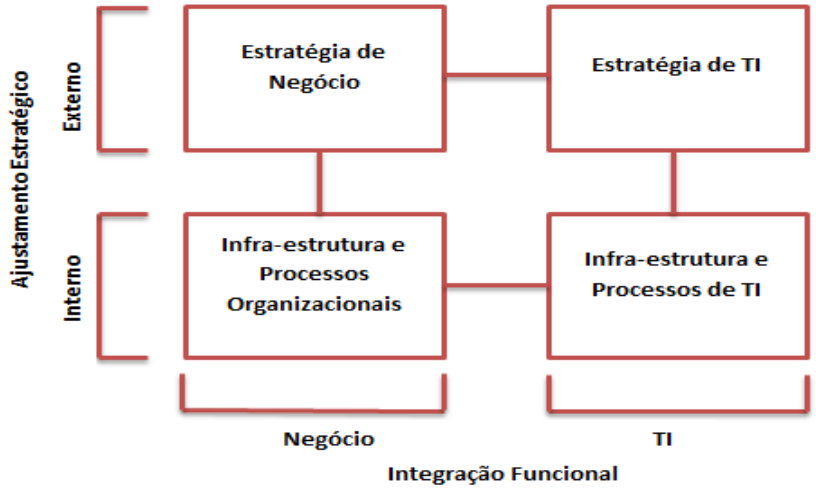


Figura 1: Modelo de Alinhamento de Henderson e Venkatraman [1993]

Conforme já se observou, o alinhamento entre a aplicaçao das TI e o negócio não é um evento, mas um processo contínuo de adaptação e transformação [Henderson e Venkatraman 1993]. Para alcançá-lo, é necessária uma mudança, muitas vezes substancial, no pensamento da gestão sobre o papel das TI na organização, assim como um entendimento da estratégia de SI e da sua importância, tanto no suporte como no direccionamento das decisões de estratégia de negócio.

Observando-se os inter-relacionamentos entre negócio e sistema de informação, ao nível das suas estratégias, processos, recursos e agentes, para a obtenção do tão almejado alinhamento entre o sistema de informação e o negócio, então poder-se-á considerar que a governação do sistema de informação da organização constitui pressuposto fundamental para o sucesso de quaisquer iniciativas que visem melhorar aquela convergência entre as TI e o negócio. Sendo a busca desta convergência um processo contínuo, então, de igual forma, se poderá argumentar sobre a relevância da gestão de sistemas de informação para a concretização, sempre imperfeita e inacabada, do alinhamento.

3. Governação e Gestão de Sistemas de Informação

A governação de sistemas de informação é um conceito relativamente recente na literatura, mas que tem vindo a conquistar relevância quer na academia quer entre os profissionais. Face à potencial aplicação das TI em todas ou quase todas as áreas organizacionais e à transformação do papel das TI de meras ferramentas administrativas em instrumentos estratégicos [Henderson e Vankatraman 1993], muitas empresas e agências governamentais avançaram para a implementação

da governação de SI por forma a alcançarem a fusão entre o negócio e as TI e obter o envolvimento da gestão sénior nas questões de SI [Haes e Van Grembergen 2009].

Embora as organizações giram muitos activos (pessoas, dinheiro, instalações, etc.), talvez a informação e as tecnologias que a recolhem, armazenam, processam e disseminam estejam entre os activos que lhes causam maior perplexidade e lhes colocam maiores desafios. Se por um lado os negócios exigem mudanças cada vez mais rápidas, os sistemas, uma vez implementados, permanecem relativamente rígidos. Frequentemente, as implementações de TI envolvem investimentos imediatos e continuados, na busca de resultados que podem ser de difícil previsão e de concretização nem sempre satisfatória. Essas incertezas e complexidades fazem muitos gestores renunciarem à responsabilidade de garantirem que os colaboradores da organização utilizam eficazmente as TI [Weill e Ross 2004].

A adopção das TI na organização é reconhecida como um processo complexo que passa pelo planeamento, avaliação custo/benefício gerado pelo sistema e pela sua adequação à realidade organizacional. Na verdade, aquela adopção é um processo de mudança que abarca não só o ambiente tecnológico, mas também o ambiente técnico, os recursos humanos e toda a estrutura da organização [Pascutti et al. 2009].

Em última análise, a governação de SI é uma actividade da responsabilidade da equipa de executivos da organização, estando as responsabilidades por essa actividade frequentemente atribuídas a uma comissão composta por executivos de topo, de linha e de SI, de modo a congregar as diferentes perspectivas e sensibilidades detidas pelos protagonistas do negócio e das TI.

A governação de SI confere liderança, estrutura organizacional e orientação para os processos que asseguram que a função SI da organização entende e sustenta as estratégias da organização e os seus objectivos. A importância desta actividade decorre, assim, do tratamento dado em sua sede às principais questões de negócio e à sua relação com os sistemas de informação, à protecção dos sistemas de informação estratégicos e às orientações que estabelece para a actividade de gestão de SI [Buckby et al. 2005].

Importa, aqui, realçar a distinção entre a actividade de governação de SI e a actividade de gestão de SI, dado por vezes se deparar com visões em que essas duas actividades são entendidas como sinónimos ou, na busca incessante da nova *buzzword*, a actividade de governação se apresenta como a nova fórmula que, potenciando benefícios e anulando dificuldades da actividade de gestão, resolverá todos os problemas deparados pelos responsáveis de SI.

Enquanto a gestão de SI se foca no fornecimento eficiente e eficaz dos serviços e produtos de SI e na gestão das operações de SI, a governação de SI enfrenta a dupla exigência de contribuir para as

operações e desempenho actuais do negócio e de transformar e posicionar as TI para atender aos desafios futuros do negócio [Peterson 2003].

Esta diferenciação entre governação de SI e gestão de SI foi explicitamente reconhecida pela versão mais recente do referencial COBIT (COBIT 5), no qual a governação de SI assegura que as necessidades, as condições e as opções das partes interessadas são analisadas para determinar objectivos balanceados e de comum acordo a serem alcançados, estabelecendo direcção através de priorização e tomada de decisão [ISACA 2012]. Para além disso, a governação também assegura a monitorização e a conformidade do direccionamento e objectivos estabelecidos previamente. Por sua vez, à gestão de SI cabe o planeamento e execução das iniciativas relacionadas com as TI, alinhando-as de acordo com as directivas recebidas.

Esta distinção entre a governação e a gestão de SI é reflectida directamente no modelo proposto pelo COBIT, o qual subdivide as práticas e as actividades relacionadas com as TI em dois domínios principais, estando a governação associada à análise, direcção e monitorização das iniciativas conexas aos sistemas de informação e a gestão associada ao planeamento, construção, execução e monitorização dos próprios sistemas de informação, conforme se ilustra na Figura 2.

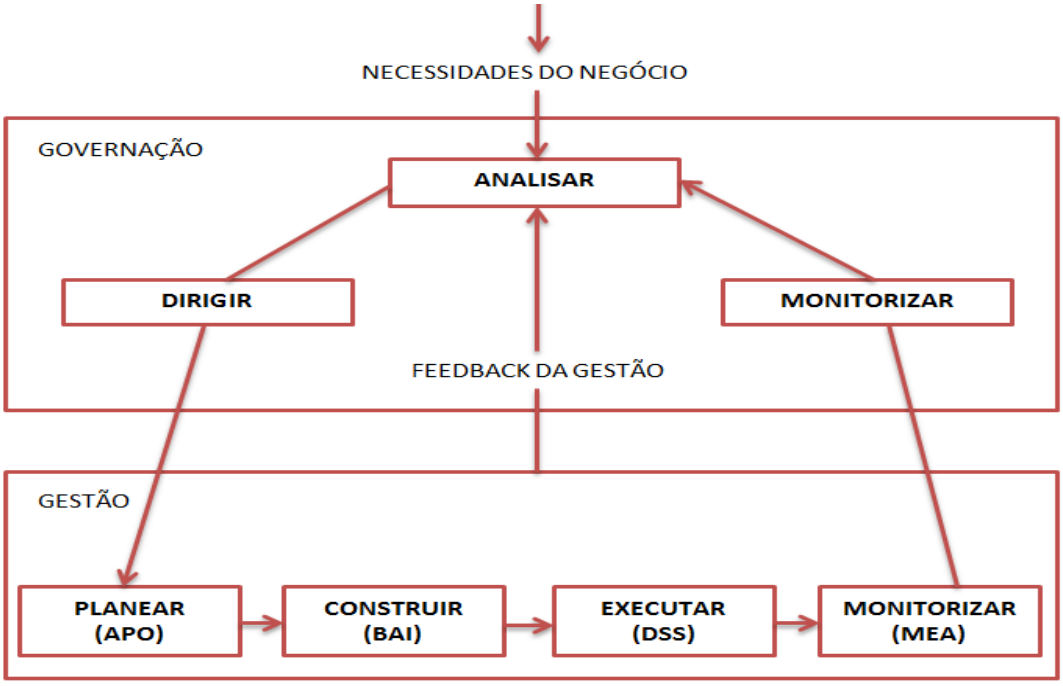


Figura 2: Relacionamento entre os Domínios do COBIT 5

Adaptado de ISACA [2012, p. 25]

A governação de SI funda-se numa estrutura de relacionamentos e processos para dirigir e controlar as TI a fim de alcançar as metas da organização pela agregação de valor. Para tal, a governação de SI visa responder às seguintes questões [Weil e Ross 2004]:

- As capacidades de TI melhoram a competitividade da organização?
- Todos os gestores da organização reconhecem as suas responsabilidades para a gestão e uso eficaz de TI ou assumem que este é um problema apenas da área de SI?
- Os investimentos em TI na organização visam atender os objectivos estratégicos ou a organização desperdiça recursos e investimentos por apenas atender a iniciativas tácticas e a necessidades operacionais?

Uma vez que a governação de SI tem por objectivo conhecer os valores e a importância estratégica das TI na organização, assegurar que as TI podem suportar as suas operações e garantir que pode implementar as estratégias necessárias face ao crescimento e à expansão da organização no futuro, as melhores práticas de governação de SI devem garantir que as expectativas de TI são atendidas e os riscos inerentes às TI minimizados [ISACA 2012].

Desta forma, o uso eficaz de TI e a integração entre a sua estratégia e a estratégia do negócio ultrapassam a ideia de TI como simples ferramentas de automação de processos e produtividade. A via para a aplicação bem-sucedida das TI na organização não está mais relacionada somente com o hardware e o software utilizados, ou com os métodos de desenvolvimento de SI adoptados, mas com o grau de alinhamento da aplicação das TI com a estratégia e as características da organização e da sua estrutura organizacional [Laurindo et al. 2001, p. 161].

Se bem que a importância do alinhamento entre a aplicação das TI e o negócio seja amplamente reconhecida quer na literatura quer na prática de SI, a avaliação de quão bem uma organização persegue ou alcança esse alinhamento levanta desafios interessantes. Neste trabalho assume-se como pressupostos que uma boa governação de SI, a par de uma boa gestão de SI, são condições necessárias para o alcance do alinhamento entre o sistema de informação e o negócio. Defende-se, assim, que o primeiro passo para aferir o grau desse alinhamento passe pela avaliação daquelas duas actividades conexas ao sistema de informação – governação e gestão. Verificando-se que essas actividades apresentam qualidade suficiente, então poder-se-á focar a atenção no procedimento de aferição do grau concreto de alinhamento e, concomitantemente, apontar-se os aspectos que necessitam de ser melhorados para intensificar a convergência entre a aplicação das TI e o negócio. Caso se constate que aquelas duas actividades carecem de qualidade, então a busca do alinhamento requererá que, previamente, se revejam e redimensionem essas actividades.

Com base no exposto, formulou-se como objectivo deste trabalho a delineação de um procedimento de auditoria (enquanto processo formal de avaliação independente e sustentado) fundado na avaliação das actividades de governação e gestão de SI como etapa prévia à aferição do grau de alinhamento entre o sistema de informação e o negócio de uma organização.

4. Descrição do Estudo

Com vista a propor e experimentar o procedimento de auditoria das actividades de governação e gestão de SI recorreu-se ao método de investigação de estudo de caso, sendo a unidade de análise um banco central que, por questões de confidencialidade, se denominará neste estudo por NSC.

O NSC é o banco central de um país situado em África. No contexto da sua estrutura organizacional é constituído por um Conselho de Administração, seis unidades de estrutura de negócio e seis de suporte, entre as quais se encontra a Direcção de Tecnologias e Sistemas de Informação (DTSI). À DTSI cabe propor e promover políticas e soluções no âmbito das tecnologias e sistemas de informação para o suporte da actividade do Banco, prestando de forma eficiente serviços de concepção, implementação, exploração e manutenção das infra-estruturas que lhes são inerentes.

Ao NSC, como banco central, compete a regulação e supervisão das instituições financeiras do seu País. Como banco regulador e supervisor, e uma vez que nos últimos anos tem crescido o número de instituições financeiras nesse país, por forma a cumprir com eficácia a sua função, por um lado, e tendo-se verificado um grande investimento em termos de TI por parte daquelas instituições, por outro lado, houve a necessidade do NSC também investir em TI. Assim, nos últimos anos, e por orientação do Governo Central, o NSC tem feito um forte investimento em TI com vista a tornar mais eficazes, seguras e céleres as suas operações de negócio, fortalecer o seu papel de supervisor do Sistema Financeiro e modernizar-se administrativamente.

Neste sentido, para aferir se os investimentos feitos em TI estão a responder aos seus objectivos, ou seja, para assegurar que as TI estão a ser aplicadas de modo a suportarem os objectivos da organização, o Conselho de Administração do NSC determinou a realização de uma auditoria aos seus sistemas de informação. Foi no âmbito desta auditoria que se enquadrou a realização deste estudo e em que se projectou e aplicou um procedimento de avaliação das condições prévias ao alinhamento entre a aplicação das TI e as operações (negócio) do NSC. Assim, para assegurar que as TI foram concebidas e encontram-se aplicadas para suportar os objectivos da organização, a auditoria propôs-se verificar as seguintes condições no Banco NSC:

i. Existência de uma estrutura responsável pela governação de SI e

ii. Existência de uma estrutura responsável pela função de SI na organização, ou seja, pela gestão de SI.

Para se dar corpo ao trabalho, decidiu-se utilizar o COBIT 5 como referencial de boas práticas de governação e gestão de SI. Para o seu cumprimento elaborou-se um questionário com perguntas fechadas, onde os respondentes deveriam apenas responder ou Sim ou Não (S/N) de acordo com as questões colocadas, que foi enviado a quadros da organização, quer das áreas de negócio, quer das áreas de SI. Antes do envio aos respondentes, o questionário foi pré-validado por um grupo de especialistas em auditoria de Tecnologias e Sistemas de Informação para se certificar que as questões apresentadas eram claras e relevantes, tendo o grupo de especialistas, depois de algumas alterações e melhorias, dado o seu parecer favorável.

As temáticas das questões inseridas no questionário, formuladas com base no COBIT 5, são aquelas que cabem à estrutura de governação de SI assegurar e as que cabem à gestão de SI promover, conforme se ilustra no Quadro 1.

ESTRUTURA DE GOVERNAÇÃO DE SI	
A organização tem definida de forma clara a estrutura de governação de SI?	S/N
Existem documentos formais que mandatam e definem as responsabilidades desta estrutura de governação de SI?	S/N
Os processos de governação de SI aplicam-se a todas as TI da organização?	S/N
A estrutura de governação de SI é integrada por gestores de topo, gestores de negócio e gestores de SI?	S/N
Os membros da estrutura de governação de SI estão cientes das suas responsabilidades e participam regularmente nas reuniões deste órgão?	S/N
As actas das reuniões desta estrutura existem e estão acessíveis aos seus membros?	S/N
A estratégia de SI é discutida e aprovada ao nível da estrutura de governação de SI?	S/N
As metas de desempenho e as métricas e níveis de serviço para as TI são definidas e monitorizadas pela estrutura de governação de SI?	S/N
Os programas e orçamentos gerais de TI são discutidos e aprovados ao nível da estrutura de governação de SI?	S/N
ESTRUTURA DE GESTÃO (FUNÇÃO) DE SI	
A organização tem definida de forma clara a estrutura de gestão de SI?	S/N
Essa estrutura está concebida para apoiar os objectivos organizacionais?	S/N
Essa estrutura assegura a responsabilização e o estabelecimento de políticas para a tomada de decisão no âmbito dos SI?	S/N
Está definido um *staff* para a função SI?	S/N
As funções de cada gestor de SI estão definidas e documentadas?	S/N

Existem métricas definidas para medir o desempenho das TI (manuais de organização, procedimentos, acordos de níveis de serviço, etc.)?	S/N
O orçamento de TI e o estado dos projectos de TI são monitorizados e reportados à estrutura de governação de SI?	S/N
Existem manuais de procedimentos e serviços de SI que fornecem detalhes razoáveis para orientar as operações e os operadores?	S/N
Os manuais de procedimentos e serviços são actualizados regularmente?	S/N

Quadro 1: Questões-Âncora sobre Governação e Gestão de SI
Adaptado de ISACA [2012]

No início dos trabalhos, realizou-se uma reunião formal com quadros da organização para aferir o seu conhecimento acerca das necessidades de SI para a execução das suas actividades operacionais. Para a área de SI foi feita a primeira reunião onde o objectivo foi compreender a estrutura e funções da DTSI. Foram colocadas várias questões tais como: saber como o departamento está organizado (isto é, quantas áreas, quantas pessoas, a quem reporta), quais os objectivos (estratégicos e operacionais) definidos para o departamento, quais os principais riscos associados ao cumprimento dos objectivos estabelecidos, quais os principais relatórios que utiliza para analisar o progresso da actividade/negócio face aos objectivos definidos, quais os processos-chave da DTSI e quais as entidades de que depende a DTSI para assegurar os seus objectivos.

Para as áreas de negócio as questões colocadas visaram apurar quais os principais sistemas e tecnologias de informação de suporte às suas actividades, como é que a DTSI tem respondido às necessidades de sistemas e tecnologias de informação das áreas, entre outas questões.

Paralelamente, no terreno verificou-se, a partir de relatórios documentais que a DTSI tem elaborado e enviado ao Presidente do Conselho de Administração, de actas de reuniões internas da área, de manuais de procedimentos e de organização e de testes substantivos, se as práticas de governação e de gestão de SI no Banco NSC estavam de acordo com as boas práticas recomendadas pelo COBIT 5. Os testes substantivos foram executados para aferir se os processos e os procedimentos de execução das actividades estão desenhados para prevenir ou detectar numa base regular erros materiais.

Para a recolha dos dados foram utilizadas as técnicas de observação, entrevistas semi-estruturadas e recolha de documentos.

5. Resultados

De acordo com o COBIT 5, a governação de SI assegura que as necessidades, as condições e as opções das partes interessadas são analisadas para determinar objectivos balanceados e de comum acordo a serem alcançados, estabelecendo direcção através de priorização e tomada de decisão.

A estrutura de governação de SI deve ser formada por executivos, gestores-chave do negócio e equipa de gestão do DTSI, que deverá reunir-se regularmente para rever as iniciativas e projectos de TI e priorizar, alinhar e reconciliar esses projectos com os planos de negócio e a orientação estratégica do Banco NSC.

A estrutura de governação de SI dever ter também a responsabilidade de assegurar que o desempenho das TI está em linha com as expectativas do negócio, constituindo os Acordos de Nível de Serviço (SLA's)[1] uma base para os níveis e os resultados esperados dos serviços de TI.

A gestão de SI, por sua vez, planeia, cria, executa e monitoriza as actividades de TI em alinhamento com a direcção definida pela Administração e emanada da governação de SI para atingir os objectivos organizacionais [ISACA 2012].

Assim, a organização deverá definir claramente uma estrutura de governação de SI, mandatando e definindo formalmente as respectivas responsabilidades. Por sua vez, esta estrutura deverá assegurar que:

- Os processos de governação de SI se aplicam a todas as TI da organização;
- Na sua composição estão integrados gestores de topo, gestores de negócio e gestores de SI;
- Os seus membros estão cientes das suas responsabilidades e participam regularmente nas reuniões do órgão;
- As actas das reuniões desta estrutura existem e estão acessíveis aos seus membros;
- No seu âmbito sejam discutidas alternativas estratégicas para o sistema de informação, consubstanciando-se, posteriormente, na formulação e aprovação da estratégia de SI;
- Na sequência das suas determinações estratégicas para o sistema de informação, defina e monitorize metas de desempenho e métricas de alto nível, bem como balizas para os níveis de serviço a observar pelas TI aplicadas;
- Ainda em articulação com a definição estratégica, estabeleça os programas e orçamentos gerais para a função SI e, concomitantemente, para as TI da organização.

Concluídos os trabalhos a que se propôs, a auditoria não identificou o envolvimento formal e estruturado das áreas de negócio na governação de SI. A auditoria verificou que não está instituída

[1] Do Inglês, Service Level Agreement

uma estrutura de governação de SI. A função SI reporta directamente ao Presidente do Conselho de Administração. Este por sua vez apresenta as questões de TI ao Conselho de Administração.

Também não foi identificado o envolvimento estruturado das áreas de negócio na definição da estratégia de SI, na priorização de projectos de negócio e na monitorização e avaliação dos níveis de serviço de TI.

A auditoria não encontrou evidências de quaisquer outras reuniões formais e regulares para discutir a estratégia de SI e as necessidades de negócio que envolvam as áreas de negócio do banco.

De igual modo, a auditoria não identificou reuniões formais regulares realizadas entre as áreas de negócio e a equipa de gestão de SI para discutir as necessidades de negócio, não estando em vigor qualquer processo formal para envolver o DTSI quando os projectos com componentes de TI são iniciados.

À gestão de SI cabe o planeamento e execução das iniciativas de SI, alinhando-as com os direccionamentos recebidos. Logo, a função SI deverá:

- Estar concebida para apoiar os objectivos organizacionais;
- Assegurar a responsabilização e o estabelecimento de políticas para a tomada de decisão;
- Possuir um *staff* perfeitamente definido;
- Garantir que as funções de cada gestor de SI se encontram definidas e documentadas;
- Garantir a existência de métricas para medir o desempenho das TI (manuais de organização, procedimentos, acordos de níveis de serviço, etc.);
- Monitorizar e reportar a execução do seu orçamento e o estado dos projectos de TI à estrutura de governação de SI;
- Assegurar a existência de manuais de procedimentos e serviços de SI que orientem as operações e os operadores conexos ao sistema de informação;
- Assegurar que os manuais de procedimentos e serviços são revistos e actualizados regularmente.

Relativamente à função SI, embora não exista uma estrutura formal de governação de SI, pode-se verificar que a função SI no Banco NSC se encontra formalmente constituída e documentada. Especificamente, encontraram-se evidências de definição formal do pessoal de SI e dos processos conexos à função SI, bem como da atribuição de responsabilidades, papéis e relacionamentos no âmbito da função SI. Estão também definidas políticas de circulação, divulgação e acesso a informação, as quais foram aprovadas pelo Conselho de Administração e divulgadas para toda a organização.

A auditoria observou que as políticas e manuais de procedimentos não estão documentados para alguns serviços de SI. As tarefas conexas aos SI são realizadas com base nos conhecimentos e experiência do pessoal, sem orientação de manuais de procedimentos ou políticas, criando-se, portanto, uma dependência excessiva relativamente a funcionários com conhecimentos detalhados sobre esses processos. No caso de uma saída ou ausência de pessoal-chave, não há nenhuma orientação documentada sobre como devem ser executadas as funções.

Verificou-se, também, que não estão definidos acordos de níveis de serviço (SLA's) entre o DTSI e as áreas de negócio e não há indicadores de prestação de contas reportados e controlados de forma regular. Para algumas funções e serviços de TI não existem manuais de procedimentos que forneçam indicações que orientem as operações e os operadores.

A título de síntese, apresenta-se no Quadro 2 as posições consolidadas nas temáticas subjacentes às questões-âncora previamente apresentadas para o banco NSC, mediante a indicação do grau de satisfação por parte da instituição nas vertentes da governação e da gestão de SI examinadas.

ESTRUTURA DE GOVERNAÇÃO DE SI	
A organização tem definida de forma clara a estrutura de governação de SI.	Não satisfaz
Existem documentos formais que mandatam e definem as responsabilidades desta estrutura de governação de SI.	Não satisfaz
Os processos de governação de SI aplicam-se a todas as TI da organização.	Não satisfaz
A estrutura de governação de SI é integrada por gestores de topo, gestores de negócio e gestores de SI.	Não satisfaz
Os membros da estrutura de governação de SI estão cientes das suas responsabilidades e participam regularmente nas reuniões deste órgão.	Não satisfaz
As actas das reuniões desta estrutura existem e estão acessíveis aos seus membros.	Não satisfaz
A estratégia de SI é discutida e aprovada ao nível da estrutura de governação de SI.	Não satisfaz
As metas de desempenho e as métricas e níveis de serviço para as TI são definidas e monitorizadas pela estrutura de governação de SI.	Não satisfaz
Os programas e orçamentos gerais de TI são discutidos e aprovados ao nível da estrutura de governação de SI.	Não satisfaz
ESTRUTURA DE GESTÃO (FUNÇÃO) DE SI	
A organização tem definida de forma clara a estrutura de gestão de SI.	Satisfaz
Essa estrutura está concebida para apoiar os objectivos organizacionais.	Satisfaz
Essa estrutura assegura a responsabilização e o estabelecimento de políticas para a tomada de decisão no âmbito dos SI.	Satisfaz
Está definido um *staff* para a função SI.	Satisfaz
As funções de cada gestor de SI estão definidas e documentadas.	Satisfaz

Existem métricas definidas para medir o desempenho das TI (manuais de organização, procedimentos, acordos de níveis de serviço, etc.).	Não satisfaz
O orçamento de TI e o estado dos projectos de TI são monitorizados e reportados à estrutura de governação de SI.	Não satisfaz
Existem manuais de procedimentos e serviços de SI que fornecem detalhes razoáveis para orientar as operações e os operadores.	Satisfaz parcialmente
Os manuais de procedimentos e serviços são actualizados regularmente.	Satisfaz parcialmente

Quadro 2: Síntese da Avaliação da Governação e Gestão de SI do Banco NSC

6. Recomendações

As TI podem contribuir significativamente para que as organizações alcancem os seus objectivos, mas para isso é preciso tornar a governação de SI mais ágil e clara na sua implantação e utilização, especialmente no que se refere aspectos tais como, equilíbrio dos riscos; controlo dos custos, de pessoas, de contractos e de fornecimento de serviços de terceiros; além de ficar claro como são tomadas as decisões e quem as toma [Mendonça et al 2013], tornando evidente que há um alinhamento entre as TI e o negócio das organizações.

Para ultrapassar as inconformidades verificadas na auditoria realizada no Banco NSC e por forma a potenciar o alinhamento entre o sistema de informação e o negócio do Banco, com base no procedimento de avaliação levado a cabo, avançaram-se as seguintes recomendações, sob a forma de tarefas que a instituição deverá realizar a breve trecho:

- Definir de forma clara a estrutura de governação de SI;
- Elaborar e divulgar os documentos formais que mandatam e definem as responsabilidades desta estrutura de governação de SI;
- Criar processos de governação de SI e aplicá-los a todas as TI da organização;
- Discutir e aprovar a estratégia de SI ao nível da estrutura de governação de SI;
- Discutir e aprovar os programas e orçamentos de TI ao nível da estrutura de governação de SI;
- Estabelecer métricas para medir o desempenho das TI;
- Definir e monitorizar as metas de desempenho e as métricas e níveis de serviço das TI;
- Monitorizar e reportar a execução orçamental e o estado dos projectos à estrutura de governação de SI por parte da estrutura de gestão de SI.

7. Considerações Finais

A temática do alinhamento entre os SI e o negócio das organizações tem sido estudada por diversos autores. As evidências teóricas e empíricas sugerem que um melhor alinhamento entre o sistema de informação e o negócio potencia maiores níveis de desempenho organizacional. Além disso, o alinhamento proporciona a visualização das informações estratégicas, bem como posiciona a área de SI e os sistemas de informação como elementos chave nas organizações [Affeldt e Vanti 2009].

Para assegurar que as TI estão concebidas para suportar os objectivos da organização é importante que haja um comprometimento entre a gestão de topo, as áreas de negócio e a área de SI (função SI), para que aquelas entendam o papel e potencial das TI no suporte do negócio e da organização e a função SI saiba quais os serviços que deve entregar ao negócio por forma a garantir que as necessidades do negócio em relação às TI estão a ser satisfeitas. Para tal, é importante que exista uma estrutura formal com responsabilidade e mandato para definir a estratégia de SI no suporte ao negócio, alinhando-a com os objectivos da organização.

A governação de SI assume, assim, um papel fundamental ao consubstanciar as actividades daquela estrutura formal, devendo ser parte integrante da gestão organizacional, liderando e criando as estruturas e processos organizacionais que asseguram que as TI sustentam as estratégias e os objectivos da organização.

Em consequência e em complementaridade com a governação de SI, a actividade de gestão de SI visará a concretização do alinhamento entre a aplicação das TI e o negócio, balizada pelas directivas emanadas da governação de SI.

Neste trabalho constatou-se que o Banco NSC não dispõe de uma estrutura de governação de SI que garanta que as necessidades de TI para suportar o negócio estão a ser bem abordadas e que a função SI está a prestar os serviços requeridos pelas áreas de negócio. Daqui decorre que não se pode assegurar que as TI estejam concebidas para suportar os objectivos e a estratégia do Banco NSC. Para além das conclusões imediatas do procedimento de auditoria, a avaliação efectuada no Banco NSC permitiu avançar com um conjunto de recomendações para a institucionalização da actividade de governação de SI, melhoria da operação da função SI, ou seja, da actividade de gestão de SI, e articulação entre a governação e a gestão de SI no NSC.

Para a avaliação das actividades de governação e gestão de SI delineou-se um procedimento de auditoria fundado no COBIT 5 e que se crê poder constituir o primeiro momento da aferição do alinhamento entre a aplicação das TI e o negócio.

Como trabalho futuro identifica-se o interesse em delinear o segundo momento de aferição do alinhamento: concluindo-se sobre a existência e qualidade das actividades de governação e gestão

de SI, importará medir o grau com que as TI suportam os objectivos e a estratégia do negócio. Posteriormente, será possível lançar-se estudos com vista a apurar o relacionamento entre o grau de alinhamento entre o sistema de informação e o negócio de uma organização e a qualidade das suas actividades de governação e gestão de SI e, assim, aferir sobre a eventual suficiência destas duas actividades para o alcance do dito alinhamento.

Agradecimentos

Este trabalho foi apoiado pela FCT – Fundação para a Ciência e Tecnologia no âmbito do Project Scope PEst-OE/EEI/UI0319/2014.

REFERÊNCIAS

Affeldt, F.S. e Vanti, A.A., (2009). Alinhamento Estratégico e Tecnologia da Informação: Análise de Modelos e Propostas para Pesquisas Futuras. *Revista de Gestão da Tecnologia e Sistema de Informação* 6(2), 203–226.

Belfo, F. and Sousa, R.D. (2012). A Critical Review of Luftman's Instrument for Business-IT Alignment. *MCIS 2012 Proceedings*. Paper 6.

Buckby, S., Best, P. and Stewart, J. (2005) The Role of Boards in Reviewing Information Technology Governance (ITG) as part of Organizational Control Environment Assessments. In Cusack, B., Eds. Proceedings 2005 IT Governance International Conference, pp. 1-14, Auckland, New Zealand.

Calle Jr, E, and Kanter, J. (1998). Aligning Information Systems and Business Strategy. A Case study. *Journal of Information Technology Management* IX(1).

Campbell, B., Kay, R., and Avison, D. (2005). Strategic alignment: a practitioner's perspective. *Journal of Enterprise Information Management* 18(6), 653-664.

Chan, Y.E., and Reich, B.H. (2007). IT alignment: what have we learned? *Journal of Information Technology* 22 (4), 297-315.

Chan, Y.E., Sabherwal, R. and Thatcher, J.B. (2006). Antecedents and Outcomes of Strategic IS alignment: An Empirical Investigation. *IEEE Transactions on Engineering Management* 53(1), 27-47.

Croteau, A. and Bergeron, F. (2001). An Information Technology Trilogy: Business Strategy, Technological Deployment and Organizational Performance. *Journal of Strategic Information Systems* 10, 77–79.

Haes, S. and Van Grembergen, W. (2009). An Exploratory Study into IT Governance Implementations and Its Impact on Business/IT Alignment. *Information Systems Management* 26(2), 123–137.

Henderson, J. and Venkatraman, N. (1993). Strategic Alignment: Leveraging Information Technology for Transforming Organizations. *IBM Systems Journal* 32(1), 4-16.

ISACA (2012). COBIT 5 – A Business Framework for the Governance and Management of Enterprise IT, ISACA.

Kaplan, R. and Norton, D. (2006). *Alignment: Using the Balanced Scorecard to Create Corporate Synergies*. Boston: Harvard Business School Publishing Corporation.

Kappelman, L., McLean, E., Luftman, J. and Johnson, V. (2013). Key Issues of IT Organizations and Their Leadership: The 2013 SIM IT Trends Study. *MIS Quarterly Executive* 12(4), 227-240.

Kappelman, L., McLean, E., Johnson, V. and Gerhart, N. (2014). The 2014 SIM IT Key Issues and Trends Study. *MIS Quarterly Executive* 13(4), 237-263.

Kearns, G.S., and Lederer, A.L. (2003). A Resource-Based View of Strategic IT Alignment: How Knowledge Sharing Creates Competitive Advantage. *Decision Sciences* 34 (1):1-29.

Laurindo, F., Shimizu, T., Carvalho, M. e Rabechini Jr. R. (2001). O Papel da Tecnologia da Informação (TI) na Estratégia das Organizações. *Gestão e Produção* 8(2), 160-179.

Luftman, J. (2000). Assessing Business-IT Alignment Maturity. *Communications of Association of Information Systems 4*, Article 14.

Luftman, J., and Derksen, B. (2012). Key Issues for IT Executives 2012: Doing More with Less. *MIS Quarterly Executive* 11(4), 207-218.

Mendonça, C., Guerra, L., Neto, M. e Araújo, A, (2013). A Governança de Tecnologia da Informação: Um Estudo do Processo Decisório em Organizações Públicas e Privadas. *Revista Administração Pública – Rio de Janeiro* 47(2), 443 – 468.

Pascutti, M., Modesto, L., Santos, T. e Aquino, W. (2009). Governança de Tecnologias de Informação: Um Estudo de Caso de Micro e Pequenas Empresas na Cidade de Apucarana, *Revista F@pciencia* 9, 89-98.

Peterson, R. (2003). Integration Strategies and Tactics for Information Technology Governance, in *Strategies for Information Technology Governance*, Ed. Wim Van Grembergen, Idea Group Publishing.

Pitassi, C. e Leitão, S. (2002). Tecnologia de Informação e Mudança: Uma Abordagem Crítica, *RAE – Revista de Administração de Empresas* 42(2), 77–87.

Reddy, G., Srinivasi, R., Rikkula, S. e Rao, V. (2009). Management Information Systems to Help managers for Providing Decision Making in an Organization. *International Journal of reviews in Computing*.

Sabherwal, R., and Chan, Y.E.. (2001). Alignment between business and IS strategies: A study of prospectors, analyzers, and defenders. *Information Systems Research* 12 (1), 11-33.

Tan, F.B., and Gallupe, R.B. (2006). Aligning business and information systems thinking: a cognitive approach. *IEEE Transactions on Engineering Management* 53 (2), 223-237.

Teo, T.S.H. and Ang, J.S.K. (1999). Critical success factors in the alignment of IS plans with business plans. *International Journal of Information Management* 19, 173-185.

Wang, J. W., F. Gao e Ip, W.H. (2010). Measurement of Resilience and Its Application to Enterprise Information System. Enterprise Information Systems 4(2), 215-223.

Weill, P. and Ross, J. (2004). *IT Governance: How Top Performers Manage IT Decisions Rights to Superior Results*. Boston: Harvard Business School Press.

Organizations and autopoiesis: a biological perspective

Ana Sara Costa[1], José Tribolet[2,3]

1) CEG-IST, Department of Engineering and Management, Instituto Superior Técnico, Universidade de Lisboa, A. Rovisco Pais, 1, 1049-001, Lisbon, Portugal

anasaracosta@tecnico.ulisboa.pt

2) Department of Computer Science and Engineering, Instituto Superior Técnico, Universidade de Lisboa, A. Rovisco Pais, 1, 1049-001, Lisbon, Portugal

3) INESC, Rua Alves Redol 9, Lisbon, Portugal

jose.tribolet@inesc.pt

Abstract

Organization studies have recently had plentiful new approaches and paradigms, such as the autopoiesis theory. Maturana and Varela developed the theory of autopoiesis − a biological theory of autopoietic systems − in order to explain the particular nature of living in the physical domain. Their main contribution has been to provide a concise specification of the defining characteristics of living systems, including humans. Several authors, especially the sociologist Luhmann, have tried to extend such theory to social systems, including organizations. Social systems are systems that are based on the interactions of living systems. The question therefore arises if organizations are also autopoietic systems. In this paper, this question and the relevance of the autopoiesis theory are discussed for understanding organizational changes and self-maintenance, from a biological point of view. Supported by the autopoiesis theory, organizations can be interpreted as networks of interactions, processes and rules of coordination.

Keywords: Organizations, Organizational Engineering, Autopoiesis, Autopoietic System, Self-maintenance.

1. Introduction

In recent years, new paradigms and approaches for organization studies have been developed. The general understanding of organizations (or enterprises) has been studied from a biological perspective, based specially on the theory of autopoiesis. However, the application of this theory, which was originated in biology, has been controversial. Hence, there is no clear agreement among organization scholars regarding the role or the place of autopoiesis in organization science.

Autopoiesis is a concept developed by two Chilean biologists, Humberto Maturana and Francisco Varela [1980], to denote a form of system organization where the system as a whole produces and replaces its own components, and differentiates itself from its environment. Their concern was not the explanation of the social systems or organizations, but rather to explain what is a living being. The main contribution of Maturana and Varela's [1980] work about

autopoiesis was to present an understanding of the defining characteristics of living systems [Kay 2001].

Autopoiesis has been an extensive subject of debate and has been applied in several areas, such as sociology [Luhmann 1990], management [Hall 2003] and law [Teubner 1993; Teubner and Febbrajo 1992], among others. The application of this perspective to social systems is really an ongoing topic of debate [Kay 2001].

In recent years, several papers regarding autopoiesis in organizations studies have been published. Many authors have adopted such perspective to study organizations applying Maturana and Varela's work, as Luhmann, for example [cf. Luhmann, 1995]. Although, several authors have taken autopoiesis straight from the area of the biological sciences to organization studies, others have applied it to the qualified approach of Luhmann, who developed the social systems theory, or have even combined autopoiesis with other approaches [Magalhães and Sanchez 2009]. Some authors are reluctant in applying autopoiesis to social systems, since Maturana and Varela wrote that autopoiesis is not a social theory. In the literature, Luhmann's social systems theory is identified as the predominant adopted approach of autopoiesis to social systems [Fuchs and Hofkirchner 2009].

Aveiro [2010] considers there is a lack of concepts and methods in Organizational Engineering (OE) for a continuous real-time management of organizational changes, through a continuous update of organizational models. In organizations, a large amount of time is wasted in dealing with unexpected exceptions, essentially because organizational models are not current or coherent with reality [Aveiro 2010]. The knowledge of organizational performance in real-time makes possible to empower all the human actors to strategically deal with organizational changes. Organizations are referred to as socio-technical systems. The active resources of an organization are: carbon-based actors (humans) and silicon-based actors (computers). In order to survive, organizations must adapt continuously to a changing environment. Emergence and unexpected exceptions are controversial when applied to organizations, due to the presence of cognitive agents. The mechanisms of emergence within organizations, which are social systems, can be expected to be different from those present in other natural systems. The application of the autopoiesis theory could explain this difference. There is also the need for a better understanding of the Organization as well as the Organizational Self-Awareness [Magalhães et al. 2008].

This paper discusses a new understanding of organizations from a biological perspective and contributes to answering the question: Are organizations autopoietic systems?

The rest of this paper is organized as follows. Section 2 is devoted to the autopoiesis theory, which outlines the basic concepts and main developments of autopoiesis, while section 3 applies them to organizations. Finally, the last section presents some concluding remarks and suggestions for future research.

2. Autopoiesis

Maturana and Varela [1980] proposed the concept of autopoiesis (from Greek: *autos* = self, *poiein* = to produce) as a definition of life. This theory was developed in order to provide an understanding of the nature and characteristics of living systems, namely biological cells and metacellular organisms.

The main idea of autopoiesis is that living systems are characterized by their self-production, which means that the components of the system produce the components of the system. Therefore, the requirements for the maintenance of self-production limit the way in which individuals can interact with their environments. All components and processes of the system produce the same components and processes, which establishes an autonomous and self-producing entity [Mingers 1995]. According to Maturana and Varela [1992], autonomy is the key concept of living beings and is the ability to specify what is proper to them. Moreover, these systems can only be characterized with reference to themselves and whatever takes place in them is necessarily determined in relation to themselves, i.e, self-referentiality. Therefore, the mechanism that makes living systems autonomous is autopoiesis.

An autopoietic system exists as a network of relations and processes, which continuously produces the components that realize such network as a concrete unity. An organization can be viewed as a distributed network of carbon and silicon actors, dynamically evolving in real-time.

A significant implication of autopoiesis is the distinction between organization and structure:

- Organization is the relation between the components of the system and determines the identity of the system;
- Structure is the realization of a system's organization in a specific realization space [Maturana and Varela 1980].

The system constantly reproduces its structure and organization, and the continuous reproduction of the system's structure and organization is fundamental to living systems. Maturana and Varela [1980] identified this phenomenon, autopoiesis, and defined it as the

ability of a system to generate its specific constitution and its components (structure), and their interplay (organization) on its own. For instance, by analogy, an algorithm for solving a certain problem can be viewed as a description of the system's organization. From this viewpoint, the corresponding computer program can be viewed as the realization of this organization, i.e.,, the structure in a certain space. In this case, such structure is the programming language [Abou-Zeid 2001]. Therefore, all systems of a particular type have the same organization, but they can have different structures, and the structure may change without changing the organization. Hence, all living systems have the same autopoietic organization realized in an infinity of different structures. Zeleny [2005] interprets an organization as a network of interactions and processes, which are identified by its rules of coordination – organization – and differentiated by applying such rules of coordination under different conditions – structure.

The model of autopoiesis presented by Maturana and Varela is presented in Figure 1. The network of operations of a living system is closed, i.e., the system is operationally closed, and in its interaction with its surrounding environment it is structure-determined.

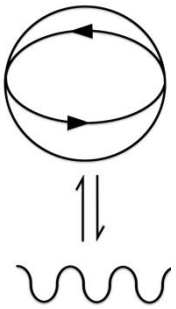


Figure 1 - Model of autopoiesis according to Maturana and Varela

From an autopoietic viewpoint, an individual's behaviour is determined by particular states of nervous system activity, which is defined by the concept of operational closure [Maturana and Varela 1980]. This presupposes that in all cases nervous system activity works in a closed cycle [Maturana and Varela 1980]. An environmental perturbation may act as a trigger for change, but the nervous system's structure always dictates which forces can be a trigger [Mingers 1991].

Autopoiesis generates a structural coupling (see Figure 2) with the environment: the structure of the nervous system of an organism generates patterns of activity, which are triggered by perturbations from the exterior, thereby contributing to the continuing autopoiesis of the organism. In other words, an autopoietic system can be viewed as a network of transformation and destruction processes and its components interact to constantly regenerate such network.

Therefore, autopoiesis is necessary and sufficient to characterize a living system [Maturana and Varela 1980].

Figure 2 shows the metabolic pathways of two autopoietic systems (circles/spheres) in the context of their surrounding environment. The reciprocating arrows between them show an interdependence that has developed over time. The structure of both systems and environment change as a result of mutual non-destructive perturbations.

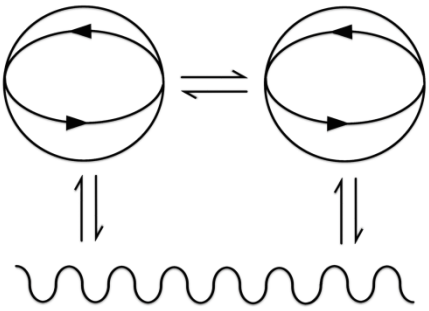


Figure 2 - Structural Coupling

Systems theory identifies that processing is the life function of all living systems, which is the act of sequentially selecting actions. Moreover, the selection can be considered as the fundamental operation enabling dynamic processing, and the potential actions are the basic building blocks dynamically chained to a process, as presented in Figure 3 [Thannhuber 2001].

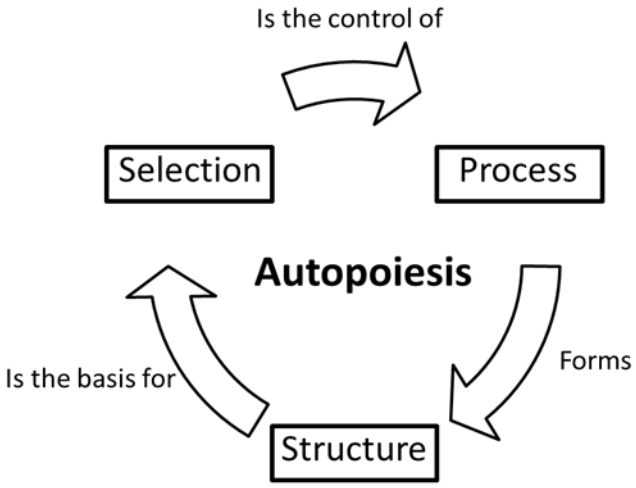


Figure 3 - Autopoiesis (adapted from Thannhuber et al. [2001])

Varela et al. [1974] enumerated six criteria, which have been discussed by other researchers (e.g. Hall et al. 2005) considered necessary and sufficient to recognize when a system could be considered to be autopoietic, as follows:

1. Bounded: the unity has identifiable boundaries;
2. Complex: the system is comprised of a set of components;
3. Mechanistic: the unity is a mechanistic system, which proprieties are generated by components' interrelations and not by components' individual proprieties;
4. Self-differentiated: the components forming the system's boundaries are intrinsically determined by their properties in the space of their interactions;
5. Self-producing: the unity generates the components of this boundary, either from existing components or by coupling of non-component elements from its environment;
6. Autonomous: all the unity's components participate in the production of its components.

If all these criteria are fulfilled, then the system is considered as being autopoietic.

3. Organizations and Autopoiesis

There has been considerable debate regarding the application of the autopoiesis theory to social systems, including organizations. Indeed, the theory of autopoiesis provides ways of thinking about mechanisms of organizations.

In the literature, few authors incorporate the theory of autopoiesis directly on Maturana and Varela's original work, even as others have adopted Luhmann's social system theory [Goldspink and Kay 2009]. Additionally, other approaches have also been combined with the autopoiesis theory. For instance, the autopoiesis theory combined with the complexity theory has demonstrated profound implications about organizations [Kay and Goldspink 2009].

Nicklas Luhmann, a German sociologist, extended this concept of autopoiesis to establish a theory of social systems, which is based on the idea that human social systems were formed by recursive networks of communications. Luhmann's work of the theory of autopoiesis in social systems is a thorough and ambitious reworking of the autopoiesis theory. According to Luhmann [1995], the basic elements of social systems are not people but communications. This perspective has the advantage of describing the system in terms of its operational characteristics, independent from the specific participants in such system as well as from time. We consider the exclusion of humans from social systems as the main problem of such theory.

The autopoiesis theory gives a comprehensive of living beings from the simplest unicellular organisms (first-order unities) to the most complex multi-cellular organisms (second-order unities), and it extends beyond organisms toward “social systems” of animals or people (third-order unities). Social or business organizations are examples of third-order unities [Beeson 2009]. When applying the autopoiesis theory to organizations, we are considering the organization itself as a unit system to which autopoiesis might be ascribed.

Morgan [1997] identifies three main features of Maturana and Varela's theory of living systems – autonomy, circularity, and self-reference – and suggests that organizations can be seen as attempting to turn their environments into extensions of their own identity, which implies a focus on establishing the organization's identity.

Fuchs and Hofkirchner [2009] argue that social systems (including organizations) are autopoietic. However, they suggest an understanding that is human-centred and that humans permanently create the unity of human actors and social structures (human sociality). This means that their interpretation of social systems and autopoiesis departs from Luhmann's social system theory.

Organizations must to be able to produce their specific organizational dynamics, as a process of the adjustment of internal chaotic dynamics to the chaotic dynamics of the environment. This crucial survival process as a self-organizing ability is essential to a sustainable organization. An organization must guarantee that it has regulation capabilities to deal with exceptions, in order to have control and self-maintenance, and to be a self-sustainable system. When dealing with organizational complexity, managers need to understand what kind of strange attractors are naturally propelling the autopoiesis in their organizations as well as to find out what kind of activity can attract and concentrate the energy of their employees.

External perturbations from the system allow for its structure changes as long as these do not affect its operations, and guarantee the system’s identity and closure. Biggiero [2001] argues that inputs cannot be distinguished from perturbations, and organizations’ boundaries are not stable and independent, thereby such a perspective is not applicable to organizations.

If all the criteria presented in Section 3 are fulfilled, then the system could be considered to be autopoietic. However, an organization as a complex socio-technical system does not fulfil all of such criteria. The boundary of an organization is its unique language, which allows to distinguish an organization itself from its environment through a linguistic space. An organization is a complex and dynamic network of two kinds of active components: carbon-based actors (humans), and silicon-based actors (computers). These components communicate through the interactions among them. Therefore, we can consider that an

organization is bounded, complex and mechanistic. System boundaries are internally determined in interaction with their surrounding environment. Nevertheless, organizations do not evidently produce their own components and boundaries through the actions of its own components. Most organizations import ready-made components from its environment. From this viewpoint, organizations are not autopoietic systems.

We consider acceptable to apply autopoiesis in a metaphorical way to other systems, and in particular, to organizations. Many concepts of the autopoiesis paradigm may be useful and applicable to study social systems and organizations, such as self-organization, self-maintenance, self-reference, autonomy, operational closure and structural coupling.

Another relevant application of the Maturana and Varela's work, in the context of organizational studies, is the humans' capacity for reflexive self-awareness. Organizational Self-Awareness is of utmost relevance on OE. An organization is self-aware when all of its actors maintain a real-time synchronization of their individual organizational models and share the view of the "whole", which supports the organizational knowledge, allowing the actors to make better decisions and increase their autonomy [Magalhães and Tribolet 2006].

Integrating autopoiesis with systems theory through the principles of dynamic systems is another approach to study of organizations as systems. Regarding the dynamic aspects of an organization, we can consider three feedback loops: (1) Feedback loop 1: stable model; (2) Feedback loop 2: model configuration change owing to a changing environment; and (3) Feedback loop 3: need for a new and suitable model due to a radical change [Tribolet et al. 2014]. In organizations, the constant change and continuous adaptation require a representation that allows to capture the overall picture of the system as an unit in all its dimensions and complexity. In this sense, an organization maintains its capability to govern itself, static and dynamically, through an informed and intelligent action of the actors, based on the Enterprise Cartography [Tribolet et al. 2014].

4. Conclusions

The theory of autopoiesis has been demonstrated to have philosophical implications for our thinking about organizations. Despite its popularity, autopoiesis theory has not achieved the expected profile in the field of organization studies.

The main contribution of this work is theoretical and it helps to answer a long standing controversial question among researchers: Are organizations autopoietic? Concisely, our position is that, applying directly the autopoiesis theory, organizations themselves cannot be considered autopoietic systems. However, we consider that many concepts of the autopoiesis

theory are useful to explain the operating system of organizations, and the autopoiesis can be metaphorically applied.

This work suggests that autopoiesis may be used as a backdrop for a new organizational and management paradigm. Hence, autopoiesis also appears to introduce a multidisciplinary approach to management by introducing biological thinking. The theory of autopoiesis can really guide managers and provide a better understanding about operations of their organizations, in order to avoid wasting time and resources.

Further work is necessary regarding autopoiesis and organizations. Besides more theoretical research is needed in this direction, autopoietic concepts can contribute in a practical way to organization management. Moreover, we consider that autopoiesis and OE have several concepts that can answer to organizational challenges. Applying the autopoiesis metaphor and the concepts of OE it is possible to develop an adequate organizational self-maintenance's framework that can make adjustments to quickly adapt to organizational perturbations, in order to maintain the control of the organization when such changes occur. We consider that the conception, implementation and maintenance of information systems in organizations should be based in the presented perspective. Additionally, for instance, given its holistic nature, autopoiesis can provide a suitable framework for the integration of IT/IS into organizations. Future research is also needed to study the underlying dynamics of the change process, to better understand how organizations as a system are or should be developed, and how they grow. Then, we can explain the successes and failures of organizations, and also define resilience strategies in order to control perturbations in organizational systems.

5. References

Abou-Zeid, A.-S., "Toward an Autopoietic Approach for Information Systems Development", in M. Rossi and K. Siau (Eds.), Information Modelling in the New Millennium, Idea Group Publishing, Hershey, 2013, pp. 34-52.

Aveiro, D.S.A., G*.O.D. (Generation, Operationalization & Discontinuation) and Control (sub) organizations : a DEMO-based approach for continuous real-time management of organizational change caused by exceptions*, PhD thesis, Universidade Técnica de Lisboa, 2010.

Beeson, I., "Information in Organisations: rethinking the autopoietic account", in R. Magalhães and R. Sanchez (Eds.) Autopoiesis in Organization Theory and Practice, Emerald Group Publishing, Bingley, 2009, pp. 185-200.

Biggiero, L., "Are firms autopoietic systems?", in Van Der Zouwen, G. and Geyer, F. (Eds.), Sociocybernetics: Complexity, Autopoiesis, and Observation of Social Systems, Greenwood, Westport (Ct), 2001, pp. 125-140.

Fuchs, C. and W. Hofkirchner, "Autopoiesis and Critical Social Systems Theory", in R. Magalhães and R. Sanchez (Eds.), Autopoiesis in Organization Theory and Practice, Emerald Group Publishing, Bingley, 2009, pp. 111-129.

Goldspink, C. and R. Kay, "Autopoiesis and organizations: A biological view of organizational change and methods for its study", in R. Magalhães and R. Sanchez (Eds.), Autopoiesis in Organization Theory and Practice, Emerald Group Publishing, Bingley, 2009, pp. 89-110.

Hall, W.P., "Organizational autopoiesis and knowledge management", Presented, ISD '03 Twelfth Int. Conf. Info. Syst. Dev., Melbourne, Aust., 25-27 Aug, 2003.

Hall, W.P. P. Dalmaris and S. Nousala, "A Biological Theory of Knowledge and Applications to Real World Organizations" in Proceedings, KMAP05, Knowledge Management in Asia Pacific Wellington, NZ 28-29 November 2005.

Kay, R., "Are organizations autopoietic: A call for new debate?", *Systems Research and Behavioral Science*, 18, 6, (2001), 461-477.

Kay, R. and C. Goldspink, "Autopoiesis: Building a Bridge between Knowledge Management and Complexity", in R. Magalhães and R. Sanchez (Eds.), Autopoiesis in Organization Theory and Practice, Emerald Group Publishing, Bingley, 2009, pp. 233-242.

Luhmann, N., *Essays on Self-reference*, Columbia University Press, New York, 1990.

Luhmann, N., *Social systems*, Stanford University Press, Palo Alto, CA, 1995.

Magalhães, R. and J. Tribolet, "Engenharia Organizacional: das partes ao todo e do todo às partes na dialéctica entre pessoas e sistemas", in S.G. Costa, L.M. Vieira and J.N. Rodrigues (Eds.). Ventos de Mudança, Editora Fundo de Cultura, Rio de Janeiro, Brasil, 2006.

Magalhães, R. P. Sousa and J. Tribolet, "The Role of Business Processes and Enterprise Architectures in the Development of Organizational", Revista de Estudos Politécnicos, Polytechnical Studies Review, Vol. VI, 9, 2008.

Magalhães, R. and R. Sanchez, "Autopoiesis Theory and Organization: An Overview", in R. Magalhães and R. Sanchez (Eds.), Autopoiesis in Organization Theory and Practice, Emerald Group Publishing, Bingley, 2009, pp. 3-25.

Maturana, H.R. and F.J. Varela, *Autopoiesis and Cognition: the Realization of the Living*, vol. 42, BSPS, Reidel, Dordrecht, 1980.

Maturana, H.R. and F.J. Varela, *The tree of knowledge*. *The biological roots of human Understanding*, Shambhala Publications Inc, Boston, MA, 1992.

Mingers, J., *Self-producing systems*. *Implications and applications of autopoiesis*. Plenum Press, NewYork, 1995.

Mingers, J., "The Cognitive Theories of Maturana and Varela", *Systems Practise*, 4, 4, (1991), 319-338.

Morgan, G., *Images of organization*, Sage Publication, Thousand Oaks, CA, 1997.

Teubner, G., *Law as an Autopoietic System*, Blackwell, Oxford, 1993.

Teubner, G. and A. Febbrajo, *State, Law, and Economy as Autopoietic Systems*, Giuffrè, Milan, 1992.

Thannhuber, M. M.M. Tseng and H.-J. Bullinger, "An Autopoietic Approach for Building Knowledge Management Systems in Manufacturing Enterprises", *Annals of the CIRP*, 50, 1, (2001), 313-318.

Tribolet, J. P. Sousa and A. Caetano, "The Role of Enterprise Governance and Cartography in Enterprise Engineering", *Enterprise Modelling and Information Systems Architectures*, 9, 1, (2014), 38-49.

Varela F.H. Maturana and R. Uribe, "Autopoiesis: the organization of living systems, its characterisation and a model", *Biosystems*, 5, 4, (1974), 187-196.

Zeleny, M., *Human systems management: Integrating knowledge, management and systems*, Singapore: World Scientific Publishing, 2005.

Categorização do Consumo de Energia em Sistemas de Povoamento de Data Warehouses

Miguel Guimarães[1], João Saraiva[2] e Orlando Belo[1]

1) Centro de I&D ALGORITMI, Universidade do Minho, PORTUGAL

pg22800@alunos.uminho.pt, obelo@di.uminho.pt

2) HasLab - Laboratório de Investigação em Software Confiável, Universidade do Minho, PORTUGAL

jas@di.uminho.pt

Resumo

Os recursos computacionais exigidos durante o processamento de grandes volumes de dados durante um processo de povoamento de um *data warehouse* faz com que a necessidade da procura de novas implementações tenha também em atenção a eficiência energética dos diversos componentes processuais que integram um qualquer sistema de povoamento. A lacuna de técnicas ou metodologias para categorizar e avaliar o consumo de energia em sistemas de povoamento de *data warehouses* é claramente notória. O acesso a esse tipo de informação possibilitaria a construção de sistemas de povoamento de *data warehouses* com níveis de consumo de energia mais baixos e, portanto, mais eficientes. Partindo da adaptação de técnicas aplicadas a sistemas de gestão de base de dados para a obtenção dos consumos energéticos da execução de interrogações, desenhámos e implementámos uma nova técnica que nos permite obter os consumos de energia para um qualquer processo de povoamento de um *data warehouse*, através da avaliação do consumo de cada um dos componentes utilizados na sua implementação utilizando uma ferramenta convencional. Neste artigo apresentamos a forma como fazemos tal avaliação, utilizando na demonstração da viabilidade da nossa proposta um processo de povoamento bastante típico em *data warehouses* – substituição encadeada de chaves operacionais -, que foi implementado através da ferramenta Kettle.

Palavras chave: Povoamento de *Data Warehouses*, Avaliação do Consumo de Energia em Componentes de Sistemas de ETL, Categorização de Consumos de Energia, Substituição Encadeada de Chaves Operacionais, Kettle.

1. Introdução

Correntemente, e em grande parte devido à mudança de paradigmas, a preocupação por questões energéticas tem vindo a acentuar-se. O aumento do preço da eletricidade bem como as questões relacionadas com o ambiente e a sua preservação, criaram a necessidade de se conceber componentes de software e de hardware energeticamente eficientes. Hoje, por exemplo, o fator de desempenho ou a eficiência de execução deixaram de ser, em exclusivo, as maiores preocupações dos gestores e administradores de centros de dados, sendo notório o peso que a eficiência energética tem vindo a ter na escolha final do software ou do hardware a utilizar. A preocupação pela produção de componentes de menor consumo energético e, consequentemente, que permitem uma maior poupança nos custos de manutenção de um qualquer centro de dados, incentivou a um claro aumento e aprofundamento do estudo de técnicas e métodos para criar sistemas economicamente mais rentáveis em termos energéticos.

Técnicas como o uso de diferentes esquemas de gestão de energia, como, por exemplo, a apresentada em (Kliazovich et al., 2010), demonstraram já a sua aplicação e utilidade na gestão do consumo de energia em centros de dados. Também, os trabalhos apresentados em

(Beloglazov et al., 2011) e em (Sitaram et al., 2015) são provas muito claras daquilo que tem vindo a ser desenvolvido para a gestão da eficiência do consumo de energia em centros de dados, sendo mais uma confirmação acerca da mudança de paradigma em relação à questão da preocupação energética e da redução do consumo de energia. Do grande volume de dados que é armazenado diariamente em qualquer centro de dados, uma parte bastante significativa está relacionada com o sustento de sistemas de suporte à decisão e, em particular, com o povoamento de *data warehouses*. Não é recente a preocupação energética neste tipo de sistemas, sendo, hoje, claramente notória ao nível dos administradores de centros de dados. Tal conduziu ao incentivo do estudo e do desenvolvimento de soluções que tornassem os sistemas operacionais mais eficientes, menos consumidores de energia. Veja-se, por exemplo, o caso do trabalho apresentado em (Poess et al., 2010).

Usualmente, o povoamento de sistemas de *data warehousing* é realizado por peças de software bastante complexas, que, basicamente, tratam de recolher a informação que este tipo de sistema de dados precisa em diversos sistemas operacionais, conciliando-a, preparando-a e integrando posteriormente no *data warehouse* em questão. Este tipo de sistema é vulgarmente conhecidos por sistema ETL (*Extract – Transform – Load*). Porém, muito embora exista a preocupação de reduzir a despesa mensal que a manutenção de um centro de dados acarreta - um terço da despesa de manutenção recai no consumo energético dos servidores e dos sistemas de refrigeração (Rasmussen, 2011) -, nada de grande relevância ainda foi feito no sentido de minimizar o consumo energético dos sistemas responsáveis pelo povoamento de *data warehouses*. Com base nas pesquisas que efetuámos, até à data, não existe nenhuma técnica ou investigação especialmente orientada para avaliar os consumos energéticos de um sistema ETL, especialmente no que diz respeito à avaliação de diferentes configurações para o sistema e dos seus correspondentes consumos energéticos. Apesar de existir desde 2007 um consórcio industrial denominado "The Green Grid" (Green Grid, 2015), formado por profissionais de todo o mundo, que lida com aspetos relevantes no que diz respeito à eficiência da utilização dos recursos envolvidos nos centros de dados, esse trabalho ainda não foi realizado.

Neste artigo, pretendemos apresentar e discutir o trabalho que temos vindo a realizar em prol da avaliação do consumo de energia que os diversos tipos de tarefas e processos apresentam quando integrados num sistemas de povoamento de um *data warehouse.* Com este trabalho pretendemos determinar a forma como conceber e desenvolver processos ETL mais eficientes em termos energéticos, com base na análise dos diversos componentes utilizados na sua implementação. A partir dessa análise será possível estabelecer uma tabela de eficiência energética para componentes ETL, permitindo, assim, aos arquitetos de sistemas ETL projetar e configurar sistemas de povoamento mais eficientes do ponto de vista energético. Nas próximas secções apresentaremos uma breve exposição de trabalhos relacionados com o processo de investigação que desenvolvemos (secção 2), bem como apresentaremos e discutiremos o caso de estudo selecionado, os dois cenários de implementação prática que levámos a cabo e os resultados alcançados, dando particular relevância, obviamente, à análise dos consumos de energia realizados pelos diversos componentes ETL integrados no sistema que implementámos (secção 3). Por último (secção 4) apresentamos algumas breves conclusões e apontamos algumas linhas de orientação para a realização de novos trabalhos no futuro.

2. Trabalho Relacionado

O tema da eficiência energética tem vindo a ser discutido e investigado por diferentes entidades nas mais diversas áreas do conhecimento. O domínio das tecnologias da informação e da comunicação também tem sido alvo deste tipo de trabalhos. Tanto ao nível do hardware como do software, tem-se vindo a verificar um aumento significativo do número de processos de investigação e da procura de técnicas e métodos que contribuam para a redução do consumo energético dos diversos tipos de componentes que integram os sistemas computacionais e, com

isso, contribuir também para poupar recursos económicos e ambientais. Ao nível dos sistemas de software, por exemplo, o trabalho apresentado em (Carção et al., 2014) permitiu adaptar a técnica de localização de falhas para identificar quais as porções de código responsáveis por um maior consumo energético, tornando possível a partir dessa informação reformular o código de forma a se tornar energeticamente mais eficiente. Outro exemplo relevante de trabalhos nesta área é aquele que foi apresentado em (Couto et al., 2014) no qual se descreveu o desenvolvimento de uma técnica para detetar consumos anómalos de energia em aplicações Android. Estes dois trabalhos, em particular, estão relacionados com a instrumentação de código em sistemas de software e indiciam claramente que a problemática do consumo energético está cada vez mais presente nos diferentes domínios das tecnologias de informação e comunicação. Outra das áreas em que encontramos alguns trabalhos de relevo neste domínio foi a dos centros de dados. Vejamos, por exemplo, o trabalho apresentado em (Sitaram et al., 2015), no qual se relevou a poupança de recursos energéticos através da criação de um algoritmo de alocação, que toma em consideração a eficiência energética bem como a disponibilidade dos diversos recursos do sistema, não criando pontos de falha, e, como tal, contribuindo para um menor consumo energético.

Também existem alguns trabalhos relacionados com a investigação e o desenvolvimento de metodologias específicas para o cálculo do consumo energético de *queries* em sistemas de gestão de bases de dados. O relatório Claremont (Agrawal et al., 2008) é um desses casos, tendo sido uma das primeiras abordagens ao consumo de energia no domínio das bases de dados. Segundo este relatório, na idealização e concepção de um sistema de gestão de bases de dados, é importante ter em consideração o consumo de energia efetuado nas suas diversas tarefas. Mais tarde, a seguir a este relatório, surgiram outros processos de trabalho que tomaram em consideração também essas mesmas preocupações (Lang et al., 2009) (Xu et al., 2010) (Lang et al., 2011). Contudo, todos estes trabalhos apenas se focaram em questões de hardware. Em termos de software, só em (Xu et al., 2010) é que encontrámos uma solução para redesenhar o *kernel* de um SGBD de forma a que este consumisse menos energia. Mais tarde, em (Kunjir et al., 2012) encontramos algumas outras alternativas para reduzir o pico do consumo de energia em sistemas de gestão de bases de dados. Mais recentemente, em (Gonçalves et al., 2014) foi apresentado um trabalho de avaliação sobre o consumo de energia de *queries* num ambiente convencional de um sistema de gestão de bases de dados. Nesse trabalho, os autores redesenharam o plano de execução de uma *query* produzido por um motor de bases de dados de forma a que nele pudesse ser incluído informação pertinente sobre o consumo de energia da query, dando particular relevância ao consumo efetuado em cada um das operações mais elementares que foram realizadas na satisfação da *query* e, consequentemente, determinar o seu consumo total de energia – a esses novos planos foi atribuída a designação de planos de consumo energético.

3. Avaliação do consumo de um Sistema ETL

3.1 O Consumo de Energia em Sistemas ETL

A crescente preocupação que atualmente se verifica relativamente aos consumos energéticos em ambientes de centros de dados tem conduzido, também, a um aumento dos processos de investigação e desenvolvimento nesse domínio. Tanto em termos de hardware como de software, várias são as iniciativas que têm vindo a ser promovidas e concretizadas. Porém, naquilo que se relaciona com o projeto e desenvolvimento de sistemas ETL, até à data, e com base naquilo que até hoje conseguimos apurar, a investigação de técnicas e métodos para avaliar o consumo energético de um sistema ETL são inexistentes. Tal circunstância incentivou-nos a desenvolver um processo de investigação e desenvolvimento que abordasse, em particular, o consumo de energia de um sistema ETL, tomando em consideração cada um dos elementos operacionais (componentes ETL) que são utilizados na sua implementação. A experiência

adquirida em trabalhos anteriores no domínio da eficiência energética em sistemas de software (Gonçalves et al., 2014) (Couto et al., 2014), garantiu-nos as bases mais essenciais para iniciar e desenvolver esse trabalho, investigando a forma como poderíamos definir uma forma de avaliar efetivamente o consumo de um sistema ETL. Com isto complementaríamos os atuais métodos e técnicas para a garantia de desempenho de um sistema, que durante muito tempo foi a principal preocupação dos arquitetos e engenheiros de sistemas de ETL, com mais valor, assegurando assim implementações mais eficientes, tanto em desempenho como em consumo de energia.

Os sistemas ETL são os maiores responsáveis pelo consumo de recursos computacionais num sistema de *data warehousing*. Tais recursos computacionais têm, obviamente, uma tradução direta para recursos energéticos. Assim, a possibilidade de medir o consumo energético de um sistema ETL permitir-nos-á comparar e analisar diferentes implementações para um mesmo sistema, tendo em conta a sua eficiência energética, avaliada a partir dos consumos de energia dos seus diversos componentes. Dessa forma poderemos, também, decidir por esta ou por aquela implementação de acordo com o seu consumo energético. Por exemplo, existem processos ETL específico que são vulgarmente utilizados na implementação de sistemas de povoamento de *data warehouses*. Referimo-nos a processos relacionados com a captura de dados modificados em fontes de informação, a verificação e garantia da qualidade de dados, a substituição encadeada de chaves operacionais, a atualização de tabelas de dimensão com variação ou o carregamento intensivo de dados. Se para um dado processo ETL puder ser realizada uma implementação energeticamente mais eficiente que uma outra qualquer, é possível optar por uma implementação com maiores ganhos energéticos (em detrimento, por exemplo, de uma implementação mais eficiente em termos de desempenho), diminuindo o consumo de energia correspondente, bem como os custos de manutenção do próprio sistema de *data warehousing*.

3.2 A Substituição Encadeada de Chaves Operacionais

Para podermos fazer a avaliação do consumo de energia de um sistema ETL num cenário de aplicação real escolhemos um caso bastante prático relacionado com um dos processos mais vulgares que usualmente podemos encontrar num sistema de povoamento de um *data warehouse*: a substituição encadeada de chaves operacionais – *surrogate key pipelining*. Basicamente, este processo tão particular permite fazer a conversão de chaves provenientes dos sistemas operacionais fonte por outras chaves correspondentes. Estas últimas são comummente designadas por chaves de substituição. Estas chaves artificiais incorporam os esquemas das tabelas de dimensão, acolhendo valores inteiros, únicos, sequenciais e sem qualquer carga semântica associada, que são utilizados, entre outras coisas, como base para a conciliação de valores referentes a uma mesma entidade, para sustentar a realização de processos de atualização de dimensões com variação, ou para facilitar a realização de processos de junção em *data warehouses*. Na Figura 1 apresentamos uma versão simplificada do processo de definição e mapeamento de uma chave de substituição.

A importância deste tipo de chaves, ao invés das ditas chaves operacionais (frequentemente referidas também como chaves naturais), reside no facto destas garantirem identificadores inequívocos para registos de uma tabela de um *data warehouse*. Quando no povoamento de um *data warehouse* estão envolvidas mais do que uma fonte de informação, mais do que um sistema operacional, é vulgar ocorrerem diversos conflitos entre as chaves operacionais de diferentes sistemas, não só causados por valores repetidos como também por valores antagónicos ou inconsistentes. Como tal, estas chaves têm que ser substituídas aquando da sua "migração" para o *data warehouse*. Além disso, num *data warehouse* as chaves das tabelas devem ser desprovidas de inteligência, de significado. As chaves de substituição garantem isso, quando substituem as chaves operacionais em simples sequências de números inteiros (se possível pequenos). Para reforçar um pouco mais a necessidade da utilização de chaves de substituição, está também o facto de que, ocasionalmente, num contexto empresarial, os valores

das chaves operacionais realmente mudarem ao longo do tempo, por este ou por aquele motivo. Num *data warehouse* não é nada "boa" a prática de modificar os valores das chaves numa tabela de dimensão ou numa tabela de factos. Usualmente, em casos como estes, são geradas novas chaves de substituição e estabelecidos os relacionamentos que garantam a ligação entre os diversos valores envolvidos.

```
procedimento defineChaveSubstituição
   (in tabelaLookup char, chaveOperacional variant; out chaveSubstituição int)

início
   chaveSubstituição ← obterChaveSubstituição(tabelaLookup,chaveOperacional)
   se chaveSubstituição = NULL
     então chaveSubstituição ← gerarChaveSubstituição(tabelaLookup,chaveOperacional)
fim
```

Figura 1 – Atribuição de chaves de substituição.

3.3 Um Processo de Substituição Encadeada de Chaves Operacionais

Um processo de substituição encadeada de chaves operacionais pode ser implementado de diferentes formas, com base na natureza do próprio processo e das chaves envolvidas como também dos resultados pretendidos – o tipo de chave de substituição que se pretende aplicar. Na Figura 2 podemos ver uma versão bastante simplificada deste processo. No caso de estudo que escolhemos, a implementação do processo de substituição foi realizada recorrendo a uma ferramenta de ETL bastante conhecida - o Kettle (Kettle, 2015) ou *Pentaho Data Integration* - e envolveu a implementação de duas soluções distintas para o processo de substituição referido.

```
procedimento substituiçãoEncadeadaChaves
   (in listaChavesOperacionais list; out listaChavesSubstituição list)

início
   para cada (tabelaLookup,chaveSubstituição) na listaChavesOperacionais fazer
      início
         defineChaveSubstituição(tabelaLookup,chaveOperacional,chaveSubstituição)
      fim
fim
```

Figura 2 – A substituição encadeada de chaves operacionais.

A primeira implementação do sistema ETL – versão A – aborda a questão da substituição encadeada de chaves operacionais da forma mais habitual. Nesta versão, o processo de substituição atua registo a registo e, para cada um deles, identifica as chaves operacionais a substituir, a partir de uma configuração base previamente estabelecida, e, chave a chave, vai verificando se existe ou não já uma chave de substituição atribuída. Para isso o processo utiliza, para cada chave operacional envolvida, uma tabela de *lookup* específica. Estas tabelas são estruturas temporárias para suporte ao mapeamento de chaves (e por vezes de outra informação necessária para a conciliação e garantia da qualidade de dados envolvidos no povoamento). Assim, caso a tabela de *lookup* em causa contenha um valor de substituição para o valor da chave operacional em questão, o processo faz a sua substituição pelo seu correspondente valor de substituição (*surrogate key*). Num processo de implementação real, caso não existisse um valor de substituição, o processo faria a geração de uma nova chave de substituição, armazenando-a na tabela de *lookup* correspondente, substituindo de seguida a chave operacional. De forma a simplificar um pouco o nosso caso de estudo e garantir a regularidade

da execução do processo (não tratando exceções ou situações de erro, por exemplo), a parte da geração de chaves de substituição não foi considerada. Assumimos, somente, que todas as chaves de substituição deveriam estar garantidas no momento da execução do sistema ETL. Todavia, aplicámos um processo de tratamento de exceções, comum a todos os casos, que trata de situações nas quais se verifique a não existência de uma chave de substituição, criando um procedimento de quarentena, algo bastante típico em sistemas de povoamento de *data warehouses*.

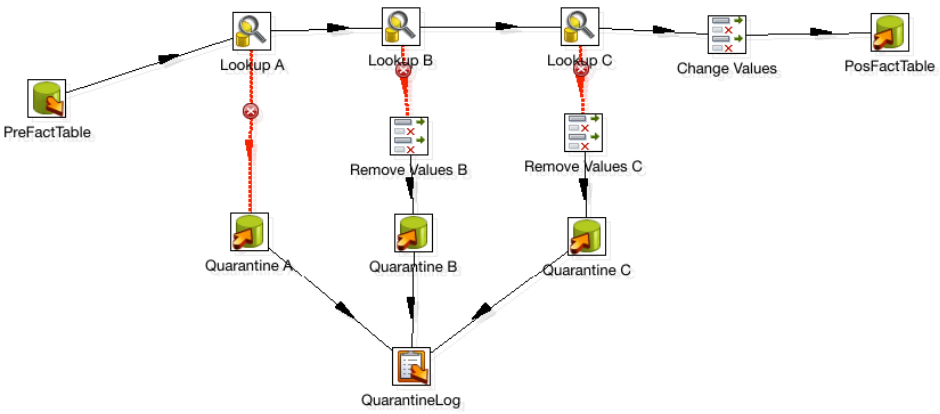


Figura 3 – Implementação do processo de substituição encadeada de chaves operacionais (versão A).

Na Figura 3, podemos observar o *package* Kettle que foi desenvolvido para acolher a primeira versão do processo que acabámos de descrever. Nela, facilmente identificamos a forma como o processo de substituição encadeada de chaves operacionais acontece. Num primeiro componente (*Table Input – 'PreFactTable'*) é realizada a leitura dos registos que temos que processar. São registos típicos de uma tabela de factos, aqueles que contêm as chaves operacionais (três atributos) cujos valores, assumimos, devem ser substituídos. De seguida, o processo inicia as tarefas de substituição de chaves operacionais, realizando sequencialmente três operações de *lookup*, uma para cada chave operacional, utilizando três componentes *DB Lookup* encadeados (*'Lookup A', 'Lookup B'* e *'Lookup C'*) para fazer, respetivamente, as necessárias conversões dos valores contidos em três atributos: A, B e C. Em qualquer um destes componentes, e seguindo o que expusemos anteriormente, caso não exista um valor de substituição para a chave operacional em tratamento, o registo dessa chave é copiado para uma tabela de quarentena utilizando um componente *Table Output*. Desta forma assinalamos e guardamos os registos cujos valores das chaves operacionais contêm valores incorretos, possibilitando a sua posterior análise e recuperação, se possível. Os restantes componentes Kettle que figuram na implementação desta primeira versão, nomeadamente *Select Values* (*'Remove Values B', 'Remove Values C' e 'Change Values'*) e *Write to Log* (*'QuarantineLog'*) foram utilizados respetivamente para fazer a seleção de atributos num dado registo e para escrever um determinado evento num dado ficheiro de log.

A versão B (Figura 4) da implementação do processo de substituição encadeada de chaves operacionais é, de certa forma, mais elaborada que a versão A. Uma diferença notória é o número de componentes ETL que envolve na sua implementação, que é claramente superior ao da versão A. Isso deveu-se ao tipo de implementação que foi realizado, uma vez que pretendíamos com esta versão alternativa implementar um processo de substituição de chaves paralelo, no qual várias chaves operacionais poderiam ser substituídas simultaneamente, em contraposição à versão A que, como referimos anteriormente, processa os registos e a substituição das chaves de forma sequencial, em mode *pipelining*. Na versão B não temos de

facto um *pipelining*. Mas a forma como o sistema está implementado permite-lhe ter um comportamento perfeitamente equivalente, em termos de resultado refira-se, à da versão A.

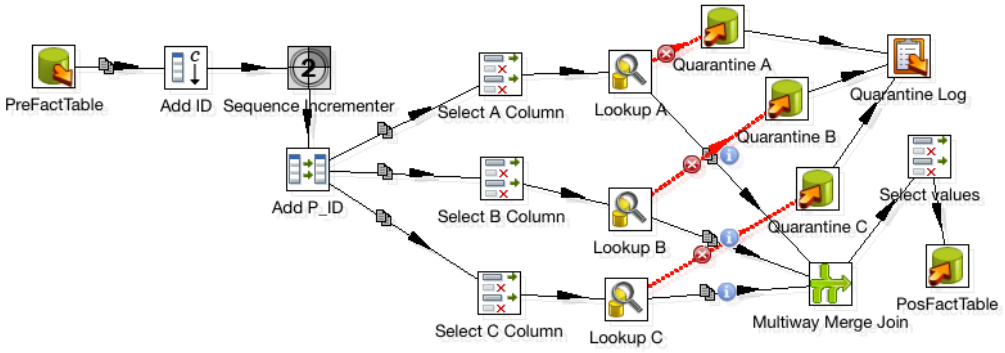


Figura 4 – Implementação do processo de substituição encadeada de chaves operacionais (versão B).

Vejamos então como é que a versão B atua. Através da Figura 4 vemos que o processo ETL se inicia da mesma forma que na versão A, com a leitura dos registos a processar. Porém, a partir daqui, o processo desenrola-se de forma bastante distinta. As três tarefas seguintes, envolvendo, respetivamente, componentes ETL *Add Constant* (*'Add Id'*), *Add Sequence* (*'Sequence Incremeter'*) e *Set Field Value* (*'Add P_Id'*), preparam os registos de forma a que os atributos das chaves operacionais sejam separados para serem processados em paralelo nos passos seguintes, de acordo com cada uma das tabelas de *lookup* envolvidas no processo. Após a separação dos atributos, são despoletados três tarefas de substituição de chaves em paralelo, envolvendo dois tipos de componentes, ETL *Select Values* e *DB Lookup* -, uma para cada um dos atributos envolvidos (A, B e C). Nesta tarefas o processo ocorre de forma similar ao seu equivalente na versão A, o mesmo acontecendo para os casos dos valores de atributos que vão para quarentena. Depois disso, os atributos que foram anteriormente separados voltam a ser reunidos por um componente ETL *Multiway Merge Join*, sendo o registo agora restaurado com as chaves de substituição já atribuídas. Por fim, os registos processados são colocados numa tabela de factos no *data warehouse*. Os restantes componentes Kettle que figuram na implementação desta segunda versão, nomeadamente *Select Values* (*'Select Values'*) e *Write to Log* (*'QuarantineLog'*) fazem a seleção de atributos num dado registo e a escrita de um determinado evento num ficheiro de log. Estes componentes desempenham um papel muito semelhante àquele que realizaram na versão A.

Componente ETL	Versão A	Versão B
Table Input	✓	✓
Table Output	✓	✓
DB Lookup	✓	✓
Select Values	✓	✓
Write to Log	✓	✓
Add Constant		✓
Add Sequence		✓
Set Field Value		✓
Multiway Merge Join		✓

Tabela 1 – Relação dos componentes ETL usados nas diferentes versões do processo ETL implementado.

Na tabela 1 podemos observar os diferentes componentes ETL que foram utilizados nas implementações das duas versões do de ETL. A designação de cada um desses componentes está apresentada de acordo com a nomenclatura seguida pela ferramenta de ETL que utilizámos. As diferenças que se verificam entre as duas implementações realizadas podem ser facilmente identificadas através dos componentes que foram utilizados em cada uma delas.

3.4 Avaliação do Consumo de Energia

De forma a garantir que a avaliação do consumo de energia do processo de povoamento em causa fosse mais correta, não se resumindo simplesmente a uma comparação de implementações, projetámos e desenvolvemos seis cenários de teste distintos, com objetivos diferentes, de forma a obter várias medições do consumo e do desempenho das duas versões implementadas para o processo ETL em causa. Para cada cenário de teste considerámos diferentes volumes de dados. De teste para teste, aumentámos significativamente o número de registos com chaves operacionais para substituir. Além disso, criámos tabelas de *lookup* com diferentes tamanhos e disposições de armazenamento, garantindo tais características no momento do seu povoamento inserindo os seus registos de uma forma ordenada ou de uma forma aleatória conforme as características de pesquisa que queríamos atribuir às tabelas. Assim, as tabelas de *lookup,* para o processamento dos valores dos atributos A, B e C, foram povoadas, respetivamente, com 10, 100 e 10 000 registos, inseridos com base nas modalidades anteriormente referidas, criando os correspondentes dados para suporte aos cenários de teste t_1, t_2 e t_3. O volume de dados a processar pelo sistema ETL implementados foi constante para todos esses cenários de teste, tendo sido fixado em 243 000 registos. De salientar que, nem todas as chaves operacionais integradas nesses registos tinham uma chave de substituição na tabela de *lookup* correspondente. Como tal, o número de registos que de facto chegarão ao *data warehouse* será inferior aos 243 000 registos iniciais. Cada cenário de teste foi executado cinquenta vezes para cada uma das implementações realizadas (versão A e versão B). Na prática isso correspondeu à repetição dos testes cinco vezes, tendo sido realizadas em cada uma das realizações dos testes dez execuções da implementação do sistema ETL em questão. Como planeado, durante a realização dos diversos cenários de teste foram coletados os valores dos consumos de energia realizados por cada um dos componentes ETL integrados nas duas versões do sistema.

	Aleatório			**Ordenado**		
	t_1	t_2	t_3	t_1	t_2	t_3
Table Input	346.38	355.44	340.32	335.12	297.28	2712.35
Table Output	359.05	374.47	354.38	348.16	302.92	2826.30
DB Lookup	359.05	373.88	352.87	348.16	301.47	2826.30
Select Values	359.05	374.47	352.87	348.16	302.48	2826.30
Write to Log	362.44	375.86	359.43	351.41	302.92	2826.30

Tabela 2 – Energia consumida (em Joules) por cada componente integrado na versão A do processo ETL.

Após a execução dos vários cenários de testes, para ambas as implementações do sistema ETL, procedemos à análise dos dados de consumo recolhidos. Nas Tabelas 2 e 3 podemos observar uma relação dos consumos realizados pelos diversos componentes ETL utilizados, respetivamente, na versão A e na versão B do sistema ETL. Nessas tabelas, os valores de consumo estão organizados segundo cada um dos componentes ETL utilizados – *Table Input, Table Output, DB Lookup, Select Values*, etc. – e de acordo com o modo como os dados estavam organizados nas tabelas de *lookup* – em modo aleatório ou ordenado. De seguida, nas Figuras 5 e 6, podemos ver, em gráfico, respetivamente, o resultado da análise do consumo energético e o resultado da análise de desempenho do sistema, com base nos valores coletados durante os testes realizados.

	Aleatório			Ordenado		
	t_1	t_2	t_3	t_1	t_2	t_3
Table Input	446.56	492.59	440.32	419.87	474.00	3916.59
Table Output	531.41	582.91	529.53	509.57	579.10	4788.87
DB Lookup	520.95	582.91	529.53	509.57	572.93	4728.21
Select Values	520.95	543.00	490.20	493.97	572.93	4505.13
Write to Log	362.44	582.91	529.53	509.57	572.93	4788.16
Add Constant	471.45	517.72	465.23	444.72	499.25	4104.24
Add Sequence	471.45	543.00	490.20	469.48	524.42	4291.49
Set Field Value	471.45	543.00	490.20	469.48	524.42	4291.49
Multiway Merge Join	531.41	582.91	529.53	509.57	579.10	4788.87

Tabela 3 – Energia consumida (em Joules) por cada componente integrado na versão B do processo ETL.

Uma das conclusões que podemos retirar, de imediato, através da análise da Tabela 2 e da Tabela 3, é que, em termos comparativos, a versão A do processo de substituição encadeada de chaves operacionais é energeticamente mais eficiente do que a versão B. Para qualquer um dos cenários de teste - t_1, t_2 e t_3 - a versão B é claramente mais consumidora que a versão A. Quanto à questão do consumo energético de ambas as versões é de notar que, para o cenário de teste t_3, que acolheu uma tabela de *lookup* com 10 000 registos, com os registos ordenados, há um consumo de energia excessivo quando comparado com aquilo que foi consumido com a configuração aleatória dos registos. Em termos de desempenho é possível observar que a versão A tem um melhor desempenho do que a versão B, demorando menos tempo a executar o processo ETL.

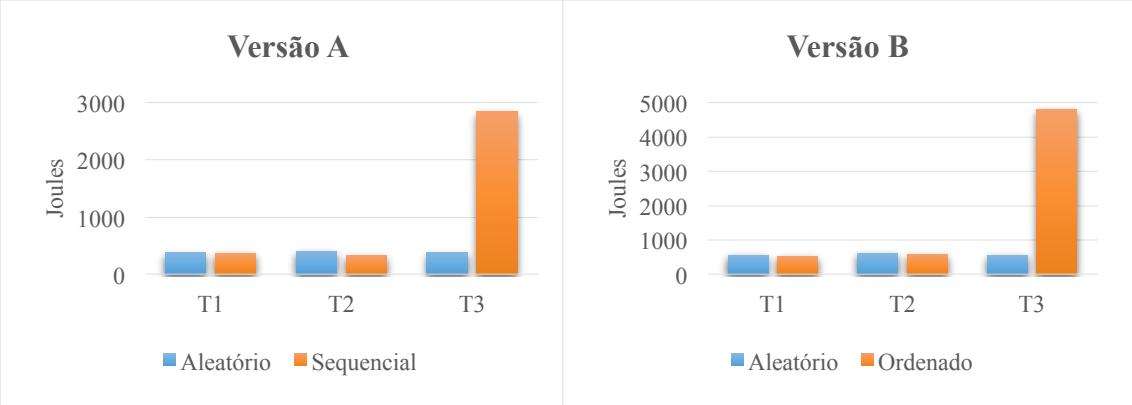


Figura 5 – Consumo energético geral do processo ETL.

O consumo excessivo de energia no cenário de teste t_3 sequencial, para ambas as versões, é explicado pelo tempo global que o processo demora a executar. O cálculo da energia é efetuado recorrendo à multiplicação da potência elétrica pela quantidade de tempo em segundos. Assim, se, neste caso, o tempo de execução aumentar, a energia consumida durante o processo de substituição encadeada de chaves naturais também aumenta. De uma forma geral, a versão A do processo ETL é, sem dúvida, muito mais eficiente do que a versão B, quer em termos de consumo energético, quer em termos de desempenho. Apesar destes resultados, é necessário avaliar um maior número de processos ETL para que seja possível criar sistemas energeticamente mais eficientes e definir, por fim, uma tabela de categorização de consumos de energia para componentes ETL. Porém, o facto de podermos medir e avaliar quantitativamente o consumo energético de um determinado processo ETL é claramente uma mais valia para qualquer projeto de implementação de um sistema de povoamento de um data warehouse, uma vez que torna possível a opção por diferentes alternativas de implementação, nas quais se ponderará o peso de uma solução mais eficiente em termos de desempenho e uma solução mais eficiente em termos de consumo de energia.

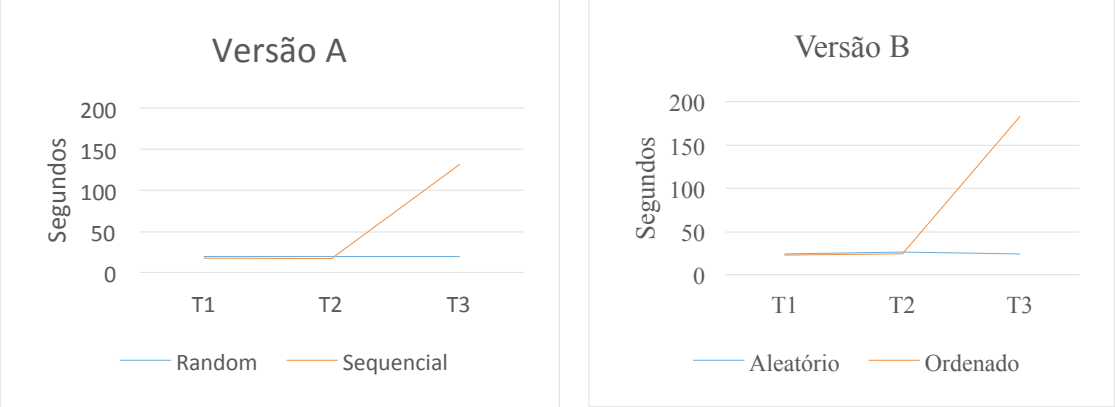


Figura 6 – Desempenho geral em segundos do processo ETL.

4 Conclusões e Trabalho Futuro

Neste artigo apresentámos uma avaliação energética para componentes ETL, com base na comparação de duas implementações distintas de um processo de povoamento de um *data warehouse* – a substituição encadeada de chaves operacionais. Este tipo de avaliação permitiu-nos conhecer o consumo de energia de um processo ETL, bem como a dos diversos componentes que realizam as tarefas implementadas. É muito interessante verificar o impacto que diferentes tipos de implementações para um mesmo caso de estudo têm em termos de consumo de energia. Isto é algo que há bem pouco tempo não se pensaria fazer. Avaliar, em particular, o consumo de energia de um sistema de software pensava-se ser irrelevante quando comparado a outro tipo de sistema, em particular os sistemas de hardware. Porém, em grandes sistemas computacionais, como é o caso dos sistemas de *data warehousing*, mesmo o consumo de pequenos componentes causa alguma diferença quando tais componentes são utilizados em grande escala. As pequenas diferenças, aqui, fazem realmente a diferença.

As preocupações de um qualquer administrador de um sistema de povoamento sempre estiveram direcionadas para a implementação e execução dos sistemas implementados, para análise da sua correção e desempenho. Todavia, hoje, com os sistemáticos alertas sobre os elevados consumos de energia dos sistemas computacionais, em especial os relacionados com a implementação de grandes centros de dados, e com a consequente degradação dos recursos naturais conjugada com a escassez de recursos económicos, tais preocupações chegaram também aos sistemas de povoamento de *data warehouses*. Não é pois de estranhar a grande cadência com que estão a aparecer estudos neste domínio. Apesar de existirem várias técnicas para obter ou determinar os consumos de energia, por exemplo, de coisas tão pequenas como uma *query* sobre uma base de dados, a obtenção de dados relativos a consumos energéticos em processos de ETL é praticamente inexistente. Neste trabalho fizemos esse estudo. A avaliação do consumo de energia de um processo de substituição de chaves operacionais, permitiu-nos abrir alguns horizontes relativamente à disponibilização de informação referente aos consumos energéticos para componentes ETL.

De facto, verificamos que este tipo de estudos tem grande interesse e viabilidade. A dimensão dos sistemas atuais requer algum cuidado e preocupação na forma como os desenvolvemos e que componentes usamos ou deixamos de usar na sua implementação tendo em conta os seus requisitos em termos de consumo de energia. Tendo em conta alguns aspetos de escala dos sistemas computacionais, uma análise minuciosa sobre o consumo de energia dos seus componentes pode conduzir, de facto, a uma poupança significativa de energia e, como tal, contribuir para a consequente poupança de recursos naturais. Futuramente, pretendemos

construir um repositório com toda a informação energética referente a processos comuns de ETL, bem como a relativa aos diversos componentes ETL que utilizam, para diferentes ferramentas comerciais atualmente disponíveis no mercado, alargando o espectro de aplicação do atual estudo para além do contexto de exploração da ferramenta Kettle. Através da informação do consumo dos componentes ETL, pretendemos construir uma tabela de eficiência energética, para que, aquando da fase de implementação de um processo ETL seja possível substituir componentes bastante consumidores, por outros com as mesmas funcionalidades, mas energeticamente mais eficientes.

5 Referências

Agrawal,R., Ailamaki,A., Bernstein,P.A., Brewer,E.A., Carey,M.J., Chaudhuri,S., Doan, A., Florescu, D., Franklin, M.J., Garcia-Molina, H., Gehrke, J., Gruenwald, L., Haas, L.M., Halevy, A.Y., Hellerstein, J.M., Ioannidis, Y.E., Korth, H.F., Kossmann, D., Madden, S., Magoulas, R., Ooi, B.C., O'Reilly, T., Ramakrishnan, R., Sarawagi, S., Stonebraker, M., Szalay, A.S., Weikum, G.: The claremont report on database research. SIGMOD Rec. 37(3), 9–19, September, 2008.

Beloglazov, A., Abawajy, J., Buyya, R., "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing". Future Generation Computer Systems, Special Section: Energy efficiency in large-scale distributed systems, 28(5), 755 – 768, 2012.

Carção, T., "Measuring and visualizing energy consumption within software code". In: Visual Languages and Human-Centric Computing (VL/HCC), 2014 IEEE Symposium on, 181–182, July, 2014.

Couto, M., Carção, T., Cunha, J., Fernandes, J.P., Saraiva, J., "Detecting anomalous energy consumption in android applications". In Pereira, F.M.Q., ed.: Programming Languages - 18th Brazilian Symposium, SBLP 2014, Maceio, Brazil, October 2-3, 2014. Proceedings. Volume 8771 of Lecture Notes in Computer Science., Springer, 77–91, 2014.

Gonçalves, R., Saraiva, J., Belo, O., "Defining Energy Consumption Plans for Data Querying Processes", In Proceedings of 2014 IEEE Fourth International Conference on Big Data and Cloud Computing (BdCloud 2014), pages 641-647, IEEE computer Society, Sidney, Australia, 2014.

Green Grid, The Green Grid, 2015. Available at: http://www.thegreengrid.org/ [Accessed June 15, 2015].

Harizopoulos, S., Shah, M.A., Meza, J., Ranganathan, P., "Energy efficiency: The new holy grail of data management systems research". In: CIDR 2009, Fourth Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 4-7, 2009, Online Proceedings, www.cidrdb.org, 2009.

Kettle, Pentaho, "Pentaho Data Integration", 2015. Available at: http://www.pentaho.com/product/data-integration [Accessed March 16, 2015].

Kliazovich, D., Bouvry, P., Khan, S.U., "Greencloud: a packet-level simulator of energy-aware cloud computing data centers". The Journal of Supercomputing, 62(3), 1263–1283, 2012.

Kunjir, M., Birwa, P.K., Haritsa, J.R., "Peak power plays in database engines". In Rundensteiner, E.A., Markl, V., Manolescu, I., Amer-Yahia, S., Naumann, F., Ari, I., eds.: 15th International Conference on Extending Database Technology, EDBT '12, Berlin, Germany, March 27-30, 2012, Proceedings, ACM, 444–455, 2012.

Lang, W., Kandhan, R., Patel, J.M., "Rethinking query processing for energy efficiency: Slowing down to win the race". IEEE Data Eng. Bull. 34(1), 12–23, 2011.

Lang, W., Patel, J.M., “Towards eco-friendly database management systems”. In: CIDR 2009, Fourth Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 4-7, 2009, Online Proceedings, www.cidrdb.org, 2009.

Poess, M., Nambiar, R., “Tuning servers, storage and database for energy efficient data warehouses”. In Proceedings of 2010 IEEE 26th International Conference on Data Engineering (ICDE), 1006–1017, March, 2010.

Rasmussen, N., “Determining total cost of ownership for data center and network room infrastructure”. Technical report, Schneider Electric Data Center Science Center, 2011.

Sitaram, D., Phalachandra, H., Gautham S, Swathi, H., Sagar, T., “Energy efficient data center management under availability constraints”. In 9th Annual IEEE International Systems Conference (SysCon), 377–381, April 2015.

Xu, Z., Tu, Y., Wang, X., “Exploring power-performance tradeoffs in database systems”. In Li, F., Moro, M.M., Ghandeharizadeh, S., Haritsa, J.R., Weikum, G., Carey, M.J., Casati, F., Chang, E.Y., Manolescu, I., Mehrotra, S., Dayal, U., Tsotras, V.J., eds.: Proceedings of the 26th International Conference on Data Engineering, ICDE 2010, March 1-6, 2010, Long Beach, California, USA, IEEE, 485–496, 2010.

Column-Based Databases: Estudo exploratório no âmbito das Bases de Dados NoSQL

José Pedro Cunha [1], José Luís Pereira [2]

1) Universidade do Minho, Portugal
a58322@alunos.uminho.pt

2) Universidade do Minho, Portugal
jlmp@dsi.uminho.pt

Resumo

O aumento da quantidade de dados gerados que se tem verificado nos últimos anos e a que se tem vindo a dar o nome de *Big Data* levou a que a tecnologia relacional começasse a demonstrar algumas fragilidades no seu armazenamento e manuseamento o que levou ao aparecimento das bases de dados NoSQL. Estas estão divididas por quatro tipos distintos nomeadamente chave/valor, documentos, grafos e famílias de colunas. Este artigo é focado nas bases de dados do tipo *column-based* e nele serão analisados os dois sistemas deste tipo considerados mais relevantes: Cassandra e HBase.

Palavras chave: *Big Data*, NoSQL, *Column-based databases*, Cassandra, HBase

1. Introdução

Depois de várias décadas de grande sucesso e bons serviços prestados às organizações, a tecnologia relacional de bases de dados tem vindo a ser desafiada por uma nova classe de tecnologias de bases de dados a que se deu a designação genérica de NoSQL (*Not only* SQL). Para este facto contribuíram decisivamente os recentes desenvolvimentos na área a que se tem vindo a chamar *Big Data*, em que o aumento da quantidade de dados gerados diariamente em diversos domínios de aplicação como a *Web*, dispositivos móveis e principalmente as redes sociais, entre outros, está atualmente na ordem das centenas de *Terabytes* [Kuznetsov & Poskonin 2014]. Como tal, tendo em conta o volume e complexidade dos dados a gerir a tecnologia relacional começa a demonstrar fragilidades substanciais. Em particular, a necessidade de gerir dados cujos formatos são dificilmente acomodáveis em sistemas relacionais, dispersos por múltiplos servidores, levou ao aparecimento das ditas Bases de Dados NoSQL. Estas são principalmente focadas no desempenho, permitindo o processamento de dados de forma rápida e

eficiente. O seu modelo de dados não necessita de seguir os padrões rígidos do modelo relacional, pelo que armazenam tanto dados estruturados como não estruturados. Dentro desta nova classe de tecnologias de Bases de Dados surgiram diferentes propostas, com distintas proveniências e áreas de aplicação, vulgarmente classificadas em quatro grupos, de acordo com o seu modelo de dados: *Column, Document, Key/Value e Graph Based databases*, sendo que cada um destes modelos possui uma grande diversidade de propostas no mercado [Hossain & Moniruzzaman 2013].

Este artigo é focado nas bases de dados NoSQL do tipo *column-based* e tem como objetivo efetuar uma análise comparativa dos produtos tecnológicos mais relevantes deste grupo assim como compará-los com a tecnologia relacional de forma a compreender quais as suas caraterísticas, diferenças e principais vantagens. Assim, tendo em conta o seu prestígio, documentação existente e a quantidade de organizações de renome que utilizam os seus serviços, as bases de dados NoSQL do tipo *column-based* consideradas mais relevantes para análise são o Cassandra e o Apache HBase.

Em termos de estrutura, o artigo começa por abordar as principais caraterísticas das bases de dados NoSQL, assim como fazer uma breve s*í*ntese acerca de cada um dos quatro tipos de base de dados já mencionados. De seguida é feita uma análise individual às ferramentas Cassandra e HBase no que diz respeito ao seu modelo de dados, modelo de consulta e ainda à sua arquitetura. A seção seguinte é acerca do trabalho prático a realizar numa segunda etapa e por fim as conclusões retiradas acerca do estudo efetuado.

2. Caraterísticas das Bases de Dados NoSQL

O termo NoSQL foi utilizado pela primeira vez em 1998 por Carlo Strozzi ao fazer referência a uma base de dados relacional criada por si denominada "Strozzi SQL" que tinha a particularidade de não utilizar a linguagem SQL para efetuar consultas aos dados. O termo voltou a ser utilizado mais tarde em 2009 numa conferência organizada por Johan Oskarsson acerca de bases de dados do tipo *open-source* com o nome "NoSQL *meetup*", sendo que a partir daí passou a ser o termo da moda para referir bases de dados não relacionais [Abramova et al. 2014].

O seu modelo de dados não possui os padrões rígidos do modelo relacional pelo que armazenam dados estruturados como não estruturados. Esta ausência de esquema permite adicionar campos aos registos da base de dados sem que antes tenham que ser feitas alterações na estrutura da mesma. Ao contrário do modelo relacional, as bases de dados NoSQL não suportam consultas aos dados armazenados através da linguagem SQL. Contudo, permitem facilmente armazenar e

recuperar dados de forma rápida e eficiente independentemente da sua estrutura e conteúdo [Leavitt 2010].

De forma a evitar a perda de dados as bases de dados NoSQL possuem uma arquitetura distribuída tolerante a falhas que se baseia na distribuição dos dados por vários servidores pelo que, caso um servidor deixe de funcionar, o sistema manter-se-á em funcionamento garantindo assim elevados índices de disponibilidade. Outra das principais caraterísticas destas bases de dados não relacionais é a escalabilidade horizontal que consiste em aumentar o número de máquinas do sistema, dividindo o processamento dos dados conforme as necessidades do mesmo de forma a obter sempre elevados níveis de desempenho. Uma das técnicas utlizadas para garantir isto é denominada por *sharding* [Sadalage & Fowler 2013, página 38].

Relativamente à consistência, enquanto o modelo relacional segue as propriedades ACID que entre outras coisas providencia uma consistência forte dos dados, no caso das bases de dados NoSQL pode verificar-se apenas uma *eventual consistency* conforme as necessidades do sistema em questão e tendo em consideração o teorema CAP. Este refere que entre as propriedades consistência, disponibilidade e tolerância à partição, apenas duas delas podem ser asseguradas em simultâneo, sendo sempre necessário abdicar de uma delas. Normalmente a grande discussão que se coloca neste tipo de base de dados é entre a disponibilidade e a consistência visto que por se tratarem de sistemas distribuídos partições de rede são uma constante e como tal torna-se necessário que os sistemas sejam tolerantes a este tipo de falhas [Leavitt 2010].

Existem assim, quatro tipos de bases de dados NoSQL:

- ***Key/Value-based*:** destacam-se como sendo as mais simples e onde se pode verificar um melhor desempenho. Cada campo é constituído por uma chave e um valor e funciona como uma tabela *hash*, sendo que o conteúdo do valor pode ser armazenado em qualquer formato de dados. O acesso aos dados é sempre efetuado através da sua chave primária que é única e através das operações *get, put* e *delete*, sendo que em geral não são suportadas operações que abranjam conjuntos de valores [Atzeni et al. 2014]. Este tipo de armazenamento suporta uma elevada quantidade de dados e tem como principais vantagens a boa escalabilidade horizontal, simplicidade e o alto desempenho, sendo particularmente adequado para situações em que os dados não se relacionam. Por outro lado não é apropriado para realizar consultas complexas aos dados dado que este modelo não possui uma boa capacidade de indexação efetuando apenas operações *read* e *write* básicas. Num contexto real, este tipo de bases de dados são apropriadas para, por exemplo, gerir *stock* e armazenar informação relativa a perfis de utilizador ou compras em plataformas *online* dada a sua simplicidade e velocidade na recuperação dos dados

[Abramova et al. 2014]. As bases de dados do tipo *key-value* mais relevantes são o Riak e o Redis;

- ***Document-based*:** este tipo de base de dados baseiam-se em modelos de dados semiestruturados onde a informação é armazenada em forma de conjuntos de documentos que podem ter diferentes formatos tais como XML, BSON ou JSON. Cada documento é constituído por um conjunto de campos armazenados de forma não estruturada e que estão associados a uma chave única que os identifica. O seu modelo de consulta é considerado superior ao das restantes categorias de base de dados NoSQL pelo que permitem consultar conjuntos de documentos, efetuar consultas agregadas, incluir restrições sobre o campo de valores pretendido, ordenar resultados, entre outras coisas. Assim, este tipo de base de dados são apropriadas para problemas que envolvem ambientes variados e focam-se no armazenamento de grandes volumes de dados e no bom desempenho na consulta dos mesmos [Kuznetsov & Poskonin 2014]. Num contexto prático, este tipo de bases de dados representam uma boa solução para trabalhar com aplicações *web* com grandes quantidades de documentos que podem ser armazenados em ficheiros estruturados tais como documentos de texto, *emails*, sistemas XML ou sistemas CRM. São também uma boa opção para aplicações de comércio eletrónico [Abramova et al. 2014]. Os sistemas deste tipo mais adotados são o MongoDB e o Apache CouchDB;
- ***Graph-based*:** neste tipo de sistema o modelo é baseado em nós que se relacionam entre si por meio de uma aresta. Os nós contêm propriedades na forma de pares chave/valor e os relacionamentos têm sempre um nome associado e uma direção contendo um nó remetente e outro destinatário. As relações também podem conter propriedades [Robinson et al. 2013]. A organização do grafo permite que os dados depois de armazenados sejam interpretados de forma diferente com base nas diversas relações, sendo que um nó pode ter um número ilimitado de relações [Sadalage & Fowler 2013, página 113]. Estes sistemas são considerados ideais para lidar com dados que estejam interligados sendo importantes para, por exemplo analisar relacionamentos entre utilizadores de redes sociais, analisar sistemas de transportes e plataformas que sugerem produtos a utilizadores com base nos seus gostos e preferências [Abramova et al. 2014]. As principais referências deste tipo de bases de dados são Neo4J e o InfoGrid;
- ***Column-based:*** este tipo de sistema NoSQL destaca-se pela sua capacidade de armazenar grandes quantidades de dados. Ao contrário dos tradicionais modelos relacionais, uma linha corresponde a um conjunto de colunas associadas à mesma chave

primária, não existindo um esquema previamente definido em que cada coluna de uma linha tem o seu espaço atribuído mesmo que seja do tipo *NULL*. Cada linha pode conter um conjunto diferentes de colunas, o que se traduz num menor esforço computacional e contribui para aumentar o desempenho do sistema. Cada coluna é constituída pelo par nome e valor, sendo que o nome funciona como a chave da coluna. Além disto cada um destes é armazenado em conjunto com um valor *timestamp* que fornece informações acerca do valor temporal em que aqueles dados foram inseridos ou atualizados permitindo entre outras coisas resolver conflitos relacionados com *writes* e lidar com dados que estejam desatualizados. Para além disto as colunas são armazenadas por famílias de colunas de forma a facilitar a organização e a distribuição dos dados. Este modelo é o ideal para lidar com estruturas de dados complexas e volumosas [Sadalage & Fowler 2013, página 115]. Num contexto prático este tipo de base de dados é a ideal para, por exemplo, lidar com sistemas em que se verifica um elevado registo de eventos constante [Abramova et al. 2014]. As bases de dados mais populares deste tipo são o Cassandra e o Apache HBase.

3. Cassandra Vs HBase

De seguida apresentam-se sucintamente os dois produtos considerados mais representativos da classe de bases de dados NoSQL do tipo column-based.

3.1. Cassandra

O Cassandra foi desenvolvido pelo Facebook em 2008 como um projeto *open-source*, com o intuito inicial de melhorar o mecanismo do motor de busca da plataforma. Em 2009 foi adotado pela Apache Software Foundation e é hoje em dia, utilizado por grandes empresas como o Ebay, o Twitter e a Cisco Systems para além de órgãos governamentais como a NASA [Weber & Strauch 2010].

Os seus criadores, Avinash Lakshman e Prashant Malik, definem-na como uma base de dados distribuída e altamente escalável criada para armazenar grandes quantidades de dados distribuídos por vários servidores oferecendo alta disponibilidade e uma eventual consistência dos dados [Laksham & Prashant 2010].

3.1.1. Modelo de Dados

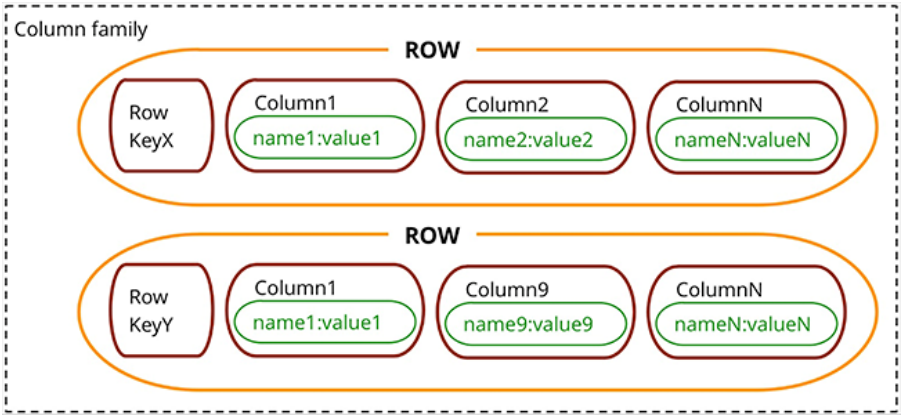


Figura 1 - Estrutura de uma família de colunas do Cassandra (retirado de [Sadalage & Fowler 2013])

O modelo de dados Cassandra é muito semelhante ao do BigTable e é constituído por [Datastax 2015]:

- **Coluna:** no Cassandra a unidade básica de armazenamento é a coluna. Uma coluna é constituída pelo par nome e valor, sendo que o nome funciona como chave da coluna. Além disto cada coluna possui um valor *timestamp* que fornece informações acerca do valor temporal em que os dados foram inseridos ou atualizados permitindo entre outras coisas resolver conflitos em operações *write* ou lidar com dados que estejam desatualizados. Uma coluna pode ser interpretada como um registo numa base de dados relacional;
- **Super Coluna:** pode ser vista como uma coluna com sub-colunas. Da mesma forma que uma coluna, uma super coluna contem um par chave/valor embora neste caso o valor esteja associado a um mapa que contem várias colunas;
- **Linha:** Ao contrário do que acontece num modelo relacional, uma linha é um conjunto de colunas associadas à mesma chave primária. Relativamente ao armazenamento dos dados enquanto num modelo relacional existe um esquema previamente definido em que cada coluna de uma dada linha tem o seu espaço atribuído mesmo que o seu valor seja *NULL*, neste caso só é atribuído espaço às colunas presentes nessa linha o que se traduz num menor esforço computacional e contribui para o aumento do desempenho;
- **Família de Colunas:** são constituídas por um número infinito de linhas, sendo que cada linha contem uma *row key* e um conjunto de colunas ordenadas pela sua chave. Uma família de colunas pode ser comparada a um conjunto de linhas de uma tabela do modelo relacional em que cada linha possui uma chave e está associada a um conjunto

de colunas. A grande diferença entre os dois modelos é o fato de no modelo não relacional não existir nenhum esquema pré-definido pelo que as linhas não têm que ter exatamente as mesmas colunas e estas podem ser adicionadas a uma só linha a qualquer instante sem que seja necessário que isso aconteça nas restantes [Sadalage & Fowler 2013, página 100];

- **Super Família de Colunas:** Semelhante a uma família de colunas, uma super família de colunas é constituída por um conjunto de super colunas, sendo que são úteis para manter dados que estão relacionados juntos [Sadalage & Fowler 2013, página 101].
- ***Keyspace*:** é semelhante a uma base de dados no sistema relacional onde se encontram armazenadas todas as famílias de colunas relacionadas com uma mesma aplicação. O *keyspace* aquando da sua criação define a forma como os dados serão replicados ao longo do *cluster*;

Relativamente às famílias de coluna, embora sejam muito flexíveis, na prática não são inteiramente *"schema-less"*, pelo que cada uma deve ser criada para contar apenas um tipo de dados [Datastax 2015].

3.1.2. Modelo de Consulta

Relativamente à API do Cassandra esta consiste apenas em três simples operações, nomeadamente:

- *get(table, key, columnName)*
- *insert(table, key, rowMutation)*
- *delete(table, key, columnName)*

Aqui o argumento *columnName* pode ser relativo tanto a uma coluna que pertença a uma família de colunas, como a uma família de colunas inteira, uma super família de colunas ou até mesmo a uma coluna pertencente a uma super coluna [Laksham & Prashant 2010].

Todos os acessos aos dados são feitos através da chave, sendo retornado o valor da coluna associada a essa chave. Existem também métodos adicionais que permitem obter vários valores assim como várias colunas tendo todos como base as chaves.

O Cassandra possui uma linguagem de consulta muito similar ao SQL denominada de CQL – *Cassandra Query Language*. Uma vez que a sua *syntax* é muito similar ao SQL torna-se mais fácil aos utilizadores interagir com os dados armazenados, verificando-se uma curva de aprendizagem reduzida [Abramova et al. 2014]. No entanto, CQL não permite efetuar *joins* ou subconsultas dos dados [Sadalage & Fowler 2013, página 114].

De forma a efetuar consultas mais complexas aos dados é possível executar trabalhos *MapReduce* na base de dados através da implementação Apache Hadoop destinada ao processamento de dados distribuídos [Datastax 2015].

3.1.3. Arquitetura do Sistema

Particionamento

Uma das principais caraterísticas do Cassandra é ser altamente escalável, ou seja, distribui os dados por um conjunto de nós pertencentes ao *cluster*, sendo que o Cassandra o faz através de um *consistent hashing*. A distribuição dos nós dentro do *cluster* pode ser compreendida mais facilmente fazendo uma analogia a um espaço circular ou um anel onde a cada nó do sistema é atribuído um valor aleatório chamado *token* que define a sua posição no anel. A cada conjunto de dados identificado por uma chave é atribuído um nó através de um *hashing* à sua chave, sendo que o sistema percorre o anel no sentido dos ponteiros do relógio e o primeiro nó encontrado que seja maior que o valor da chave define a sua posição no anel. Esse nó é considerado o nó coordenador para essa chave. Cada nó torna-se assim responsável pela região do anel entre si e o seu antecessor [Karger et al. 1997]. A principal vantagem da função *consistent hashing* é que a adição ou remoção de um nó apenas afeta os seus nós vizinhos. Esta função também apresenta alguns desafios nomeadamente em relação à distribuição não uniforme dos dados e o facto de ser alheio à heterogeneidade em relação ao desempenho dos nós. O Cassandra de forma a lidar com estes problemas tem como auxilio o denominado *"Gossip Protocol"* que garante que todos os nós do *cluster* estejam em constante comunicação, estando sempre atualizados relativamente ao estado dos outros nós [Hewitt 2010, página 88]. Assim o sistema faz uma análise dos dados presentes no anel e move os nós que tenham poucos dados de forma a aliviar aqueles que estejam sobrecarregados equilibrando a quantidade de dados presentes em todos os nós [Laksham & Prashant 2010].

Esta é uma solução utilizada antes da versão 1.2 do Cassandra, sendo que após a mesma o paradigma foi alterado e em vez de um nó estar associado a um *token* e um intervalo de valores, passou a estar associado a vários *tokens*, sendo este paradigma denominado de nós virtuais (*VNodes*). Este permite que cada nó possua pequenas partições que estão distribuídas por todo o *cluster* e utiliza igualmente um *consistent hashing* para distribuir os dados, sendo que neste caso não necessita de gerar e atribuir um *token* [Datastax 2015].

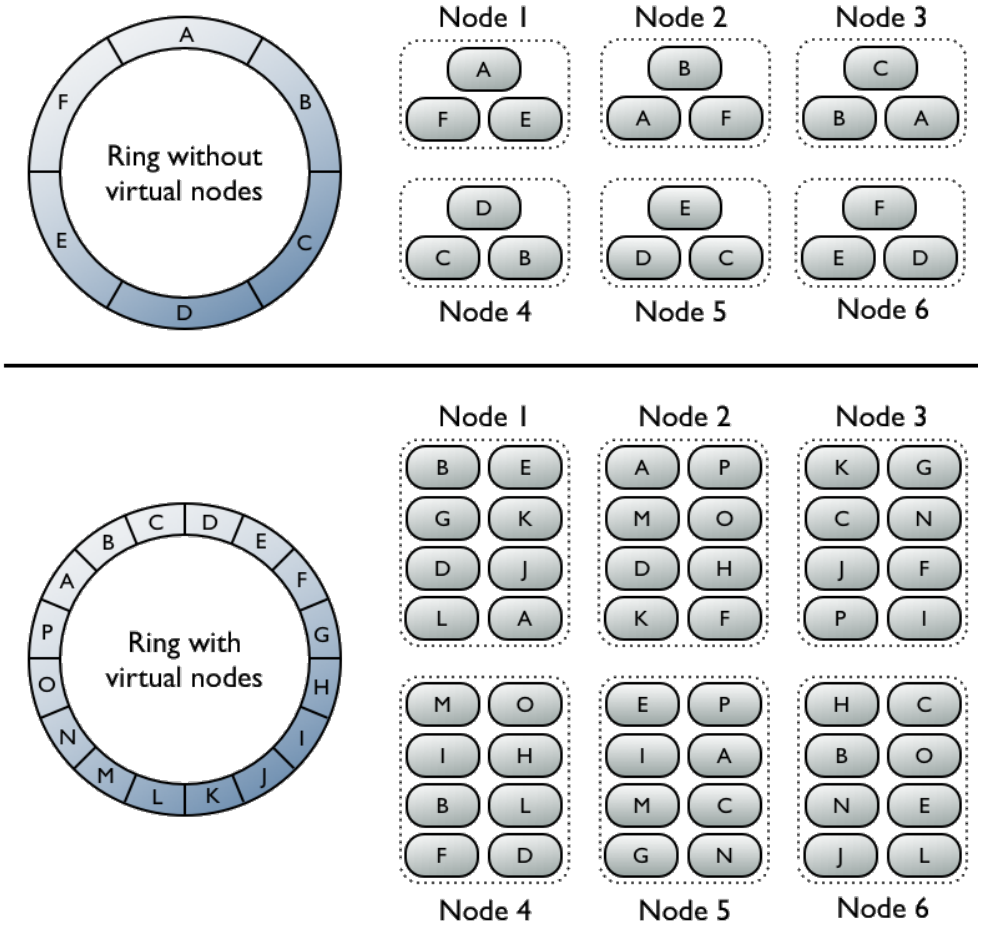


Figura 2 - Estrutura de Cluster Cassandra antes e pós versão 1.2 (retirado de [Datastax 2015])

Na figura 2 pode-se verificar ambas as estruturas, sendo que a parte superior representa o *cluster* ainda sem *VNodes*. Aqui pode-se verificar que por exemplo o intervalo "A" é replicado para os nós um, dois e três e também que cada nó possui exatamente partições consecutivas no anel, seguindo a direção do ponteiro dos relógios. A segunda parte da figura representa o paradigma *VNodes*, sendo estes escolhidos aleatoriamente e sem seguir qualquer sentido lógico [Datastax 2015].

Consistência

O Cassandra utiliza uma abordagem *quorum-based* configurável [Gifford 1979]. O cliente é capaz de especificar o nível de consistência por consulta o que lhe permite gerir a taxa de disponibilidade em relação à consistência. Um elevado nível de consistência implica que mais nós tenham que responder à consulta, oferecendo assim garantias de que os dados presentes em cada réplica são os mesmos. Caso dois nós respondam a um mesmo pedido com valores diferentes é feita uma análise ao valor *timestamp* de cada um e aquele que for mais recente é sempre retornado ao cliente, ocorrendo posteriormente uma operação *read repair* em que o nó que respondeu ao pedido com

um valor obsoleto é atualizado, de forma que todas as réplicas estejam consistentes [Hewitt 2010, página 130].

Tanto para operações *write* como *read* verifica-se a existência de vários níveis de consistência, sendo que estes têm interpretações diferentes conforme o tipo de operação. Para ambas as operações o nível que apresenta maior consistência é o *ALL*, enquanto que aquele onde se verifica uma menor consistência em prol de uma maior disponibilidade é o nível *ANY*. Para ambos os casos no nível *QUORUM* verifica-se um equilíbrio entre consistência e disponibilidade, sendo que existe a garantia de que 51% dos nós irão responder ao pedido efetuado [Datastax 2015].

O Cassandra permite assim que o cliente regule o nível de consistência que pretende tanto para operações *read* como *write* conforme as suas necessidades [Datastax 2015].

Replicação

O Cassandra recorre à replicação para alcançar elevados níveis de disponibilidade e de durabilidade do seu *cluster*, sendo cada conjunto de dados replicados por N *hosts* em que N diz respeito ao fator de replicação que pode ser configurável por instância [Sadalage & Fowler 2013, página 109].

O Cassandra oferece várias estratégias de replicação que devem ser definidas aquando da criação do *keyspace* entre as quais se destacam [Datastax 2015]:

- ***Simple Strategy*** – esta estratégia é recomendada quando se utiliza um único *data center*, sendo este definido como um grupo de nós relacionados no *cluster* com o objetivo de replicar. Nesta estratégia, a primeira réplica é definida pelo administrador do sistema e as restantes são colocadas nos seguintes nós do anel, seguindo a direção do ponteiro dos relógios e sem considerar qualquer topologia, como por exemplo, a localização do *data center;*
- ***Network Topology Strategy*** - é a estratégia recomendada para a maioria das implementações pois é mais fácil de se expandir por vários *data centers* quando necessário e especifica o número de réplicas que o utilizador pretender para *data center*. Esta estratégia tenta colocar réplicas em diferentes *racks*, de forma a prevenir problemas resultantes de quebras de energia ou falhas de rede que ao atingir um *rack* afetam todos os nós do mesmo. A decisão do número de réplicas a configurar por *data center* deve ser feita com base em algumas considerações como ser capaz de satisfazer operações *read* localmente sem problemas de latência e tendo em conta possíveis cenários de falha. As maneiras mais comuns que se verificam são duas ou três réplicas

em cada *data center* de forma a manter o sistema tolerante a falhas e também consistente.

No Cassandra verifica-se uma replicação do tipo *peer-to-peer* pois todos os nós se regem pelas mesmas regras, não havendo nenhum nó *master*, o que garante alta disponibilidade e escalabilidade [Sadalage & Fowler 2013, página 105].

Armazenamento Físico

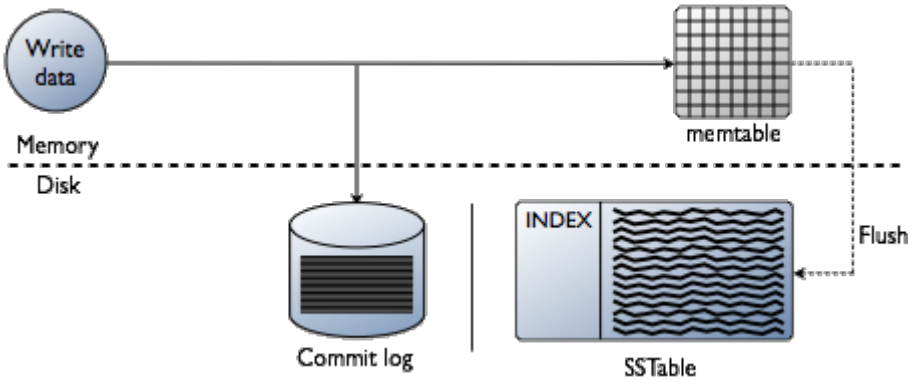


Figura 3 - Estrutura de armazenamento de *writes (*retirado de [Datastax 2015]*)*

Quando o Cassandra recebe um pedido *write*, os dados são inicialmente guardados numa estrutura do disco designada de *commitlog* e de seguida armazenados na memória RAM numa estrutura denominada *memtable*. Uma operação *write* apenas é bem-sucedida quando armazenada nas duas estruturas. Quando as operações *write* são executadas na base de dados o Cassandra disponibiliza automaticamente mais memória para a *memtable*, sendo as *writes* agrupadas e periodicamente descarregadas no disco numa estrutura designada por *SSTable*. Uma família de colunas pode ter associada vários *memtables*, sendo que apenas um se encontra em utilização e os restantes em espera para serem libertados. Caso ocorram alterações nos dados, é escrita uma nova *SSTable*, sendo que para que estas sejam apagadas é necessário que em primeiro lugar o seu estado seja definido como "apagada", para mais tarde durante o processo de compactação serem removidas do disco. O processo de compactação depois de eliminar as tabelas, agrupa os fragmentos de linhas existentes de forma a otimizar a procura de dados no disco [Hewitt 2010, página 91].

Disponibilidade

Como já foi referido anteriormente, no Cassandra todos os nós têm o mesmo peso, pelo que não existe nenhum nó *master*. Isto traduz-se em alta disponibilidade dos dados, sendo que o Cassandra

permite diversas configurações que permitem aumentar a disponibilidade em detrimento de alguma consistência nos pedidos, dependendo das necessidades da aplicação. Considere-se [Hewitt 2010, página 275]:

- R – número mínimo de nós que devem responder de forma síncrona a um pedido *read;*
- W – número mínimo de nós que devem responder de forma síncrona a um pedido *write;*
- N – corresponde ao fator de replicação, ou seja, é o número de nós que participam na replicação dos dados.

A disponibilidade é regulada pela formula (R + W) > N, sendo que sempre que esta é cumprida os dados a serem acedidos encontram-se na sua versão mais recente. Caso este cenário não se verifique a resposta aos pedidos *read* podem não se apresentar na versão mais recente dos dados pois a operação *write* associada pode ainda estar a ser replicada pelos respetivos nós, dando-se um caso de "*eventual consistency*" [Hewitt 2010, página 275].

Por outro lado, sempre que W assume o valor W=1, é dada a garantia que operações *write* estarão sempre disponíveis sem falhas a menos que todos os nós do *cluster* estejam indisponíveis, sendo que o mesmo se sucede com R. Quanto maiores os valores de R e W maior é o nível de consistência, o que nem sempre é o indicado para o bom desempenho das operações *read* e *write*. No que diz respeito ao N quanto maior o seu valor maior é o nível de durabilidade dos dados [Sadalage & Fowler 2013, página 106].

3.2. HBase

HBase é uma base de dados NoSQL que tem como base o BigTable da Google e foi desenvolvido em *Java* pela Powerset tendo-se tornado mais tarde parte do projeto Hadoop da Apache Software Foundations. O HBase é uma implementação *open-source* que executa sobre um sistema de ficheiros distribuídos denominado *Hadoop Distributed File System* (HDFS) fornecendo capacidades do BigTable ao Hadoop e que proporciona alta tolerância a falhas ao armazenar grandes quantidades de dados esparsos [Carstoiu et al. 2010].

3.2.1. Modelo de Dados

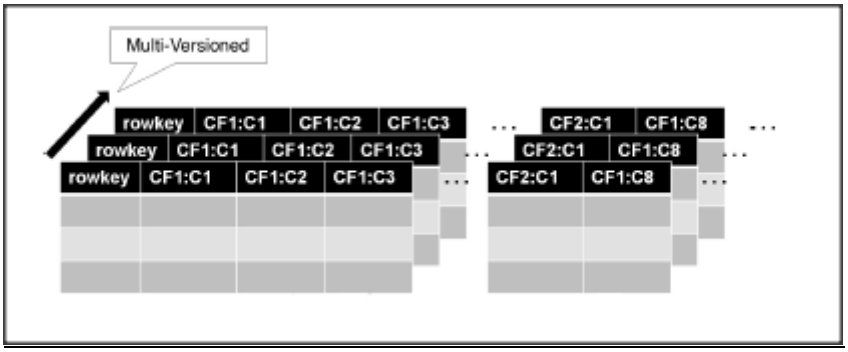


Figura 4- Modelo de Dados HBase (retirado de [Silva 2011])

O modelo de dados do HBase, também ele similar ao do BigTable é constituído pelos seguintes elementos [HBase 2015]:

- ***Namespace* –** É constituído por tabelas e corresponde a uma base de dados no modelo relacional;
- **Tabela** – Os dados no HBase encontram-se organizados por tabelas, sendo cada tabela constituída por várias linhas;
- **Linha** – Dentro de cada tabela os dados estão armazenados de acordo com a sua linha, sendo que cada linha é composta por uma *row key* e um número arbitrário de colunas. As linhas encontram-se ordenadas lexicograficamente pela sua *row key* pelo que a conceção das mesmas se torna um fator importante de forma a armazenar linhas que estão relacionadas próximas umas das outras de forma a otimizar o desempenho;
- **Família de coluna** – Os dados dentro de cada linha estão agrupados por famílias de colunas, sendo que estas devem ser definidas aquando da criação das tabelas. Cada linha de uma tabela possui as mesmas famílias de coluna, embora algumas não armazenem qualquer informação. As famílias de coluna possuem um *column qualifier* de forma a classificar o formato dos dados como por exemplo *pdf* ou *html*, sendo que estes podem variar bastante;
- **Coluna** – As colunas são agrupadas em famílias de colunas, sendo que todas as colunas da mesma família de colunas possuem o mesmo prefixo. As colunas possuem o seguinte formato: "*column family : column qualifier";*
- **Célula** – Uma célula representa a combinação de uma linha, uma família de coluna e um *column qualifier* e contem um valor aliado a um *timestamp* que representa a versão deste. Desta forma quando os valores são consultados é sempre retornada a versão mais

recente dos dados, sendo que o utilizador pode especificar qual versão ou quais versões pretende.

As atualizações das linhas são consideradas atómicas, abrangendo deste modo operações *read* e *write* ao mesmo tempo. As tabelas HBase são particionadas horizontalmente de forma automática em regiões pelo que cada região contém um conjunto de linhas delimitadas pela primeira linha e a última, sendo que esta já não pertence à região. Cada região possui ainda um identificador gerado aleatoriamente [HBase 2015].

3.2.2. Modelo de Consulta

O HBase fornece uma *API Java* que pode ser utilizada para realizar consultas, fornecendo operações básicas como *get*, *put*, *delete* e *update*. A função *scan* permite ainda lidar com intervalos de linhas, sendo capaz de selecionar quais as colunas a serem retornadas ou o número de versões de cada célula. Através da utilização de filtros é possível combinar colunas e selecionar as versões pretendidas, utilizando intervalos de tempo em que o utilizador deve especificar o tempo de início e de fim para a consulta pretendida. O HBase não possui nenhuma linguagem específica tal como o SQL para efetuar consultas aos dados pelo que recorre a outro tipo de técnicas [Lars 2011, página 75].

Para consultas mais complexas recorre a trabalhos *MapReduce* que executam subjacentemente a uma infraestrutura Hadoop denominada *Pig* que permite efetuar consultas aos dados através de uma linguagem de programação de alto nível muito semelhante ao SQL, conhecida como *Pig Latin*. A combinação desta linguagem com a *framework MapReduce* permite o processamento de grandes quantidades de dados em períodos de tempo razoáveis [Lars 2011, página 263].

Também é possível aceder ao HBase através das API *REST*, *Thrift* e ainda através de uma *Shell* HBase que permite aos utilizadores uma interação direta com o mesmo [HBase 2015].

3.2.3. Arquitetura do Sistema

Armazenamento Físico

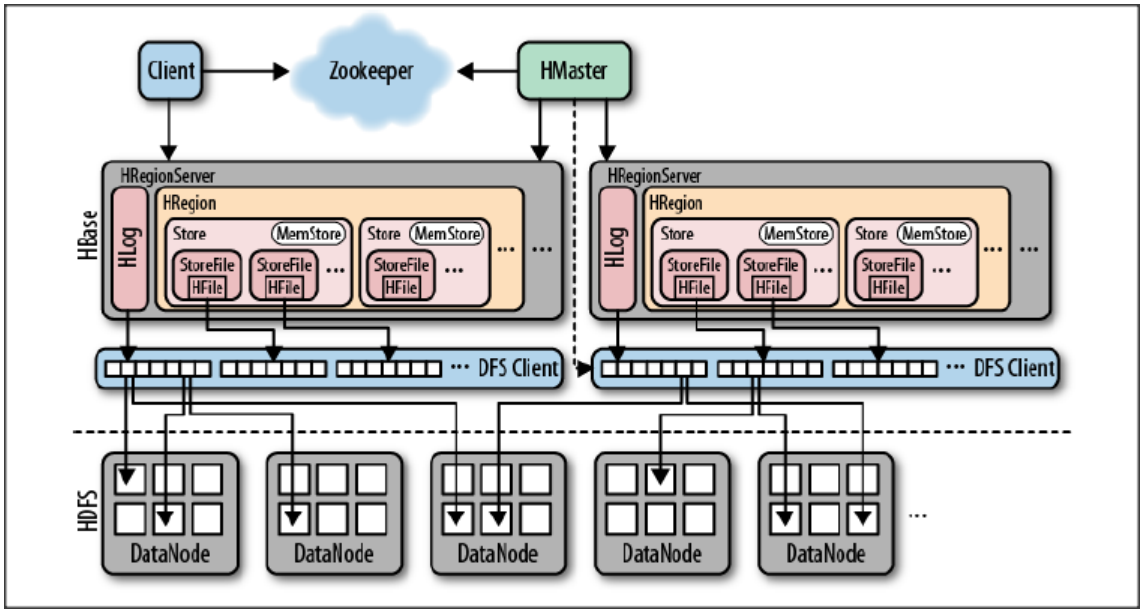


Figura 5- Visão geral de como o HBase armazena os seus ficheiros no HDFS (retirado de [Lars 2011])

Fisicamente no HBase as tabelas encontram-se distribuídas por intervalos de linhas denominados regiões. Uma região é a menor unidade de armazenamento distribuído, sendo que regiões diferentes são armazenadas em diferentes *RegionServers* pelo *master*. Uma tabela quando é criada está associada a apenas uma região, sendo que à medida que os dados vão sendo inseridos na tabela a região vai crescendo e quando o seu tamanho atinge o limite, o intervalo de linhas que a região abrange é divido para metade. Desta forma, à medida que mais dados vão sendo inseridos, mais regiões vão sendo criadas. Cada região possui ainda um *Hlog* e um *HStore*. Um *HStore* é o ponto central do armazenamento no HBase e consiste num ficheiro *MemStore* e um ou mais *StoreFiles*. Um *HStore* gere as famílias de coluna, daí ser importantes as colunas localizadas nas famílias de coluna terem caraterísticas comuns. Os dados inicialmente são gravados na *MemStore* e quando esta está cheia são libertados para o disco, criando um novo *StoreFile*. Aqui um indicador é enviado para o *HLog* que é onde são registadas todas as alterações que ocorrem aos dados no HBase. Assim que o número de *StoreFiles* atinge um limite previamente configurado é realizada uma compactação que reúne todos os *StoreFiles* num só [Yang et al. 2011].

O Hbase possui ainda uma tabela “hbase:meta” armazenada no Zookeeper que contem uma lista atualizada do estado, histórico e localização de todas as regiões do sistema [HBase 2015].

Sistema de Ficheiros Distribuídos Hadoop

Dado que o HBase executa sobre um HDFS torna-se necessário compreender a arquitetura deste, principalmente em relação ao armazenamento de ficheiros, à forma de lidar com falhas e à replicação de blocos.

O HDFS adota uma arquitetura *master-slave* em que um nó é o *master*, também conhecido como *NameNode* enquanto os restantes nós são denominados *DataNodes*. O *NameNode* coordena o *namespace* e operações no sistema de ficheiros como abrir, fechar ou alterar o nome dos mesmos. Além disto tem também a função de manter os metadados associados ao sistema tais como as permissões de acesso aos dados e localização dos ficheiros. No que diz respeito aos *DataNodes* encontram-se espalhados pelo *cluster* e têm como responsabilidade o armazenamento de dados que é feito em blocos [Borthakur 2013].

Um *DataNode* é responsável por atender aos pedidos *read* e *write* dos clientes aos ficheiros, sendo também encarregue de criar os blocos, apagá-los e replicá-los conforme as instruções dadas pelo *NameNode*. Dado que o HDFS foi desenvolvido em *Java* qualquer máquina que suporte *Java*, pode executar *software* para *NameNode* e *DataNode*. As decisões relacionadas com a replicação de blocos são sempre tomadas pelo *NameNode*, que recebe regularmente de cada *DataNode* do *cluster* informações acerca do movimento e proporção dos blocos utilizados. O mecanismo de replicação consiste em manter cópias redundantes dos dados no sistema, de forma que em caso de ocorrência de alguma falha exista sempre uma cópia disponível. O HDFS, por defeito, armazena três cópias em cada bloco para assegurar a confiabilidade, disponibilidade e a tolerância a falhas do sistema. A política de replicação habitual para um *data center* é ter duas máquinas réplica para o mesmo *rack* e uma réplica para um nó localizado noutro *rack*, sendo que isto limita o tráfego de dados entre *racks* e a hipótese de falha de um *rack* inteiro é muito menor que a hipótese de falha de um nó. Para minimizar a latência no acesso aos dados, o HDFS tenta aceder aos dados a partir da réplica mais próxima, pelo que se esta se encontrar no mesmo *rack* terá prioridade de escolha [Vora 2011].

De acordo com um estudo realizado por Zhou Yinan, o HDFS oferece o melhor desempenho no processamento de grandes volumes de dados. A estratégia de replicação é o seu ponto central e exerce grande influência no desempenho das operações *read* e *write*, sendo o facto de a posição dos dados redundantes poder ser ajustada, a grande diferença entre o HDFS e outros sistemas de ficheiros distribuídos [Yang et al. 2011].

Replicação

O HBase fornece um mecanismo de replicação do *cluster* que permite manter os estados entre *clusters* sincronizados utilizando um *write-ahead log* (*WAL*) do *cluster master* para propagar as alterações aos dados. Um *WAL* garante que as alterações aos dados sejam inicialmente gravadas num *Hlog* antes de serem aplicadas, fornecendo assim garantias de atomicidade e durabilidade dos dados [HBase 2015].

O modelo de replicação utilizado pelo HBase é do tipo *master-slave* em que um *cluster* HBase é o *master*, ou seja, o responsável por criar e replicar os dados e os restantes são denominados *slaves*, ou seja, os que recebem os dados por via da replicação. A replicação é feita de forma assíncrona, o que significa que os *clusters* podem estar geograficamente distantes, que a ligação entre eles pode falhar durante algum tempo e que as linhas inseridas no servidor *master* podem não estar disponíveis ao mesmo tempo nos servidores *slave*, verificando-se uma *eventual consistency* dos dados [HBase 2015].

A figura 6 representa o processo de replicação no HBase.

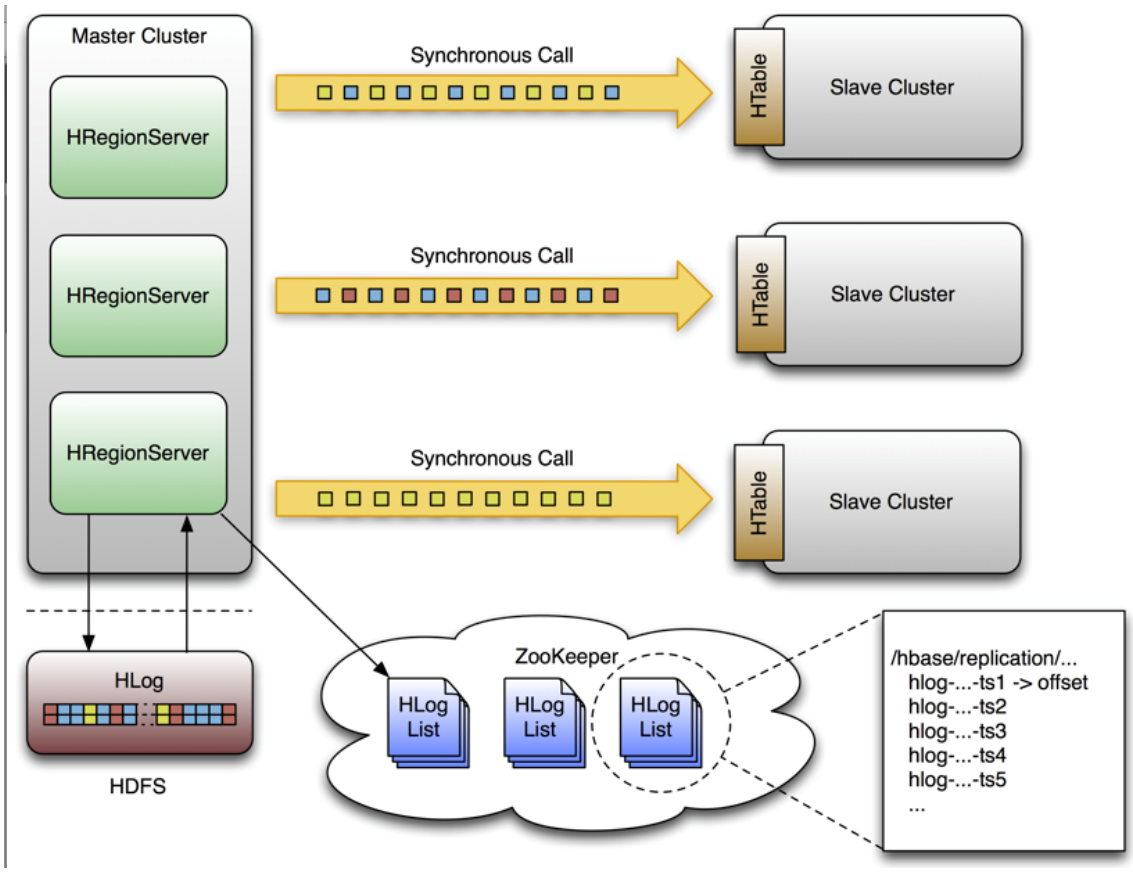


Figura 6 - Processo de Replicação HBase (retirado de [HBase 2015])

Os *WALs* de cada *RegionServer* são a base da replicação do HBase e devem manter-se no *HDFS* enquanto são necessários para replicar dados para qualquer *cluster slave*. Cada *RegionServer* lê a partir do *log* mais antigo que necessite de replicar e mantem o registo do seu progresso dentro do Zookeeper de forma a simplificar a recuperação de falhas. O Zookeeper é um serviço que faz o papel de coordenador e gere as principais atividades de replicação de forma a simplificar o processo de recuperação de falhas. O marcador da posição que indica o progresso, assim como a quantidade de *WAL's* por processar para todos os *clusters slave* pode ser diferente para todos [HBase 2015].

Consistência

Relativamente à consistência o HBase apresenta-se por defeito como estritamente consistente, dado que todos os valores surgem numa única região definida por um intervalo de chaves e cada região está atribuída a um único *RegionServer* de cada vez, o que garante que todas as operações *write* se realizam numa ordem e que todas as operações *read* acedem aos dados confirmados mais recentes.

No entanto, uma vez que a replicação no HBase é feita de forma assíncrona, encontrando-se a replicação ativa, quando são efetuados *reads* através de um servidor *slave*, apenas é garantida uma *eventual consistency* pois as atualizações aos dados por parte do *master* não chegam ao mesmo tempo a todos os *slaves* [HBase 2015].

4. Conclusões e trabalho futuro

Apesar de serem várias as semelhanças entre o Cassandra e o HBase existem também algumas diferenças importantes a verificar. Ao nível da infraestrutura o Cassandra possui apenas um único tipo de nó em que todos eles possuem as mesmas responsabilidades, enquanto no HBase isso não se verifica, existindo um nó considerado *master* e os restantes *slaves* com diferentes responsabilidades. Como tal, o HBase carateriza-se por ter uma forte consistência e ser otimizado para operações *read*. Por sua vez o Cassandra carateriza-se pela alta disponibilidade, sendo capaz de regular o nível de consistência pretendido, garantindo uma *eventual consistency* dos dados. O sistema Cassandra é otimizado para operações *write*. Assim, relativamente ao teorema CAP o HBase é considerado um sistema do tipo CP (consistente e tolerante a falhas) e o Cassandra um sistema AP (disponível e tolerante a falhas).

Relativamente à comunicação entre os nós, o Cassandra utiliza uma estratégia denominada *Gossip Protocol* enquanto o HBase recorre a um sistema externo denominado *Zookeeeper*. Quanto ao particionamento dos dados o HBase adota uma estratégia ordenada enquanto o Cassandra o faz

aleatoriamente. A replicação dos dados no HBase é feita de forma assíncrona enquanto no Cassandra se verifica de forma síncrona. Em relação ao modelo de consulta o Cassandra possui uma linguagem CQL muito semelhante ao SQL que permite ao utilizador uma rápida aprendizagem e suporta índices secundários por coluna de forma a melhorar a consulta aos dados. O HBase apesar de não suportar índices secundários suporta outros mecanismos que fornecem as mesmas funcionalidades, sendo também capaz de filtrar o acesso aos dados o que resulta em pesquisas mais rápidas. O HBase, ao contrário do Cassandra é também capaz de suportar nas suas consultas agregados como SUM, AVG, MAX. Independentemente disto, tanto num sistema como noutro é importante que colunas com padrões de acesso semelhantes sejam mantidas na mesma família de colunas de forma otimizar o desempenho das consultas. Daí ser importante planear previamente a organização das famílias de colunas.

Posto isto o trabalho futuro passa por identificar um caso de estudo, com dados a rondar os milhões de registos e comparar o desempenho dos dois sistemas ao nível do acesso e consulta aos dados em diferentes cenários e retirar conclusões acerca do desempenho de cada um dos sistemas.

5. Referências

Abramova, V., Bernardino, J. e Furtado, P.," Experimental Evaluation of NoSQL Databases" *International Journal of Database Management Systems*, 6, 3 (2014), 1–16.

Atzeni, P., Bugiotti, F. e Rossi, L., "Uniform access to NoSQL systems". *Information Systems*, 43, 2014, 117–133.

Bordhakur, D., *HDFS Architecture Guide*, http://hadoop.apache.org/docs/r1.2.1/hdfs_design.html, (20 de Julho de 2015), 2013.

Carstoiu, D., Cernian, A. e Olteanu, A., "Hadoop Hbase-0.20.2 performance evaluation", *New Trends in Information Science and Service Science (NISS), 2010 4th International Conference on*, 2010, 84-87.

Datastax, *DataStax Cassandra 2.2 for Linux*, http://docs.datastax.com/en/cassandra/2.2/cassandra/cassandraAbout.html, (5 de Maio de 2015), 2015

Gifford, D.K.,"Weighted voting for replicated data", *Proceedings of the seventh symposium on Operating systems principles - SOSP '79*, 1979, 150–162.

HBase, A., *Apache HBase ™ Reference Guide*, http://hbase.apache.org/book.html, (5 de Maio de 2015), 2015.

Hewitt, E., *Cassandra - The Definitive Guide*, O'Reilly Media, Inc., 2010

Hossain, S.A. e Moniruzzaman, A., "NoSQL Database : New Era of Databases for Big data Analytics - Classification , Characteristics and Comparison", *International Journal of Database Theory and Application*, 6, 4 (2013), 1–14.

Karger, D. et al., "Consistent Hashing and Random Trees", *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing - STOC '97*, 1997, 654–663.

Kuznetsov, S.D. e Poskonin, A. V., "NoSQL data management systems", *Programming and Computer Software*, 40, 6 (2014), 323–332.

Laksham, A. & Prashant, M., "Cassandra: a decentralized structured storage system", *ACM SIGOPS Operating Systems Review*, 2010, 1–6.

Lars, G., *HBase: The Definitive Guide*, O'Reilly Media, Inc., 2011.

Leavitt, N., "Will NoSQL Databases Live Up to Their Promise?", *Computer*, 43, 2010, 12–14.

Robinson, I., Webber, J. e Eifrem, E., *Graph Databases*, O'Reilly Media, Inc., 2013.

Sadalage, P. e Fowler, M., *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*, Addison-Wesley, 2013.

Silva, C., *Data Modeling with NoSQL: How, When and Why*, Tese de mestrado integrado, Engenharia Informática e Computação, Faculdade de Engenharia da Universidade do Porto, 2010.

Vora, M.N., "Hadoop-HBase for large-scale data", *Proceedings of 2011 International Conference on Computer Science and Network Technology, ICCSNT 2011*, 2011, 601–605.

Weber, S. e Strauch, C., "NoSQL Databases", *Lecture Notes Stuttgart Media*, 2010, 1–8.

Yang, J., Tang, D. e Zhou, Y., "A distributed storage model for EHR based on HBase", *Proceedings - 2011 4th International Conference on Information Management, Innovation Management and Industrial Engineering, ICIII 2011*, 2011, 369–372.

Cruzamentos e Encruzilhadas – a importância de um SI para um serviço de psicologia escolar

Pedro Miguel de Oliveira[1], Abílio Oliveira[2], António Martins[3]

1) Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
pmoal@iscte-iul.pt

2) Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
abilio.oliveira@iscte.pt

3) Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
antonio.martins@iscte.pt

Resumo

Os desafios que se colocam hoje em dia aos adolescentes são cada vez mais complexos. Desde os variados estímulos externos, às características individuais, muitos são os contributos que vão moldando a sua forma de pensar, aprender e (re)agir. Neste sentido, os contextos em que eles se inserem assumem, progressivamente, mais relevância no seu dia-a-dia, nomeadamente no que toca ao meio escolar. É neste enquadramento que se destaca a importância de existir um sistema de informação que permita optimizar o trabalho dos psicólogos, e que se traduza numa clara mudança no desempenho organizacional de um serviço de psicologia escolar. Pretendemos assim modelar e propor um SI que permita reportar informação relevante, cruzando dados quantitativos e qualitativos resultantes de diferentes fontes, nomeadamente testes psicotécnicos aplicados num programa de orientação escolar e vocacional. Suportado na qualidade da informação gerada, cada psicólogo poderá desenvolver um acompanhamento personalizado, mais próximo dos interesses do aluno.

Palavras-chave: sistema de informação, serviço de psicologia escolar, orientação vocacional

1. Introdução

"Porque o destino do ser humano é ser feliz… porque a alegria é possível!" Assim espelha Rubem Alves a importância do professor despertar, no aluno, a satisfação de conhecer o desconhecido, na capa do seu livro *A Alegria de Ensinar*. Consequentemente "é necessário que os educadores tenham tempo, condições e disposição para interagir com os educandos, ouvindo-os e dando-lhes espaço à sua imaginação, criatividade, aprendizagem e desenvolvimento." [Oliveira 2011]

Este artigo não tem propriamente por base o professor, procura sim a relevância dos diferentes contextos no desenvolvimento de cada criança, de cada adolescente, e a sua vontade de aprender,

a sua capacidade de ser feliz. “A pedagogia, arte e ciência de ensinar, não pode reduzir-se simplesmente a uma metodologia. Deve incluir a visão do mundo e da pessoa humana a ser educada.” [GRACOS 1994] Ou como diria Agostinho da Silva, “É melhor aprender latim ou melhor aprender matemática? É melhor não ser estúpido.”

É de acordo com um paradigma, que vai para além das paredes de uma sala de aula, onde o ser e o estar se confrontam tantas vezes na procura de um saber, que é proposto um sistema de informação que procure optimizar as dinâmicas de um serviço de psicologia escolar. A possibilidade de levar por diante uma melhoria na organização processual deste serviço pode permitir que a acção do psicólogo seja renovada, possibilitando-lhe um trabalho de campo, junto de cada aluno, mais efectivo e consequente. “O pedagogo-terapeuta deve propor-se a descobrir o que é o funcionamento mental, desligando-se do preconceito escolástico e retrogrado de que se funciona mentalmente apreendendo as técnicas que resultam da utilização de uma cultura escrita e de que a inteligência é uma função mecânica.” [Santos 1976]

Esta acção directa no desenvolvimento do adolescente, organizada e centrada nele mesmo, acabará por produzir efeitos numa cadeia de elos académicos que terá como sentido primário o próprio indivíduo. Ou seja, tal evolução acabará por se repercutir numa melhor integração académica do mesmo. Consequentemente, e de um modo extrapolado, poderá afirmar-se que estaremos a promover uma desejada inclusão social do adolescente numa sociedade que procura a evolução científica em comunhão com a globalização do equilíbrio e da integração social do cidadão no seu meio envolvente.

Indo ao encontro destes princípios, apresentamos aqui algumas fundamentações sobre uma realidade de um sector tão importante no seio escolar, como são os serviços de psicologia. Partindo de um estudo de caso, realizado em contexto escolar, propõe-se a modelização de um sistema de informação (SI) que consequentemente dará origem a um protótipo. Recorrendo à versatilidade do *Microsoft Excel*, será assim criado um modelo de simulação, como mecanismo facilitador do trabalho do psicólogo, libertando-o de um conjunto de tarefas administrativas.

Esta ferramenta está a ser desenvolvida no âmbito de uma dissertação de mestrado e, para além do seu enquadramento académico, visa contribuir para uma reflexão sobre a importância de existir um sistema de informação que responda às necessidades apresentadas por um serviço de psicologia escolar. Pretende-se assim reforçar a importância que os serviços de psicologia têm no contexto atual e futuro das nossas escolas.

2. Um sistema de informação num serviço de psicologia escolar… Alternativa ou necessidade?

Quando se começa a abordar este assunto pela primeira vez é inevitável colocar-se a questão: *Em que medida pode um sistema de informação organizar e optimizar um serviço de psicologia escolar?*

A resposta pode ser encontrada ainda no final do século passado… "A Content Management System allows content to be stored, retrieved, edited, updated, controlled, then output in a variety of ways such that the incremental cost of each update cycle and output production shrinks dramatically over time. It is a great concept." [Kartchner 1998]

Mais recentemente pode ser vista outra resposta quando lemos que "vivemos hoje numa sociedade da informação e do conhecimento, também caracterizada por ser uma sociedade em rede, onde a dependência da informação e interdependência entre organizações são uma realidade. (…) Quase todos os tipos de profissionais necessitam de utilizar sistemas de informação nas suas actividades diárias, e, em muitos casos, são chamados a participar na implementação dos mesmos." [Alturas 2013]

Neste sentido, e levando estas respostas para a esfera da psicologia educacional, é importante salientar em primeiro lugar que "Serviços-escola de psicologia constituem um importante veículo de transposição de conhecimentos da universidade para a comunidade. Eles são espaços privilegiados para a realização de pesquisa de desenvolvimento e avaliação de práticas fundamentadas no conhecimento científico." [Elias, Marturano e Motta-Oliveira 2013]

Por conseguinte, e perante o papel da escolaridade na vida do adolescente, "existe uma forte correlação entre manter e continuar a escolaridade, por um lado, e a existência de dificuldades na adolescência, por outro, estando alguns destes problemas correlacionados com a qualidade dos estudos." [Braconnier e Marcelli 2000].

Consequentemente, e talvez o destaque mais importante face à necessidade de existência de um serviço de psicologia escolar de qualidade, é o facto de cada aluno ter direito a crescer em harmonia consigo próprio e com o meio que o rodeia. "Os psicólogos educacionais têm agora o desafio de poder contribuir para a conceptualização, adequação e eficácia da educação para todas as crianças" [Lindsay 2007]. Além disso, e segundo recomendação do Conselho Nacional da Educação, "a orientação escolar e profissional deve desempenhar um papel estratégico na elevação dos níveis de qualificação da população portuguesa, ao facilitar o acesso à informação sobre a oferta de educação e formação disponível e ajudar jovens e adultos na construção de uma identidade pessoal e vocacional" [Conselho Nacional da Educação 2011].

Assim sendo, perante as potencialidades de um sistema de informação, e face à importância que a medição acaba por ter no âmbito da psicologia, em particular no da orientação escolar e vocacional do indivíduo, pode-se aglutinar conceitos afirmando que "a construção de sistemas de observação e medida dotados de qualidades adequadas é essencial ao avanço do conhecimento (…)" [Moreira 2004]. Por outro lado, e concorrendo num sentido mais lato, salienta-se que "a orientação escolar e profissional representa uma actividade de interface entre Educação/Formação Profissional e inserção na vida activa e constitui o veículo que poderá permitir uma adaptação harmoniosa e equilibrada entre Sistema Educativo e Sistema Económico" [Pinho 1997].

É assim perante a clara necessidade de recolha, organização, tratamento e arquivo de diferentes fontes de informação que vão surgindo ao longo do percurso escolar de uma criança, ou adolescente, que urge a necessidade de conhecer melhor o trabalho do psicólogo escolar. Ao ser desenvolvida uma pequena pesquisa sobre o funcionamento de alguns serviços de psicologia escolar, fica-se com uma sensação empírica de que muito do trabalho desenvolvido continua a ser arquivado em processos físicos que se acumulam como resmas de papel. A rede de comunicação e de obtenção de informação não tem a agilidade desejada, os gastos associados são maiores e a potencial perda de informação e de tempo acaba por ser uma realidade, na medida em que, a transição de processos de um sector para outro, ou mesmo de uma escola para outra, nem sempre prima pela eficácia adequada.

De qualquer modo, e no que toca a aplicação isolada de diferentes tipos de testes psicotécnicos, consegue-se perceber que alguns destes podem, em grande parte, ser trabalhados manualmente ou informaticamente pelo psicólogo. Contudo, depois de aplicados no terreno, estes entram muitas vezes numa lógica de impressão e arquivamento físico relegando-se para segundo plano a importância da existência de um bom sistema de informação, capaz de centralizar toda a informação gerada. É por isso que a sua existência permitirá complementar, de um modo mais eficiente, a acção do psicólogo na comunidade escolar em que se encontra inserido. No entanto, e face a um sentimento de que será mesmo necessária a "imposição" de um sistema centralizado, eficiente e ágil, é necessário ter em linha de conta um cuidado relevante em todo este processo. Salvaguardar a confidencialidade que é devida a cada aluno uma vez que "os/as psicólogos/as têm a obrigação de assegurar a manutenção da privacidade e confidencialidade de toda a informação a respeito do seu cliente, obtida directa ou indirectamente, incluindo a existência da própria relação, e de conhecer as situações específicas em que a confidencialidade apresenta algumas limitações éticas ou legais." [Código Deontológico da Ordem dos Psicólogos Portugueses 2011]

É por estas e outras razões que haverá uma necessidade clara de conhecer, aprofundar e enriquecer ainda mais este tema. O desenvolvimento e a implementação de um sistema de

informação que minimize as encruzilhadas sentidas no seio de um serviço de psicologia escolar, e que possibilite um bom cruzamento da informação existente. Em suma, um sistema que permita organizar e optimizar o trabalho desenvolvido pelo psicólogo, despoletando sentidos e razões com repercussões efectivas no desenvolvimento de cada aluno.

3. Conceptualização de uma alternativa

É importante intuir sobre as dinâmicas de base que o sistema de informação tem que oferecer, procurando perceber os requisitos que decorrem das acções desenvolvidas por diferentes serviços de psicologia. Apesar da concepção apresentada partir de um levantamento de requisitos que resultou de um estudo caso, não será falacioso generalizar alguns dos conceitos base num projecto genérico que vá ao encontro das necessidades apontadas.

Sabendo que é necessário preservar a identidade dos alunos e informações pessoais associadas, sempre que o psicólogo tentar aceder ao SI deverá proceder à respectiva validação. A partir daí as suas ações serão as mais diversas, como pode ser observado no diagrama de casos de uso.

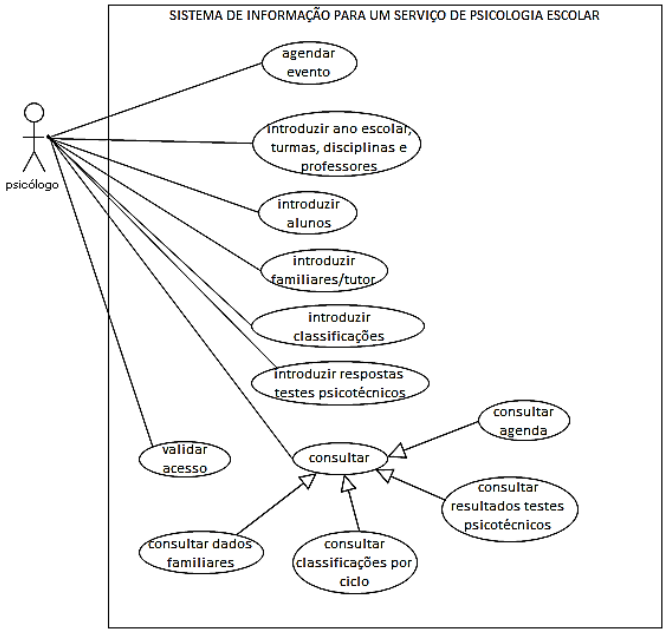


Figura 1 – Diagrama de casos de uso

Uma vez dentro do sistema, o psicólogo poderá dar início ao ano lectivo em causa, apontando as principais datas que referenciam o decurso do mesmo. De seguida, deverá introduzir e associar a informação relativa às turmas, disciplinas e professores, de acordo com o ano lectivo em causa. No que concerne aos alunos, e de acordo com a sinalização que vai sendo feita dos casos que necessitam de um maior acompanhamento, deverão ser registadas diferentes informações

relativas à sua situação académica, ao seu encarregado de educação, familiares próximos e respectivo agregado familiar.

O sistema de informação permite também o acesso a uma agenda própria, onde o psicólogo poderá introduzir ou consultar um evento de acordo com a sua data ou assunto. Para além disso, há uma mais-valia que possibilita ao psicólogo, após introduzir as respostas que o aluno dá em determinados testes psicotécnicos, obter os resultados finais desses mesmos testes. Este é um processo que tem um grande ganho de eficiência uma vez que, aliado à poupança de tempo no tratamento da informação gerada, os resultados extraídos apresentam um maior grau de fiabilidade. Na continuada procura de ganhos de produtividade, todo o trabalho produzido pode ser armazenado numa *cloud* podendo deste modo ser consultado em qualquer parte.

Em suma, a existência desta possibilidade poderá libertar o psicólogo de um espaço confinado, que muitas vezes é limitador de um centro de actuação que deve ser muito mais vasto. Vejamos agora o diagrama de classes que desenvolve alguns dos requisitos que foram sendo levantados.

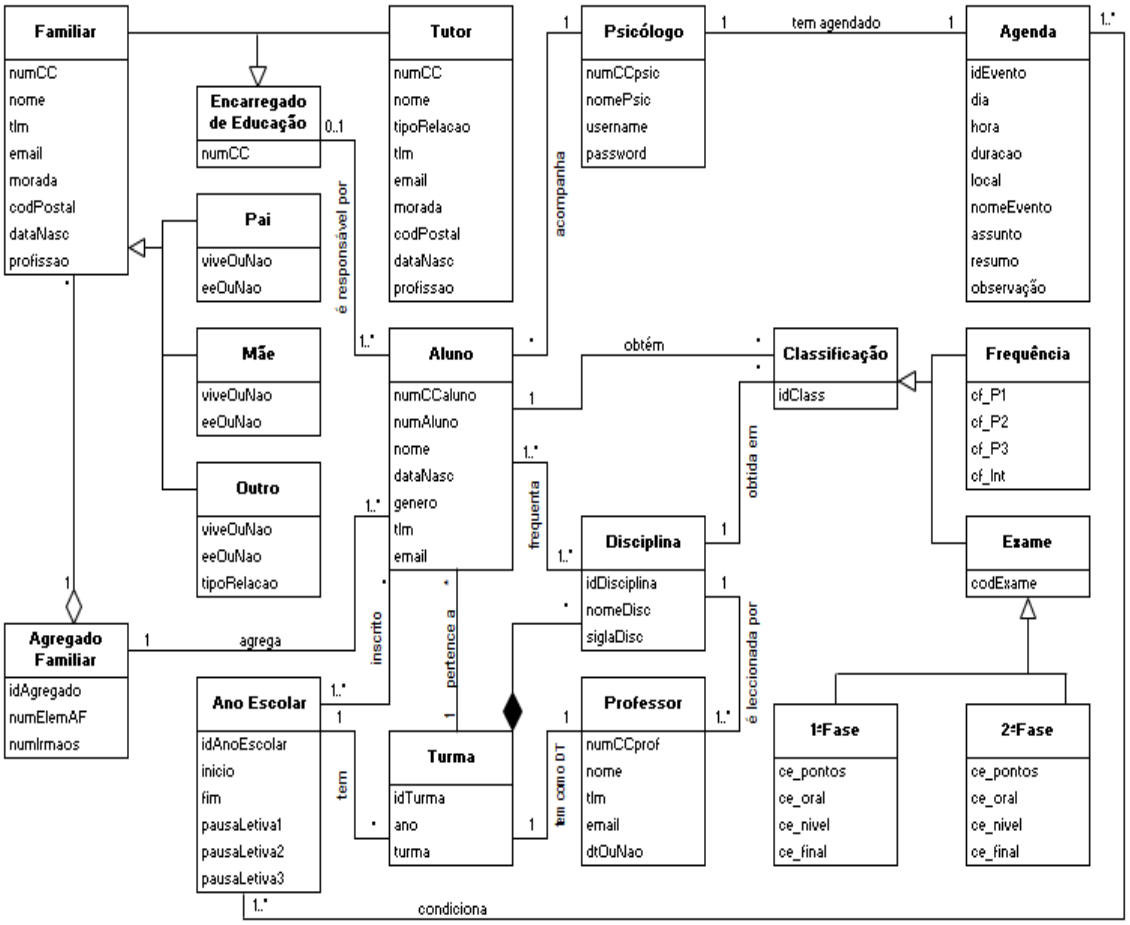


Figura 2 – Diagrama de classes

4. Protótipo em desenvolvimento

A concepção do protótipo está, neste momento, em fase de discussão, tendo em conta o modelo apresentado. A sua construção terá por base o *Microsoft Excel*, nomeadamente diferentes funcionalidades que relacionam bases de dados e funções de procura. Também as funções estatísticas serão utilizadas em determinados eventos, enquadrando um conjunto de dados como as classificações escolares ou os resultados dos testes psicotécnicos, de cada estudante.

Recorrendo ao *Microsoft Visual Basic for Applications*, serão desenvolvidas algumas macros que possibilitarão uma melhor usabilidade do sistema de informação. Consequentemente há ainda a preocupação de conceber um interface *user-friendly*, para que o psicólogo sinta o gosto e a vontade de utilização desta plataforma de trabalho.

5. Conclusão possível – fase aos desenvolvimentos esperados

O resultado final pretendido será o desenvolvimento de um protótipo de um sistema de informação, adequado às necessidades e requisitos de um serviço de psicologia escolar, que vise, em particular, uma melhor gestão de um programa de orientação escolar e vocacional. A qualidade da informação gerada, com ganhos de tempo e de eficácia, serão os grandes objectivos técnicos que suportarão um trabalho que se pretende diferenciado e de proximidade.

A partir da desejada implementação do referido protótipo terá de ser feita uma análise sob o ponto de vista do utilizador, recorrendo a um pequeno questionário, que servirá de base a uma análise e discussão dos problemas, das carências e das mais-valias observadas. Individualmente e/ou em grupo, serão ainda desenvolvidas propostas de melhoria do produto e de exploração do mesmo.

Consequentemente, e tendo como suporte toda a fundamentação teórica e prática, será desejável que a criação e a avaliação do protótipo seja o sustento de uma proposta válida e exequível na oferta de uma inovação progressista, que possibilite uma evolução num importante serviço escolar. Por via de um bom sistema de informação, bem suportado pelas tecnologias da informação, teremos então, como última instância, uma clara melhoria no trabalho levado a cabo por cada psicólogo. Deste modo, a possibilidade de um acompanhamento personalizado de cada adolescente será um contributo directo para o seu crescimento individual, académico e profissional.

6. Referências

Alturas, B., *Introdução aos Sistemas de Informação Organizacionais*, Lisboa, Edições Sílabo, 2013.

Alves, R., *A alegria de ensinar*, Porto, Edições ASA, 2003.

Braconnier, A., Marcelli, D., *As mil faces da adolescência*, Lisboa, Confrontações e Climepsi Editores, 2000.

Conselho Nacional de Educação, *Estado da Educação 2011 – A qualificação dos portugueses*, Portugal, Editorial do Ministério da Educação e Ciência, 2011.

Elias, L. C. S., Marturano, E. M. & Motta-Oliveira, A. M. A., *Teoria, pesquisa e prática em serviço-escola de psicologia*, Saúde e Transformação Social, v. 4, n. 3, p. 121-129. http://incubadora.periodicos.ufsc.br/index.php/saudeetransformacao/article/view/1978/3024 , (15 de Junho de 2015), 2013.

Grupo de Reflexão e Análise dos Colégios da Companhia de Jesus, *Pedagogia inaciana – uma abordagem prática*, Companhia de Jesus, 1994.

Kartchner, C., *Content Management Systems: Getting from concept to reality*, http://dx.doi.org/10.3998/3336451.0003.408, (4 de Janeiro de 2015), 1998

Lindsay, G., *Educational psychology and the effectiveness of inclusive education/mainstreaming*, British Journal of Educational Psychology 77, 2007, 1-24.

Moreira, J. M., *Questionários: Teoria e prática*, Coimbra, Livraria Almedina, 2004.

Oliveira, A., *O desafio da vida*, Lisboa, Coisas de Ler, 2011.

Ordem dos Psicólogos Portugueses, *Código Deontológico*, Ordem dos Psicólogos Portugueses, 2011

Pinho, M. F. D. A., *A orientação escolar e profissional numa dimensão europeia*, www.ipv.pt/millenium/ect5_pinho.htm , (4 de Janeiro de 2015), 1997.

Santos, J., *Pedagogia terapêutica*, http://www.casadapraia.org.pt/pedagogia/pedagogia_terapeutica.pdf , (15 de Junho de 2015), 1976

Document-Based Databases: Estudo exploratório no âmbito das Bases de Dados *NoSQL*

Gonçalo Sousa [1], José Luís Pereira [2]

1) Universidade do Minho, Portugal
a58333@alunos.uminho.pt

2) Universidade do Minho, Portugal
jlmp@dsi.uminho.pt

Resumo

Com este artigo pretende-se dar um contributo para a clarificação de uma nova área da tecnologia de bases de dados que emergiu, recentemente, como resposta às dificuldades apresentadas pelo tradicional modelo relacional, no que diz respeito ao suporte do chamado *Big Data* – as bases de dados *NoSQL*. Em particular, pretende-se abordar uma das quatro classes de bases de dados em que o mundo *NoSQL* se divide – as *Document-Based databases*, de que produtos como o MongoDB e o *CouchDB* são alguns dos principais representantes. Como propósito do trabalho em que este artigo se enquadra, pretende-se perceber quais as diferenças, a nível de desempenho, entre soluções suportadas por estes dois produtos, quando utilizados no suporte de bases de dados de dimensão crescente, procurando evidenciar também as diferenças entre estas e soluções construídas recorrendo à bem conhecida tecnologia relacional.

Palavras chave: *Big Data*, bases de dados *NoSQL*, *Document-Based Databases*.

1. Introdução

O crescimento exponencial da chamada Web 2.0 resultou na geração de cada vez maiores volumes de dados, caraterizados por uma grande diversidade de formatos, de que é necessário tirar partido. O armazenamento, a gestão e o processamento destes novos dados veio evidenciar as limitações das bases de dados relacionais, tanto em termos de capacidade de lidar com os grandes volumes de dados, como dos tipos de dados a suportar.

Atualmente são gerados enormes volumes de dados a todo o instante e por todo o tipo de dispositivos, desde eletrodomésticos inteligentes, até ao comércio *online*. Existem cada vez mais sistemas ligados à internet que comunicam entre si. Nos últimos anos, os avanços nas tecnologias Web, o incremento na utilização das redes sociais e a proliferação de sensores e dispositivos móveis conectados à internet, resultou na geração de imensos dados que necessitam de ser armazenados e posteriormente processados. Esta situação forçou uma evolução ao nível das tecnologias de bases de dados para responder aos novos desafios:

- Grandes níveis de concorrência na leitura e escrita de dados;

- Para não defraudar as expectativas dos utilizadores, as bases de dados têm de fornecer tempos de resposta suficientemente baixos;
- Armazenamento de grandes volumes de dados (*Big Data*) eficiente e com determinados requisitos de acesso;
- Aplicações muito exigentes em termos de processamento, como motores de pesquisa e plataformas para redes sociais, têm de responder em tempo útil às solicitações de milhões de utilizadores;
- Para além da escalabilidade é fundamental uma alta disponibilidade das soluções;
- Cada vez mais o volume de dados e os pedidos concorrentes aumentam, pelo que as bases de dados necessitam de ser facilmente expandidas sem interrupções;
- A juntar a tudo isto, exigem-se baixos custos de gestão e operação das soluções.

Neste contexto, uma nova tecnologia de bases de dados emergiu e está a tornar-se cada vez mais importante no mercado – as bases de dados *NoSQL* [Han et al. 2011]. Interessa, pois, conhecer esta nova tecnologia e perceber o contributo que pode dar para suportar as necessidades do *Big Data*. Neste contexto será aqui abordada a classe das bases de dados *NoSQL* designadas por *Document-Based*.

2. Big Data

Atualmente o armazenamento e processamento de grandes volumes de dados na área da computação é um grande desafio. O mundo digital está a crescer a olhos vistos e a tornar-se cada vez mais complexo, quer relativamente ao volume de dados a tratar (do *Terabyte* ao *Petabyte*), à sua variedade (dados estruturados, semiestruturados e não-estruturados) e à velocidade a que têm de ser processados (ritmos cada vez mais rápidos).

A este fenómeno chama-se *Big Data*, e pode ser definido como uma grande coleção de dados que cresceu tanto e de forma tão diversificada que não pode ser gerido ou explorado eficientemente com as tradicionais ferramentas da tecnologia relacional [Moniruzzaman & Hossain 2013].

O termo *Big Data* surgiu como referência a um conjunto de problemas derivados das enormes quantidades de dados produzidos por empresas, organizações e pessoas particulares. Posteriormente, passou a descrever "o conjunto de soluções que visam tratar esses dados, desde o seu armazenamento até à transformação destes em informações relevantes para auxiliarem os gestores nas tomadas de decisões" [Basso & Padua 2014].

Segundo Alexandre e Cavique [2013], o fenómeno *Big Data* assenta nos chamados 3V's:

- **Volume**: A necessidade de lidar, de forma sustentada, com o crescimento exponencial de dados gerados;
- **Velocidade**: A necessidade de processar os dados em tempo-real, devido à importância que isso tem para as organizações atuais;
- **Variedade**: O facto de os dados se apresentarem nos mais variados formatos, alguns deles dificilmente acomodáveis em estruturas rígidas.

A quantidade de dados gerada diariamente está-se a tornar abismal, desde redes sociais, web, entre outros. Isto leva a problemas ao nível do processamento, armazenamento e manipulação de dados, que faz com que os sistemas relacionais não sejam os mais indicados para lidar com o problema do *Big Data*. Por essa razão foram então criados os Sistemas *NoSQL*, que dispõem de melhorias ao nível do armazenamento e processamento de grandes quantidade de dados, principalmente híbridos e não estruturados [Vieira et al. 2012].

O *NoSQL* veio resolver o problema da grande quantidade de dados que são gerados e têm de ser tratados. Para além disso deve levar em consideração a necessidade das organizações manterem esses dados durante longos períodos de tempo [Alexandre & Cavique 2013].

3. Surgimento do conceito *NoSQL*

Para entender melhor este novo conceito é importante perceber em que consiste uma base de dados. Esta pode ser definida, muito simplesmente, como um repositório onde se armazenam dados de vários tipos, interligados através de algum mecanismo que permita manter as referências entre si.

E.F.Codd propôs o modelo relacional de base de dados em 1970. Na altura, este veio simplificar o armazenamento de dados e o seu processamento, ao migrar os dados para tabelas relacionadas entre si por campos comuns, manipulados por uma linguagem de alto-nível – a SQL. O que autor não previu foi que a quantidade de dados que se teria de armazenar no futuro se tornaria tão gigantesca como é hoje [Cardoso 2012].

Em relação ao conceito de *NoSQL*, este foi inicialmente usado em 1998 por Carlo Strozzi para referir uma base de dados *open source* que não utilizava a referida interface SQL [Abramova et al. 2014; Strauch 2010].

Pozzani [2013] refere que esta tecnologia foi desenvolvida inicialmente na Google, sendo que o primeiro *research paper* foi publicado em 2003. Uma dos primeiros movimentos de mudança em direção ao *NoSQL* ocorreu em 2007, quando a *Amazon* apresentou o *NoSQL*

Dynamo. Esta foi das primeiras empresas a guardar os seus dados numa base dados de características não-relacionais [Leavitt 2010]. O exemplo do sistema *Dynamo* veio depois inspirar vários outros projetos com um armazenamento de dados alternativo ao modelo relacional.

Strauch [2010] acrescenta que, com o aparecimento da *web2.0*, empresas famosas como *Google*, *Amazon*, entre outras, começaram a largar as tradicionais bases de dados relacionais como o *Oracle* e o *DB2* para começarem a construir os seus próprios sistemas, isto para armazenar e processar grandes quantidades de dados.

Tomando como exemplo uma das maiores redes sociais do momento, o *Facebook*, só nesta rede 2.4 biliões de conteúdos são partilhados entre amigos todos os dias [Grolinger et al. 2013; Parikh 2013]. Neste tipo de contextos a tecnologia relacional começou a apresentar limitações especialmente relacionadas com a escalabilidade e o desempenho.

As novas exigências fizeram com que a indústria das TIC surgisse com novas propostas para estes dois problemas, resultantes das limitações dos sistemas de armazenamento relacional. Empresas como o *Facebook* e a *Google* concluíram que seria impossível gerir grandes quantidades de dados com as bases de dados relacionais e começaram então a procurar soluções alternativas aos sistemas relacionais. A alternativa que surgiu denominou-se *NoSQL* [Cardoso 2012; Guimarães et al. 2013; Hecht & Jablonski 2011]. A tecnologia *NoSQL* não foi criada com o intuito de substituir o modelo relacional mas sim preencher uma necessidade do mercado que este não servia adequadamente.

Diferentemente das bases de dados relacionais que, de uma forma ou outra, são todas relativamente semelhantes, a área das bases de dados *NoSQL* caracteriza-se por possuir diferentes tipos que variam muito entre si, tendo cada uma as suas próprias caraterísticas e finalidades [Vieira et al. 2012].

4. NoSQL

NoSQL é um termo utilizado para descrever uma classe de tecnologias que fornecem uma alternativa para o armazenamento de dados em comparação com os sistemas de gestão de bases de dados relacionais. Caracterizam-se por fornecer soluções efetivas e de baixo custo para os problemas de disponibilidade e escalabilidade [Schram & Anderson 2012].

Inicialmente, o termo NoSQL era suposto significar "No SQL", mas devido a poucos sistemas eliminarem completamente a tecnologia relacional, foi então atualizado para "Not only SQL" [Schram & Anderson 2012]. Esta tecnologia é uma alternativa viável relativamente às bases de

dados relacionais, pois garante melhor desempenho nas consultas enquanto mantêm um certo nível de escalabilidade e flexibilidade [Lee et al. 2013].

Mesmo que não exista um consenso no que exatamente constitui uma solução NoSQL, o seguinte conjunto de caraterísticas é normalmente atribuído a estas [Grolinger et al. 2013]:

- São completamente livres de esquemas, de modo a poderem lidar com uma grande variedade de tipos de dados;
- Possibilitam escalar horizontalmente as soluções através da adição incremental de novos nodos (commodity servers);
- Focadas na disponibilidade e na tolerância ao particionamento. Sendo assim, estão apontadas para cenários de alta distribuição. Portanto, escolhem comprometer a consistência em favor da disponibilidade;
- Normalmente, não suportam transações ACID, tal como são fornecidas pelas bases de dados relacionais. Neste âmbito as bases de dados NoSQL são normalmente referidas como sistemas BASE [Pritchett, 2008]. No entanto, algumas bases de dados NoSQL, como o CouchDB, suportam transações ACID.

Este tipo de tecnologia começa a encontrar-se em operação em grandes centros de dados, onde as falhas são uma ocorrência constante, tanto devido a problemas de disco, como comunicação de rede. Esta tecnologia torna-se apropriada nesses contextos já que possui recursos para permitir a replicação de dados em várias máquinas, o que se traduz num aumento da disponibilidade dos sistemas [Queiroz et al. 2013].

5. ACID VS BASE

Enquanto o ACID fornece um conjunto de propriedades que garantem que as transações das bases de dados são processadas com confiança, o oposto modelo BASE, que é derivado diretamente do teorema CAP, aponta para fornecer um diferente conjunto de propriedades das que o ACID fornece [Silva 2011].

O acrónimo ACID significa, *Atomicity*, *Consistency*, *Isolation* e *Durability*, e é usado pela generalidade dos sistemas de gestão de base de dados relacionais para assegurar a integridade das bases de dados. As suas caraterísticas constituintes significam o seguinte:

- ***Atomicity*** – Caso exista uma falha na transação, o efeito das suas operações não é aplicado na base de dados;
- ***Consistency*** – A integridade da base de dados é garantida no final da execução de

uma transação;

- ***Isolation*** – O resultado de uma execução de uma transação concorrente é o mesmo do que uma transação isolada;
- ***Durability*** – Após a confirmação de uma transação, mesmo que ocorram falhas os resultados são refletidos na base de dados.

ACID fornece um conjunto de propriedades que garantem que as transações das bases de dados são processadas com confiança e que a base de dados é consistente em acessos concorrentes e falhas do sistema [Silva 2011].

Em relação ao modelo BASE, Toth [2011] refere que para que seja possível trabalhar com um sistema distribuído foi proposto enfraquecer ou relaxar o requisito de consistência permanente de dados, focando antes na disponibilidade e na tolerância ao particionamento. Deste modo surge o conceito BASE *(Basically Available, Soft state, Eventual consistency)* com propriedades e filosofia oposta.

Toth [2011] refere ainda que é considerado um cenário que lida bem com transações distribuídas, tolerância a falhas de consistência, e replicação otimista num sistema distribuído.

Vieira et al. [2012] assim como Toth [2011], referem que o paradigma BASE foi criado a partir do teorema CAP. O primeiro autor ainda refere que o facto de não se controlar a consistência, faz com que exista uma sensível diminuição no custo computacional para a garantia de consistência dos dados em relação às bases de dados relacionais. Já o segundo autor refere que este modelo de consistência está normalmente associado a sistemas distribuídos. Refere também, que todas as atualizações pendentes serão propagadas por todos os nós do sistema, caso durante um determinado tempo não existam atualizações, ficando assim no fim, todo o sistema consistente. A partir deste conceito vários tipos de bases de dados *NoSQL* foram criadas que, como já foi referido, não garantem temporariamente a consistência dos dados de uma aplicação. Toth [2011] refere grandes empresas, como *Google* e a *Amazon*, que aproveitaram a vantagem deste paradigma, isto porque utilizaram um modelo *NoSQL*, aplicado ao conceito de distribuição horizontal. Como não existe garantia de consistência, as configurações de fragmentação no modelo são mais fáceis de efetuar e, por conseguinte, o modelo torna-se mais escalável e custo reduzido.

Resumindo, *Basically Available* significa que as bases de dados estão disponíveis sempre que são acedidas, mesmo que uma parte esteja indisponível. *Soft state* diz que não precisam sempre de ser consistentes e podem tolerar inconsistência por um certo período de tempo; e *Eventual*

consistency enfatiza que, depois de um certo período de tempo, a base de dados vem para um estado de consistência.

6. Teorema CAP

O CAP (*Consistency, Availability, Partition tolerance*) foi introduzido em 2000 pelo cientista de computação Eric Brewer [2000] como uma conjetura e foi formalmente provada em 2002 por Seth Gilbert e Nancy Lynch [2002], sendo assim, estabelecida como teorema [Cardoso 2012; Silva 2011].

Toth [2011] refere que o teorema CAP, enquadrando-se na área da computação distribuída, assenta em três requisitos que descreve da seguinte forma:

- **Consistência**: Todos os nós de um sistema distribuído necessitam que a versão dos dados que armazenam seja a mesma;
- **Disponibilidade**: Ainda que um nó da rede esteja desligado, todos os clientes podem encontrar, sempre que pretenderem, pelo menos uma cópia dos dados que necessitam;
- **Tolerância ao particionamento**: Caso o sistema seja implementado em servidores diferentes, continua com os seus dados, propriedades e caraterísticas, sempre transparentes para o utilizador.

De acordo com este teorema, num sistema distribuído torna-se impossível garantir, simultaneamente, mais que dois daqueles três requisitos. Na figura seguinte estão representadas as combinações possíveis.

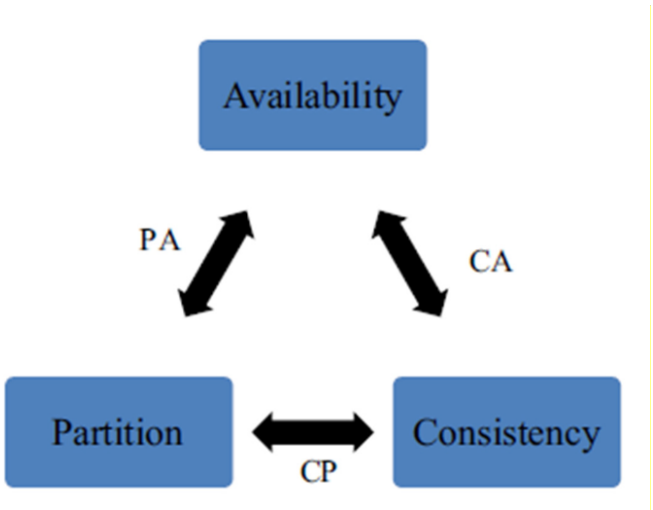


Figura 1 - Teorema CAP [Indrawan-Santiago 2012, pag.47]

Segundo Dimitrov [2010], as combinações possíveis dos três requisitos processa-se da seguinte forma:

- CA - Consistência e Disponibilidade
  - *Clusters* únicos, fácil de assegurar devido a todos os nodos estarem sempre em contacto;
  - Exemplo: 2 PC;
  - Quando um particionamento ocorre o sistema bloqueia.
- CP - Consistência e Tolerância ao particionamento
  - Alguns dados podem ficar inacessíveis (disponibilidade sacrificada), mas o resto é consistente;
  - Exemplo: *sharded database*.
- PA - Disponibilidade e Tolerância ao particionamento
  - O sistema está disponível mesmo durante o particionamento, mas alguns dados podem-se revelar inconsistentes;
  - Exemplo: *DNS*, *caches*, *Master*/*Slave replication*;
  - Necessita de alguma estratégia de resolução de conflitos.

Tipicamente as bases de dados *NoSQL* são PA (*Partition tolerance*/*Availability*) enquanto a maioria das relacionais são CA (Consistency/Availability).

7. Tipos de bases de dados *NoSQL*

Segundo diversos autores (Toth [2011], Bonnet et al. [2011] ou Guimarães et al. [2013], entre outros), as bases de dados *NoSQL* dividem-se em quatro tipos principais:

- *Document-based databases*, que têm como finalidade armazenar documentos em coleções, coleções estas que contêm objetos idênticos. Não possuem esquema desde que o número de campos não seja limitado e pode ser dinamicamente adicionada a um documento. (por exemplo, o *CouchDB* e o MongoDB);
- *Column-based Databases*, que têm como finalidade armazenar os dados como

triplas contendo linha, coluna e rótulo de tempo (por exemplo, o Cassandra e o HBase);

- *Key/Value-based Databases*, que têm como finalidade armazenar os dados como chaves e valores (por exemplo, o Redis e o DynamoDB);
- *Graph-based Databases*, que têm como finalidade armazenar relacionamentos entre os dados (por exemplo, o Neo4J e o *BigData*).

Neste trabalho interessam-nos particularmente as *Document-based Databases*, pelo que serão sucintamente apresentadas de seguida.

8. Document-based databases

Estas bases de dados são organizadas em coleções de documentos, ao contrário das bases de dados relacionais em que se usam tabelas estruturadas com campos para cada registo. Também é permitido ao utilizador adicionar vários campos de diferentes tamanhos a um documento [Leavitt 2010].

Nestas bases de dados o objetivo é gerir e armazenar documentos. Estes documentos podem assumir os formatos *XML, JSON (Javascript Object Notation) ou BSON (Binary JSON)*, que são formatos padrão na troca de dados [Moniruzzaman & Hossain 2013].

Uma vez que estas bases de dados não possuem um esquema definido para lidar com os dados faz com que sejam uma boa solução para as típicas aplicações da *web 2.0*, assim como para gerir conteúdos associados às redes sociais [Queiroz et al. 2013].

As *Document-based databases* agrupam pares de chave-valor em documentos. A cada documento é atribuída uma chave especial "ID", sendo que esta é atribuída a cada um destes separadamente. Desta forma, é possível tratar dados estruturados complexos como *Nested objects* mais convenientemente [Hecht & Jablonski 2011].

Dentro de uma coleção, os documentos podem ser vistos de forma sequencial, podem ser recuperados individualmente utilizando o seu identificador, ou efetuar uma consulta respeitando a sua estrutura de campos. É igualmente possível introduzir um novo documento a qualquer altura numa coleção [Atzeni et al. 2013].

O facto de as relações entre os dados estarem integradas dentro do próprio documento pode originar dados duplicados, no entanto possibilita uma consulta direta. Desta forma, tendo como base uma estrutura que tenha uma relação Cliente com Encomendas (um cliente pode ter várias encomendas), se uma consulta fosse criada para retornar as encomendas do Cliente "X", neste

tipo de base de dados bastaria obter o documento correspondente ao Cliente "X", que este por sua vez já traria os documentos correspondentes às suas encomendas. Já nas bases de dados relacionais seria necessário fazer uma junção das tabelas Clientes e Encomendas, o que levaria a uma diminuição de desempenho na sua execução, mais sentida caso as tabelas tenham alguma dimensão [Almeida & Brito 2012]. Uma outra vantagem de possuir dados duplicados é que "facilita a distribuição do sistema, já que a quantidade de nós a serem consultados numa busca envolvendo várias entidades relacionadas é menor caso elas estejam próximas" [Diana & Gerosa 2010].

Relativamente aos documentos abaixo apresentados pode-se visualizar que eles são similares mas podem ter diferenças em termos de estrutura e conteúdo. Isto é permitido nas *Document-based databases*. Ainda que a estrutura de dados possa diferir entre documentos, estes podem pertencer à mesma coleção, ao contrário das bases de dados relacionais em que cada linha de uma tabela tem de seguir a mesma estrutura [Sadalage & Fowler 2012].

```
{
    nome: "Gonçalo",
    idade: 24,
    morada: "Azurém, Guimarães",
    Gostos: ["Ginásio", "Futebol"]
    Data_Nascimento: "04-06-1991",
    Emprego: "Estudante",
    Estado Civil: "Solteiro",

}
```

```
{
    nome: "Luís",
    idade: 45,
    morada: {Freguesia: "Azurém",
             Cidade: "Guimarães",
             CódigoPostal: 4800},
    Gostos: ["Bicleta", "Futebol"]
    Data_Nascimento: "03-06-1970",
    Emprego: "Engenheiro Informático",
    Estado Civil: "Casado",

}
```

Figura 2 - Exemplo de dois documentos de uma mesma coleção

As bases de dados que serão alvo de atenção neste trabalho são o MongoDB e o CouchDB, já que são as mais representativas no mercado e consideradas como sendo as mais completas da família de bases de dados *NoSQL* do tipo *Document-based databases*.

8.1 MongoDB

O MongoDB é uma *Document-based database*, *open-source*, escrita na linguagem C++. A companhia 10gen também fornece suporte a esta base de dados. Não possui qualquer tipo de esquema e gere documentos JSON como o CouchDB. Foca-se no alto desempenho e desenvolvimento ágil, fornecendo ao programador um conjunto de caraterísticas para facilmente consultar dados, ao mesmo tempo que facilita a escalabilidade do sistema.

Em relação ao modelo de dados, o MongoDB utiliza uma base de dados composta por coleções

de documentos no formato BSON, que é uma versão muito semelhante ao JSON mas numa representação binária, por razões de eficiência e por disponibilizar tipos de dados adicionais em comparação com o JSON" [Strauch 2010]. Segundo Queiroz et al. [2013], estas coleções de documentos relembram os conceitos de tabelas e linhas da tecnologia relacional, mas com a particularidade dos documentos não necessitarem de ter o mesmo esquema. Sendo assim, os documentos podem conter campos que são diferentes de outros documentos (ver Figura 3).

```
{
      "_id" : "001",
      "tipo": "hospital",
      "nome": "Santa Casa",
"endereco" : {
        "logradouro": "R. das Rosas",
            "numero": 95,
            "cidade": "Ouro Preto" }
},
{
         "_id": "999",
        "tipo": "foco_queimada",
   "long_lat": [-46.76, -20.54],
   "satelite": "GOES-12",
 "observacao": "2012-07-05 05:15:00"
}
```

Figura 3 - Exemplo de documentos diferentes de uma mesma coleção [Queiroz et al. 2013, pág.485]

Cada documento possui um identificador único (_id), que pode ser fornecido pelo utilizador quando um documento é criado, ou automaticamente gerado pela base de dados. No campo _id é automaticamente criado um índice, embora outros índices possam ser criados manualmente com o objetivo de melhorar o desempenho das consultas.

Em relação ao modelo de consultas o MongoDB suporta consultas dinâmicas (ad hoc) sobre todos os documentos existentes numa coleção, em analogia com as consultas SQL disponibilizadas pelos sistemas relacionais. Segundo Silva [2011], "*as consultas são expressadas como objetos JSON e são enviadas para o MongoDB por database drivers (normalmente usando o método "find").*" Por exemplo, na figura seguinte apresenta-se um exemplo simples de uma consulta a uma base de dados de todos os livros contendo o título "Flatland".

```
1 db.books.find({'title': 'Flatland'})
```

Figura 4 - Consulta de todos os livros com o título Flatland [Silva 2011, pág.17]

Como se pode constatar na Figura 5, também é possível fazer a ordenação e a contagem de resultados. Neste caso, na primeira consulta todos os livros que estão escritos em inglês são recuperados, ordenando-os pelo seu título. Já a segunda consulta fornece o número de livros escritos em inglês. Do mesmo modo, estão também disponíveis operadores de agregação ($max, $min, $avg, $sum, etc.), que é possível fazer atuar sobre grupos de documentos.

```
1 db.books.find({'language': 'English'}).sort({'title':1})
2 db.books.find({'language': 'English'}).count()
```

Figura 5 - Exemplo de duas consultas [Silva 2011, pág.17]

Em relação ao modelo de replicação o MongoDB fornece replicação *master-slave* e *replica sets*, sendo que os dados são replicados assincronamente entre servidores. Em ambos os casos, para operações de escrita, apenas um servidor é usado, enquanto as operações de leitura podem ser redirecionadas para os servidores *slave*. Os dados são replicados através dos servidores do conjunto.

Em relação ao modelo de consistência o facto de o MongoDB possuir apenas um *master* ativo, a consistência total é possível se todas as operações de leitura forem realizadas no master. Se a replicação para os servidores *slave* for feita assincronamente e o MongoDB não necessitar de fornecer controlo de concorrência das versões, a leitura nos servidores *slave* será de *eventual consistency*. O facto de poder existir leitura nos servidores *slave* possibilita o balanceamento de carga e, sendo assim, o cliente também pode fazer com que a escrita seja replicada para servidores *slave*. Isto ajuda a lidar com escritas importantes, em que a *eventual consistency* possa não ser adequada, porque assim consegue-se flexibilidade de leitura para ler através dos servidores *slave*.

8.2 *CouchDB*

A base de dados Apache CouchDB é também uma *Document-based database open-source*. Foi criada pela Fundação Apache e desenvolvida em Erlang - uma linguagem de programação utilizada para aplicações concorrentes e distribuídas. As propriedades ACID são cumpridas nesta base de dados, fornecendo atualizações serializadas. É já, pela sua natureza, distribuída e suporta vários tipos de esquemas de replicação e resolução de conflitos.

Em relação ao seu modelo de dados uma base de dados CouchDB é formada por uma coleção independente de documentos em formato JSON [Queiroz et al. 2013]. Estes documentos são

compostos por pares de chave:valor. Os valores podem ser de vários tipos: números, strings, boleanos, listas e dicionários. Os documentos não possuem qualquer tipo de esquema e, sendo assim, não têm de seguir nenhuma estrutura, bastando respeitar a sintaxe do JSON.

A cada documento são atribuídos pelo menos dois atributos: um número de revisão (_rev), necessário para a resolução de conflitos durante atualizações, e um identificador único (_id). Na Figura 6 está representado um exemplo de um livro nesta linguagem.

```
{
 "_id":"282e1890a8095bcc0cb1318bed85bcb377e2700f",
 "_rev":"946B7D1C",
 "Type":"Book"
 "Title":"Flatland",
 "Author":"Edwin A. Abbot",
 "Date":"2009-10-09",
 "Language":"English",
 "ISBN":"1449548660",
 "Tags":["flatland", "mathematics", "geometry"],
}
```

Figura 6 - Documento com a descrição de um livro [Silva 2011, pág.14]

Nas palavras de Silva [2011], esta base de dados "*é um simples contentor de uma coleção de documentos e não estabelece nenhuma relação obrigatória entre eles*." (Anderson et al. 2010)
Embora o CouchDB não imponha nenhum tipo de relação entre os documentos, é importante possuir algum tipo de estrutura que retrate como os documentos são organizados. Sendo assim, o CouchDB fornece um *view model*.
Segundo Silva [2011], *views* são métodos de agregar documentos existentes na base de dados. Já Queiroz et al. [2013] refere que estas permitem a filtragem de partes específicas ou criar índices para acelerar a procura de um documento. As *views* não afetam os dados de uma base de dados, simplesmente mudam a maneira como os dados são representados, definindo assim o design da aplicação. Silva [2011] refere também que "*para definir uma view o utilizador deve fornecer uma função javascript que atua como uma função map da operação Mapreduce. Esta função leva um documento como um argumento e é permitido manipular os dados do documento da maneira que se desejar. A view resultante é um documento que pode ter múltiplas linhas*."
Na figura anterior (Figura 6) foi apresentado um extrato de uma coleção de documentos de livros. Partindo do princípio que é pretendido obter o título de todos os livros que estão armazenados na base de dados, criamos uma função *map* correspondente para ser usada na *view* como é mostrado na Figura 7.

```
1 function(doc) {
2   if (doc.Type == "Book") {
3     emit(null, {Title: doc.Title};
4   }
5 }
```

Figura 7 - Exemplo de uma função utilizada para obter o título de todos os livros armazenados na base de dados [Silva 2011, pág.14]

Em relação ao modelo de consultas o CouchDB utiliza uma RESTful HTTP API para efetuar operações CRUD básicas em todos os dados armazenados e utiliza métodos, como os HTTP POST, GET, PUT e DELETE para as realizar. Consultas mais complexas podem ser traduzidas através de *views*, como se viu anteriormente, e o resultado pode ser lido utilizando esta *REST API* [Anderson et al. 2010; Silva 2011].

Relativamente ao modelo de replicação o CouchDB é um sistema de base de dados distribuído do tipo *peer-based*, permitindo *peers* para atualização e alteração de dados, sincronizando bidireccionalmente as mudanças. Portanto, utilizando o *master-slave setup*, em que as sincronizações são feitas unilateralmente, ou o *master-master setup*, em que as mudanças podem acontecer em qualquer um dos nodos, estes devem ser sincronizados bidireccionalmente.

Para isto acontecer, o CouchDB utiliza uma replicação otimista e tem a capacidade de lidar com conflitos, que podem acontecer devido a replicações de alterações. Como já foi referido anteriormente, um número de revisão (_rev) é atribuído a cada documento, isto porque sempre que um documento é atualizado, a versão antiga é mantida e atribuído à versão atualizada um número de revisão diferente. Portanto, quando um conflito é detetado, a versão que prevalecer é gravada como a mais recente e a versão que não prevalecer é também gravada no histórico do documento. Isto é realizado de forma consistente em todos os nós do sistema. Existe também a possibilidade da aplicação escolher lidar com o conflito por ela própria, ignorando uma versão [Anderson et al. 2010].

Em relação à consistência, o CouchDB fornece *eventual consistency*. Como vários *masters* são permitidos, as alterações têm a necessidade de serem propagadas aos nodos restantes. Deste modo não existem bloqueios de escrita. Portanto, até estas alterações se propagarem de nodo para nodo, a base de dados permanece num estado de inconsistência. Como são suportados *master setups* únicos com múltiplos *slaves*, a consistência total (*strong consistency*) pode ser alcançada (Anderson et al. 2010).

9. Conclusões e trabalho futuro

Durante este trabalho, a revisão de literatura realizada permitiu clarificar alguns aspetos relativos às bases de dados NoSQL. Nomeadamente, as razões da sua entrada fulgurante no mercado e o papel que o fenómeno *Big Data* teve no impulsionamento destas tecnologias.

Exploradas as características principais das duas *Document-based databases* selecionadas é já claro que, embora pertencentes à mesma família NoSQL, o MongoDB e o CouchDB, distinguem-se uma da outra em alguns aspetos importantes. Nomeadamente, em algumas particularidades do modelo de dados, no modo como efetuam a replicação, como fazem as consultas e como implementam a consistência.

Como trabalho futuro pretende-se comparar os dois produtos, em termos do desempenho que é possível obter, relativamente ao suporte de bases de dados com dimensões sucessivamente maiores. Este estudo, para cada um dos volumes de dados em análise, envolverá a definição de algumas consultas de complexidade diversa, verificando-se como as duas bases de dados se comportam. Adicionalmente, pretende-se também comparar este tipo de soluções NoSQL com as soluções que lhes corresponderiam usando a tradicional tecnologia relacional, de modo a tirar outras conclusões.

10. Referências

Abramova, V., Bernardino, J., & Pedro, F., "*Experimental Evaluation of NoSQL Databases*", International Journal of Database Management Systems (IJDMS), (2014), 1-16.

Alexandre, J., & Cavique, L., "*NoSQL no suporte à análise de grande volume de dados*", Revista de Ciencias de Computação, 8, (2013), 37 – 48.

Almeida, R. C. De, & Brito, P. F. De., *Utilização da Classe de Banco de Dados NOSQL como Solução para Manipulação de Diversas Estruturas de Dados*, Palmas, 2012.

Anderson, J. C., Lehnardt, J., & Slater, N., *CouchDB: The Definitive Guide*, O. M. Inc., Ed, 2010.

Atzeni, P., Bugiotti, F., & Rossi, L., "Uniform access to *NoSQL* systems", Information Systems, 43, (2013), 117–133.

Basso, C., & Padua, R. de., *O USO DE BIG DATA NAS UNIVERSIDADES*, Seminário

de Iniciação Científica, 94, 2014.

Bonnet, L., Laurent, A., Sala, M., Laurent, B., & Sicard, N., "*Reduce, you say: What NoSQL can do for data aggregation and BI in large repositories*", Proceedings -

International Workshop on Database and Expert Systems Applications, DEXA, (2011), 483–488.

Brewer, E., "*Towards Robust Distributed Systems"*, In Proceedings of the nineteenth annual ACM symposium on Principles of distributed computing, Vol. 19, (2000), 1–12.

Cardoso, R., *Bases de Dados NoSQL*, Instituto Superior de Engenharia do Porto, 2012.

Diana, M. De. Gerosa, M. A., *NoSQL na Web 2.0: Um estudo comparativo de bancos Não-Relacionais para Armazenamento de Dados na Web 2.0*, WTDBD 2010-IX Workshop de Teses E Dissertações Em Banco de Dados, 2010.

Dimitrov, M., *NoSQL Databases*, Ontotext, 2010.

Guimarães, T., Nakai, A., & Vieira, L., *Banco de dados NoSQL para integração de bases de dados de gases de efeito estufa*, MOSTRA DE ESTAGIÁRIOS E BOLSISTAS DA EMBRAPA INFORMÁTICA AGROPECUÁRIA, 9, Brasília, 2013.

Gilbert, S., & Lynch, N., *Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services*, ACM SIGACT News, 33, 51, 2002.

Grolinger, K., Higashino, W. a, Tiwari, A., & Capretz, M. A., *Data management in cloud environments: NoSQL and NewSQL data stores*, Journal of Cloud Computing: Advances, Systems and Applications, 2, 22, (2013).

Han, J., Haihong, E., Le, G., & Du, J., "*Survey on NoSQL database"*, Proceedings International Conference on Pervasive Computing and Applications, ICPCA, (2011), 363– 366.

Hecht, R., & Jablonski, S., "*NoSQL evaluation: A use case oriented survey"*, International Conference on Cloud and Service Computing, (2011), 336–341.

Leavitt, N., "*Will NoSQL Databases Live Up to Their Promise?"* Computer, 43, 2 (2010), 12–14.

Lee, K. K. Y., Tang, W. C., & Choi, K. S., "*Alternatives to relational database: Comparison of NoSQL and XML approaches for clinical data storage"*, Computer Methods and Programs in Biomedicine, 110, 1 (2013), 99–109.

Moniruzzaman, A., & Hossain, S., "*NoSQL database: New era of databases for big data analytics-classification, characteristics and comparison"*, arXiv Preprint arXiv: 1307.0191, 6, 4 (2013), 1–14.

Parikh, J. (2013). *Facebook Newsroom : A new Datacenter for Iowa*, http://newsroom.fb.com/ (11 de Fevereiro de 2015), 2013.

Pozzani, G., *Introduction to NOSQL*, *NoSQL* Seminar, 2013.

Queiroz, G. R. De, Miguel, A., Monteiro, V., & Câmara, G., "Geographic Databases and *NoSQL* : Accomplishments and Future Directions", Revista Brasileira de Cartografia, (2013), 479 – 492.

Sadalage, P. J., & Fowler, M., *NoSQL Distilled*, 2012.

Schram, A., & Anderson, K. M., "*MySQL to NoSQL"*, Proceedings of the 3rd Annual Conference on Systems, Programming, and Applications: Software for Humanity - SPLASH '12, 191 (2012).

Silva, C. A. R. F. O. Da., *Data modeling with NoSQL: How, when and why,* Faculdade de Engenharia da Universidade do Porto, 2011.

Strauch, C., *NoSQL Databases*, Stuttgart, 2010.

Toth, R. M., *Abordagem NoSQL- Uma real Alternativa*, Sorocaba, 2011.

Vieira, M., FIGUEIREDO, J., Liberatti, G., & Viebrantz, A., *"Bancos de Dados NoSQL: conceitos, ferramentas, linguagens e estudos de casos no contexto de Big Data*." , Simpósio Brasileiro DE Banco de Dados, 1 (2012), 1 – 30.

Ferramentas Gratuitas para Desenvolvimento de Soluções de Business Intelligence

CAPSI/2015

Tiago Alexandre Marques da Silva [1], Jorge Alexandre de Albuquerque Loureiro [2,3]

1) Escola Superior de Tecnologia e Gestão de Viseu,
Instituto Politécnico de Viseu,
Campus Politécnico, 3504-510 Viseu, Portugal
tmarquesdasilva@gmail.com

2) Escola Superior de Tecnologia e Gestão de Viseu,
Instituto Politécnico de Viseu,
Campus Politécnico, 3504-510 Viseu, Portugal
3) Centro de Estudos em Educação, Tecnologias e Saúde (CI&DETS/IPV)
Campus Politécnico, 3504-510 Viseu
jloureiro@estv.ipv.pt

Resumo

As organizações guardam grandes quantidades de dados, seja essa informação relacionada directamente com a(s) área(s) de negócio de actividade ou apenas para gestão documental. A capacidade de extrair, transformar e carregar os dados de maneira a identificar oportunidades de negócio é um processo conhecido com BI (*Business Intelligence)*. Informação por si só tem pouco significado, mas a capacidade de extrair informação relacionada e relevante para o negócio poderá ser crucial para a sobrevivência da organização. Informação relacionada e com aspectos de negócio facilmente identificados têm grande importância, mas a BI vai mais além, identificando padrões, por vezes, completamente surpreendentes, que se podem traduzir em novas oportunidades de negócio trazendo, de forma directa ou indirecta, lucros à organização. O desenvolvimento de aplicações de BI pode ser feito com diversas ferramentas disponíveis no mercado, desde ferramentas pagas, gratuitas para uso comercial e *open source*. Este trabalho apresenta várias ferramentas gratuitas e/ou *open source* que disponibilizam várias funcionalidades para desenvolver soluções de BI sem custos tecnológicos. Serão listadas as funcionalidades que disponibilizam., será feita uma análise comparativa das mesmas e serão ainda apresentados casos de uso das ferramentas mais promissoras.

Palavras chave: *Business Intelligence*, *Data Warehousing*, ETL, *Dashboarding*, *Self-Service* BI.

1. Introdução

Business Intelligence [Wingyan 2013] é um conjunto de processos realizados no negócio de uma empresa para torná-la mais eficiente, de modo a que as decisões partam de informação que é analisada atentamente, em vez de serem tomadas por intuição ou pela percepção do que

acontece. Antigamente, a BI era tratada como sendo um luxo dentro de certas organizações [Watson e Wixom 2007], e as soluções que ela prometia, muitas vezes, eram vistas com desconfiança. Hoje, com o aumento do volume de dados nas empresas, em paralelo com o panorama da concorrência no mercado comercial (o que gera a necessidade de diferenciação estratégica), a *business intelligence* tornou-se uma necessidade [Negash 2004]. O tempo é outro e as organizações, os processos e os sistemas evoluíram e, como consequência, a BI também evoluiu, perante as exigências da globalização e da economia.

As soluções de BI nascem das necessidades de tomada de decisões por parte das organizações. Usualmente, o primeiro passo consiste na identificação da informação relevante para análise e a partir desse momento, começa a fase de desenvolvimento da solução propriamente dita [Gangadharan e Swami 2004]. Um quesito imperativo neste cenário é a grande diversidade de ferramentas de BI no mercado. Cada ferramenta possui um conjunto de características particulares como preço, suporte, funcionalidade, facilidade de uso e implementação, etc. Uma organização que decida implementar uma solução de BI, enfrentará um grande desafio na escolha da ferramenta a utilizar.

O desenvolvimento de uma solução de BI deve considerar todos os custos tecnológicos envolvidos, desde o arranque do projecto até à sua finalização [Lönnqvist e Pirttimäkib 2006]. A utilização de ferramentas ferramentas gratuitas (para uso comercial) e/ou *open source*, é um modo de redução de custos, mas quando decidimos optar por uma ou mais ferramentas deste tipo, a questão do preço não deve ser a única questão de análise. O tempo de aprendizagem, a facilidade de utilização, o suporte técnico e a importância do código aberto devem ser questões analisadas de forma cuidada, de modo a que a utilização destas ferramentas não seja uma aposta perdida, ou, no pior dos cenários, comprometer todo o projecto. Dada a grande diversidade de ferramentas existentes no mercado, surgiu a necessidade de perceber se a utilização deste tipo de ferramentas é exequível e produz bons resultados nas as organizações que decidam enveredar por uma solução deste tipo. Para tal, foi feito um levantamento, análise comparativa e estudo de dois casos onde se utilizam duas destas ferramentas.

2. Metodologia de Investigação

Foi necessário fazer um estudo do estado da arte das ferramentas de BI existentes no mercado que sejam gratuitas, para isso foi feita uma investigação de maneira a fazer um levantamento das opções disponíveis. Foram feitas pesquisas *online* [Anselmo 1988] de modo a perceber quais as ferramentas mais populares nos últimos anos. Os autores recorreram a opiniões em

blogs, artigos de opinião e dissertações de mestrado [Tereso 2011] e [Pereira 2011] para identificar as soluções mais abordadas e mais usadas para fazer BI.

Após a fase de pesquisa das ferramentas disponíveis, foi preciso fazer uma avaliação experimental das mesmas, de maneira a estudar a sua viabilidade, considerando-se viáveis aquelas que contenham um conjunto de módulos que:

- Possibilitem extrair dados de várias fontes;
- Permitam transformar e carregar os dados para diferentes fontes de dados;
- Permitam fazer *reporting* e *dashboarding*;
- Apresentem um conjunto de funcionalidades integradas (ferramentas do tipo *suite*).

Foram experimentadas diversas ferramentas. No caso de duas ou mais satisfazerem os requisitos anteriores, serão escolhidas aquelas que, na opinião dos autores, apresentarem um tempo de desenvolvimento menor.

3. Como Fazer Business Intelligence

Construção de um Data Warehouse Multidimensional

Com a necessidade da Business Intelligence nas organizações, surgiu também a necessidade da criação de ferramentas e repositórios para o armazenamento e manipulação de dados. O DW (*Data Warehouse*) [Kimball 2004] responde a essas necessidades, sendo um sistema utilizado para armazenar informações relativas às atividades de um negócio de forma consolidada. O DW possibilita a análise de grandes volumes de dados, recolhidos dos sistemas transacionais, e o desenho da base de dados favorece os relatórios, a análise de grandes volumes de dados e a obtenção de informações estratégicas que podem facilitar a tomada de decisão.

Um DW é modelado de uma forma diferente das bases de dados relacionais O modelo multidimensional é o mais utilizado. Este modelo traz novos conceitos à modelação de uma base de dados. A Figura 1 mostra alguns destes novos conceitos, assim como a hierarquização dos mesmos.

A atracção central do modelo dimensional é a sua simplicidade. Como os dados não estão normalizados [Codd 1972], os utilizadores interpretam a base de dados de forma mais simples.

O modelo de dados multidimensional é composto por cubos lógicos, medidas, dimensões, hierarquias, níveis e atributos. A simplicidade do modelo é inerente porque ele define objectos que representam entidades de negócio do mundo real. Os utilizadores devem procurar saber quais as medidas negócio que estão interessados em explorar, que dimensões e atributos tornam

os dados significantes, e como as dimensões do negócio são organizadas em níveis e hierarquias.

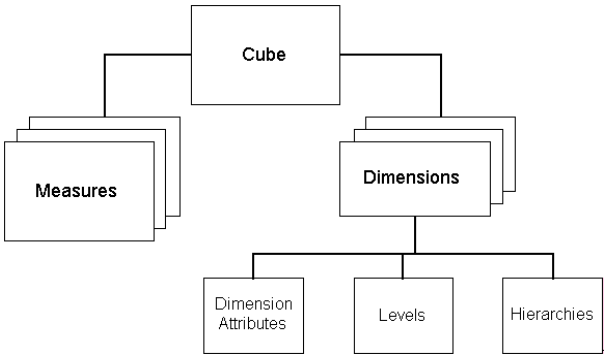


Figura 1 – Composição de um modelo multidimensional (fonte: http://web.stanford.edu/dept/itss/docs/oracle/10g/olap.101/b10333/multimodel.htm)

ETL (*Extract*, *Transform* and *Load*)

Para a construção de um DW é necessário passar por diversas fases. Uma delas é a extracção e processamento de dados. O processo de ETL destina-se à extracção e transformação dos dados e termina com a inclusão destes no DW. Esta fase caracteriza-se por englobar procedimentos de limpeza, integração e transformação de dados. Segundo a literatura, este é o processo mais crítico e demorado na construção de um DW [Bonomo 2009]. A Figura 2 mostra o processo típico de extracção de dados de uma ou mais fontes transaccionais, a sua transformação e por fim o carregamento destes noutros sistemas:

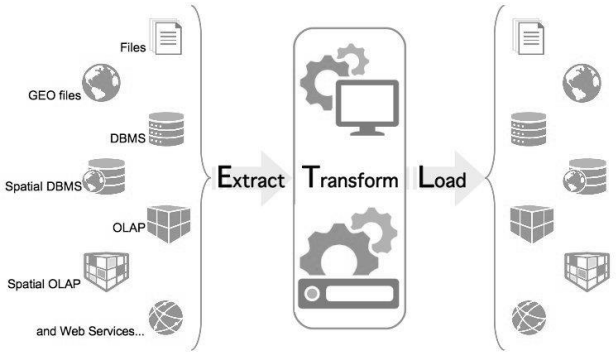


Figura 2 - Processo de ETL (fonte: http://www.dbbest.com/blog/extract-transform-load-etl-technologies-part-1/)

O processo de ETL deve ser pensado e executado com toda a atenção possível, visto que é uma fase do projecto de BI importante, sendo determinante na forma como os dados são inseridos no DW. Estudos relatam que o ETL e as ferramentas de qualidade de dados consomem um terço do orçamento num projecto de BI, podendo, no que respeita ao tempo de desenvolvimento, chegar a consumir 80% do projecto. Outros estudos mencionam, ainda, que o processo de ETL tem custos na ordem dos 55% do tempo total de execução do projecto de DW [Shilakes 2008], [Demarest 1997] e [Inmon 2000].

Construção de um Soluções Analíticas de BI

As aplicações de BI são responsáveis por mostrar informações de negócio aos utilizadores responsáveis por tomar decisões. A forma como estas mostram os dados é determinante para que os utilizadores tenham noção da evolução e do estado de certos indicadores. Estes indicadores são importantes para a tomada de decisões relevantes para o negócio.

Existem diversas maneiras de estruturar uma aplicação de BI. Usualmente recorre-se a um conjunto de *dashboards*, relatórios e consultas OLAP (*Online Analytical Processing*) [Codd 1993]. A evolução destas aplicações, ao longo dos anos, foi tal que é agora o utilizador que pode escolher o que quer ver e de que forma. A este conceito é dado o nome de *Self-Service* BI [Rouse 2012].

A. Dashboarding

Compreender o estado das organizações, do passado até o momento atual, é fundamental para prever tendências e agir rapidamente para se manterem competitivas no mercado. Promovendo uma visão geral de toda a organização ou negócio, o *dashboard* contribui para uma avaliação de cenários bastante completa e dá a oportunidade para que a organização planeie o futuro com maior segurança.

Com *dashboards* a organização fica com uma forma de comunicação mais efectiva, pois elas são baseadas em recursos visuais dinâmicos como gráficos, infográficos, linhas de tendência, entre outros, para ilustrar o desempenho do negócio de forma mais clara e objetiva. A Figura 4 mostra um protótipo de uma *dashboard*:

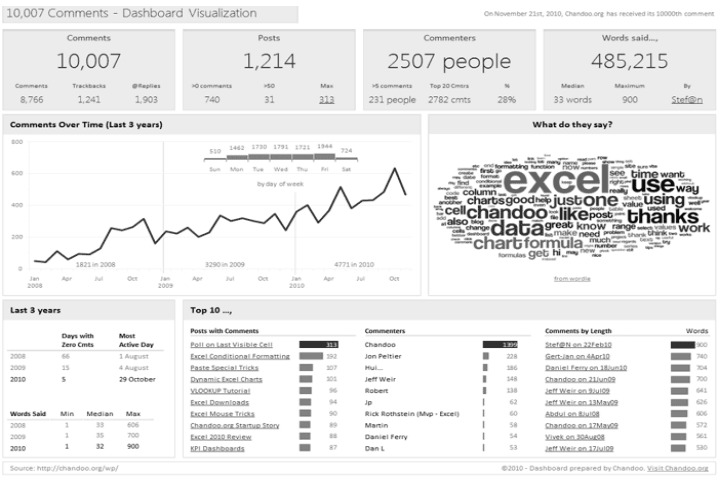


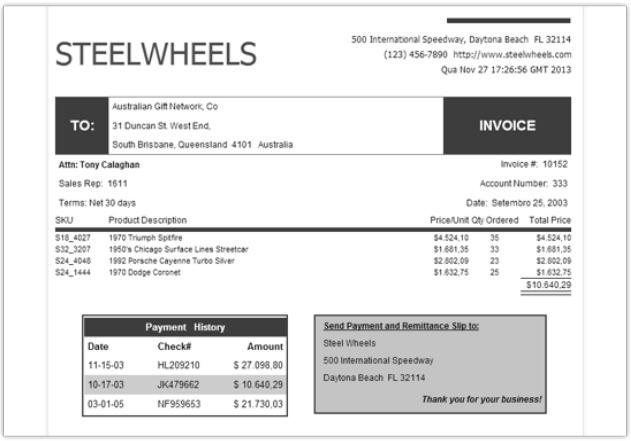
Figura 3 - *Dashboard* executiva (fonte: http://chandoo.org/wp/excel-dashboards/examples/)

Os resultados devem ser acompanhados diariamente, e isto pode ser feito através destas *dashboards* que permitem uma visão global dos indicadores e métricas do negócio. Esta

informação é o que indica se a organização ou negócio está no caminho certo, portanto, deve ser tratada como prioridade, e as *dashboards* podem dar conta disso.

Reporting e Ad-Hoc Reporting

As aplicações de *reporting* justificam-se pela monitorização e pelo acompanhamento das operações da organização numa base de continuidade, muitas vezes, diária. As aplicações de BI terão, neste contexto, de assimilar e transmitir, com grande regularidade, os dados fundamentais (quantidades, valor, alertas para valores de exceção, etc.) ao acompanhamento e análise. O principal factor de sucesso destas aplicações será o seu *modus operandi* com as aplicações de *frontoffice* de registo das actividades. A Figura mostra um relatório com dados de facturação:



STEELWHEELS

500 International Speedway, Daytona Beach FL 32114
(123) 456-7890 http://www.steelwheels.com
Qua Nov 27 17:26:56 GMT 2013

TO: Australian Gift Network, Co
31 Duncan St. West End,
South Brisbane, Queensland 4101 Australia

INVOICE

Attn: Tony Calaghan
Invoice #: 10152
Sales Rep: 1611
Account Number: 333
Terms: Net 30 days
Date: Setembro 25, 2003

SKU	Product Description	Price/Unit	Qty Ordered	Total Price
S18_4027	1970 Triumph Spitfire	$4.524,10	35	$4.524,10
S32_3207	1950's Chicago Surface Lines Streetcar	$1.681,35	33	$1.681,35
S24_4048	1992 Porsche Cayenne Turbo Silver	$2.802,09	23	$2.802,09
S24_1444	1970 Dodge Coronet	$1.632,75	25	$1.632,75
				$10.640,29

Payment History

Date	Check#	Amount
11-15-03	HL209210	$ 27.098,80
10-17-03	JK479662	$ 10.640,29
03-01-05	NF959653	$ 21.730,03

Send Payment and Remittance Slip to:
Steel Wheels
500 International Speedway
Daytona Beach FL 32114
Thank you for your business!

Figura 4 - Relatório com Dados de uma Factura.

Outra forma de fazer *reporting* é sob a forma de *ad-hoc reporting* [Rouse 2010]. *Ad-hoc reporting* é um conceito usado em aplicações que permitem responder a questões específicas de negócio ou obter mais detalhes do mesmo. O utilizador em vez de ter apenas um *report* estático, tem a possibilidade de obter informação filtrada ou navegar sobre o próprio relatório de modo a obter mais detalhes da informação a analisar.

C. OLAP

OLAP é um conceito que se refere a aplicações informáticas que permitem efectuar, de forma rápida e partilhada, a análise de informação multidimensional, originária de diversas fontes de dados. As tecnologias OLAP permitem reestruturar os dados de uma base de dados numa perspectiva multidimensional. Podem-se fazer assim consultas sobre esta estrutura, com vista a detectar tendências e a tirar conclusões. Pesquisas e questões que nas bases de dados tradicionais teriam um tempo de resposta longo, podem ser realizadas em apenas alguns segundos na abordagem OLAP. A Figura 5 mostra a organização de vendas sobre uma

perspectiva multidimensional, em que os produtos, clientes e tempo são cruzados de modo a obter informações relevantes sobre as tendências das vendas:

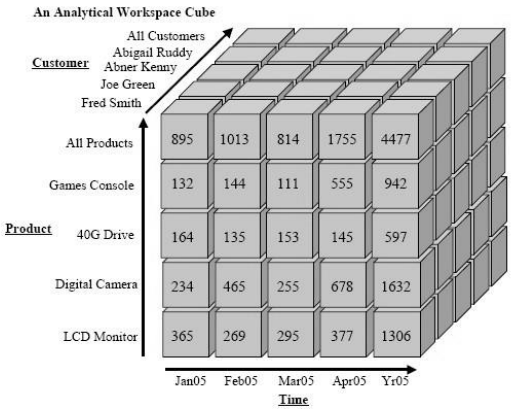


Figura 5 - Consulta OLAP a um Modelo Multidimensional (fonte: http://gerardnico.com/wiki/database/oracle/oracle_olap)

Através de uma funcionalidade denominada *drill*, um utilizador final pode navegar nos dados que interessam à sua análise, aumentando o grau de detalhe (*drill down*) e descendo, consequentemente, a dados cada vez mais próximos dos dados operacionais, ou diminuindo o detalhe (*drill up*), subindo, obtendo-se a uma informação cada vez mais agregada.

D. Self-Service BI

Self-service BI é uma forma de *business intelligence* na qual o utilizador tem a possibilidade de satisfazer, de forma independente, as suas próprias necessidades de informação. Com o *self-service* BI os profissionais não técnicos podem gerar os seus próprios relatórios, executar as suas próprias consultas e dirigir as suas próprias análises sem a assistência do pessoal das TI (Tecnologias de Informação). Há muitos benefícios que podem advir da implementação de sistemas de BI no modo de *self-service*. Os funcionários operacionais podem tomar melhores decisões e de forma mais rápida, porque já não têm que esperar a longa fila de solicitações de relatórios.

Para que os ambientes de BI no modo de *self-service* sejam eficientes, estes devem ser extremamente intuitivos e fáceis de manusear. A maior parte dos utilizadores comerciais ou de negócio não dispõe dos conhecimentos técnicos necessários para trabalhar com ferramentas complexas ou interfaces sofisticadas de BI. Uma aplicação de *self-service* BI apenas será adotada pelo seu público-alvo se lhe oferecer uma forma fácil e simples de aceder à sua própria informação personalizada, sem necessidade de uma longa formação.

4. Ferramentas Gratuitas para o Desenvolvimento de Soluções de BI

Existem no mercado um conjunto alargado de tecnologias e ferramentas que têm como objectivo a BI, ora apenas gerando reports, ou permitindo gerar gráficos para serem embutidos noutras aplicações, fazer consultas OLAP, etc. As ferramentas existentes podem-se dividir em duas grandes categorias, no que toca a custos financeiros para as organizações: ferramentas pagas e ferramentas gratuitas. No que toca a ferramentas pagas existe uma ampla variedade na oferta. As ferramentas mais destacadas são: Information Builders BI Solutions, Oracle Business Intelligence Standard, Microsoft Business Intelligence, SAP Business Intelligence, SAS Business Intelligence e IBM Cognos. O custo das licenças, por utilizador final, das ferramentas pagas, pode chegar a milhares de euros por mês. No entanto, não se pode classificar à partida como sendo um preço elevado para as organizações que as usam, pois o preço até pode ser considerado baixo, dependendo do problema e capacidade financeira das organizações. Fica fora do âmbito deste documento comparar estas ferramentas pagas, no entanto é importante realçar que existem muitas ferramentas deste tipo e o custo das licenças pode chegar aos valores já referidos.

Actualmente existem diversas ferramentas de BI que não têm qualquer custo de licença, pois são gratuitas. Existem ferramentas gratuitas que chegam mesmo a ter o código aberto para que os *developers* possam ver o mesmo e modificá-lo livremente. Estas são consideradas ferramentas *open source*. O facto de uma licença de *software* ser gratuita não significa que a mesma possa ser utilizada para criar um produto vendável. Existem licenças que restringem situações deste tipo. Outras permitem a venda de um produto desenvolvido com elas, desde que a própria ferramenta tenha uma referência no produto final, etc. A importância de análise do tipo de licença [Choose 2015] é outro factor muito importante no momento da escolha de uma ferramenta gratuita de BI.

Serão apresentadas diversas ferramentas gratuitas ao longo desta secção. Para cada uma delas serão listadas as suas funcionalidades principais.

A. Jaspersoft

A ferramenta Jaspersoft é muito conhecida pelas suas funcionalidades de *reporting*, OLAP e ETL. O Jaspersoft começou por ser gratuita, sendo adquirida pela empresa TIBCO. Actualmente, continua a ter uma versão gratuita suportada por uma larga comunidade de utilizadores. No entanto, o seu motor de *reporting* é utilizado na versão paga disponibilizada pela TIBCO. A versão gratuita tem uma média de 40.000 *downloads* semanais, tendo várias *releases* ao longo dos últimos anos. Os componentes que integram esta versão são: o JServer (servidor operacional); o JStudio (ambiente de desenvolvimento para programadores); o

JLibrary (integração de *reporting* embutido em aplicações externas) e o JETL (responsável pelos processos de ETL).

As críticas ao Jaspersoft são muito positivas, desde a facilidade com que se criam relatórios, com o editor JStudio, até à parametrização de diversas variáveis que podem tornar os relatórios bastante ricos na forma como trabalham os dados. A Figura 6 mostra a criação de um relatório com o JStudio, onde se permite observar que existe um considerável leque de componentes para inserir no relatório (à direita da imagem, existem objectos como: tabelas; gráficos; texto; texto estático; imagens; campos de dinâmicos, etc.), assim como um editor de configurações (parte inferior da imagem), expressões para campos dinâmicos (parte esquerda da imagem) e o editor gráfico do relatório a criar (parte central da imagem).

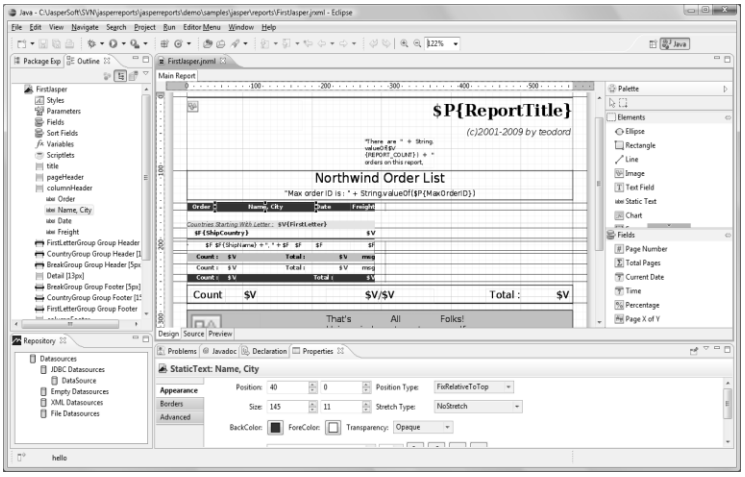


Figura 6 - Criação de um relatório no JStudio (fonte: https://marketplace.eclipse.org/content/jaspersoft-studio)

O JETL é a componente do motor Jaspersoft responsável por realizar os processos de ETL. No entanto, esta componente não foi criado pela própria Jaspersoft. O motor responsável pelo JETL é o Talend, mas esta ferramenta é incluída nos componentes constituintes do Jaspersoft. O Talend será apresentado, neste documento, numa secção mais à frente.

B. BIRT

A ferramenta BIRT é essencialmente conhecida pelas suas funcionalidades de reporting. Tal como o Jaspersoft, o BIRT começou por ser uma ferrrament gratuita, sendo adquirida pela empresa Actuate. Actualmente continua a ter uma versão gratuita suportada por uma larga comunidade de utilizadores. No entanto, à semelhança da ferramenta anterior, o seu motor de *reporting* é utilizado na versão paga disponibilizada pela Actuate. A versão gratuita tem uma média de 284 *downloads* semanais, tendo várias *releases* ao longo dos últimos anos, mas menos que o Jaspersoft. Os componentes que integram esta versão são: o Report Engine (servidor operacional); e o Report Designer (ambiente de desenvolvimento para programadores).

Esta ferramenta apenas faz *reporting*, não inclui qualquer motor para fazer ETL, nem OLAP. Apesar disso, a grande facilidade com que se criam e editam relatórios faz com que seja uma alternativa ao Jaspersoft, que apesar de também ser positiviamente avaliado pela facilidade com que se criam relatórios, é ao mesmo tempo criticado pela edição e importação de relatórios já existentes.

A Figura 7 mostra a criação de um relatório no Report Designer.

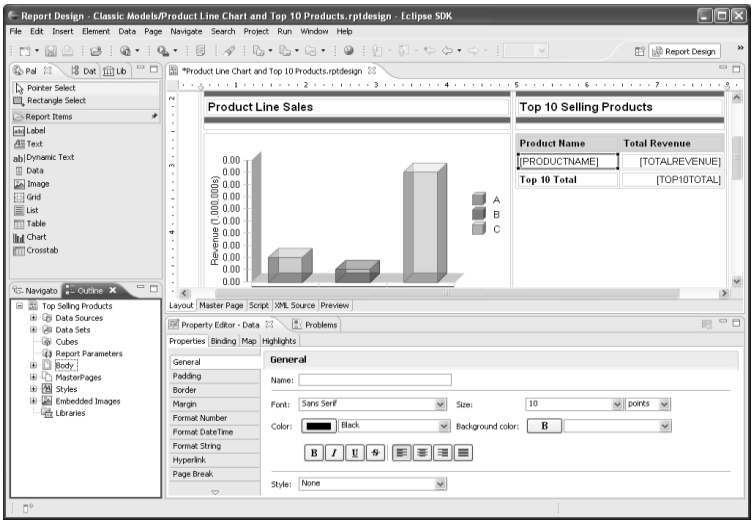


Figura 7 - Criação de um relatório no Report Designer (fonte: https://commons.wikimedia.org/wiki/File:Eclipse_BIRT_Report_Designer.png)

Como se pode constatar, existe um editor gráfico do relatório (parte central da imagem), um conjunto de componentes para adicionar ao relatório (na esquerda da imagem), assim como um editor de configurações (na parte central da imagem).

C. Talend

O Talend é constituído por um conjunto de aplicações para o tratamento de dados, desde o ETL até à Data Quality. Os componentes de ETL e de Data Quality são distribuídos, de forma separada, de duas formas diferentes: uma versão paga e outra gratuita. A versão paga diferencia-se da gratuita, pelo número de funcionalidades disponibilizadas ao utilizador, mas estas podem ser consideradas como extras, já que as funcionalidades base são livremente disponibilizadas nas versões gratuitas. A empresa, com o mesmo nome, diz-se líder em ETL no mercado europeu. Esta ferramenta (componente de ETL) tem um largo conjunto de utilizadores, pois para lá de poder ser utilizada independentemente, esta é incluída noutras ferramentas como por exemplo: o Jaspersoft e o SpagoBI (ferramenta apresentada, neste documento, numa secção mais à frente).

A Figura 8 mostra um exemplo de utilização do Talend Open Studio. É neste editor que se podem criar os processos de ETL. Como se pode verificar pela ilustração, o editor é composto por: editor gráfico (parte central da imagem), editor de configurações (esquerda da imagem), conjunto de componentes responsáveis pelo processo de ETL (parte direita da imagem) e editor de configurações (parte inferior da imagem).

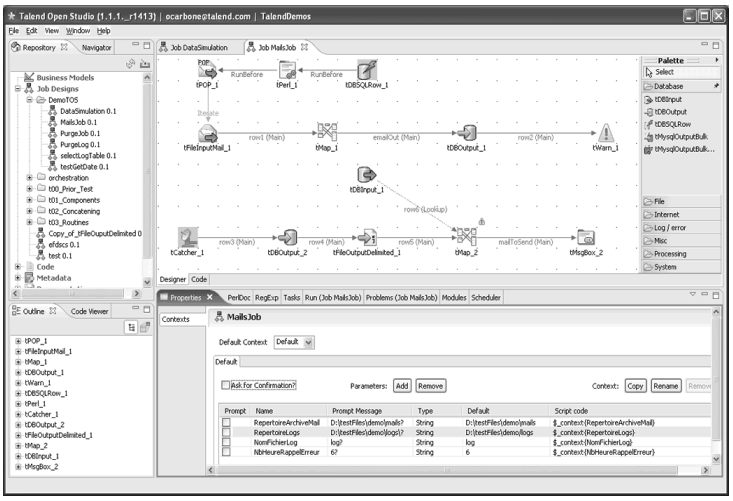


Figura 8 - Criação de um processo de ETL no Talend Open Studio (fonte: http://www.informatiquenews.fr/talend-de-letl-aux-outils-hadoop-32206)

O Talend Open Studio é uma ferramenta que disponibiliza ao utilizador vários componentes, configurações e metadados para criar processos de ETL. A criação de um processo de ETL é feita através dos *jobs*. Um *job* pode ser visto como uma caixa negra que tem entradas e saídas de dados, e no seu corpo são efectuadas operações sobre os dados. Um *job* pode agregar outros *jobs* e com isto um processo de ETL pode ser dividido em diversos *sub-jobs*, cada um tratando de um sub-processo em particular. A criação de jobs é feita através da ligação de vários componentes, uns responsáveis pela extracção de dados, outros pelo seu tratamento e, por fim, os componentes responsáves pela escrita de dados (seja ficheiro ou base de dados).

O processo de ETL pode ser editado de forma gráfica, mas o Talend para correr o *job* usa código Java ou Perl, e enquanto o utilizador está a desenvolver o *job*, a ferramenta está em tempo real a gerar código fonte, e com isto o utilizador pode ver o código, copiá-lo e corrê-lo à parte, num programa isolado. O código fonte é responsável por processar o ETL desenvolvido, e por isso, no final do processo desenvolvido, o utilizador para executar o processo de ETL terá que fazer o *deploy* para um servidor que corra código numa destas linguagens e, por fim, executar o mesmo.

D. Pentaho Data Integration

O Pentaho Data Integration é ferramenta responsável por processos de ETL. Tal como o Talend, a sua distribuição é feita de duas formas diferentes: uma versão paga e outra gratuita. A versão paga diferencia-se da gratuita, tal como o Talend, no número de funcionalidades disponibilizadas ao utilizador. Apesar disto, estas também podem ser consideradas como extras, já que as funcionalidades base são livremente disponibilizadas na versão gratuita. O Pentaho Data Integration é a ferramenta de ETL com mais utilizadores no mercado.

O seu editor gráfico é bastante parecido com o Talend Open Studio, como se pode ver pela Figura 9.

A criação de processos de ETL é muito idêntica ao Talend Open Studio, com a criação de *jobs*, ligação de diversos componentes com funções específicas e com a geração de código fonte que deve ser executado num servidor que corra o mesmo.

O Pentaho Data Integration faz parte da solução completa Pentaho. Esta solução será apresentada, mais à frente, neste documento.

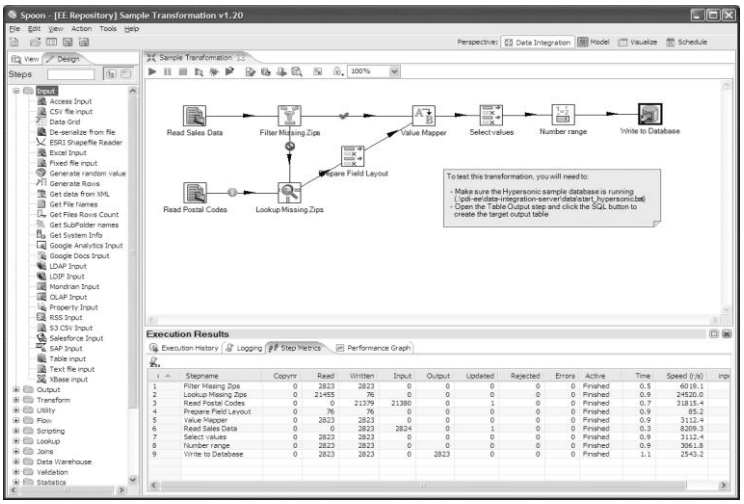


Figura 9 - Criação de um processo de ETL no Pentaho Data Integration (fonte: http://www.matera.com/br/2012/03/20/integracao-de-dados-na-matera-systems/)

Soluções Integradas

As ferramentas apresentadas anteriormente são soluções que podem ser utilizadas de forma independente. Contudo, existem outras que agregam um conjunto de funcionalidades num único *package*, sendo assim soluções integradas. Uma ferramenta integrada tem à partida um conjunto de funcionalidades com fins diferentes, que juntas proporcionam ao utilizador uma solução

completa. As ferramentas, desta categoria, apresentadas de seguida, agregam um conjunto de funcionalidades que permitem: realizar ETL, *reporting*, OLAP, *dashboarding*, *self-service* BI, etc. Tanto a ferramenta SpagoBI como o Pentaho evitam o "síndrome de funcionalidades", isto é, em vez de existir um arquipélago de funcionalidades, estas são agregadas, formando uma ilha de funcionalidades.

A. SpagoBI

O SpagoBI é uma ferramenta disponibilizada gratuitamente e com o código completamente aberto (permitindo a sua modificação e distribuição). O desenvolvimento, manutenção e apoio, ao contrário do Pentaho, não é apenas feito por uma comunidade de utilizadores. A própria empresa (a Engineering Group) é a responsável por toda a solução. A organização que decida adoptar esta ferramenta tem o suporte técnico e os serviços de consultoria garantidos. Apesar de estes serviços serem pagos, dão segurança aos *developers*, pois qualquer problemas e/ou dúvidas podem ser esclarecidas com os técnicos de manutenção do SpagoBI.

A Figura 10 mostra o conjunto de funcionalidades disponibilizadas pela solução SpagoBI. A imagem é a montra, dada pela Engineering Group, da solução SpagoBI. Na figura podem-se ver todas as funcionalidades que a solução oferece, são elas: OLAP, gráficos, KPI's (*Key Performance Indicators*), *dashboarding*, *data mining*, consultas *what-if* (*self-service* BI), ETL, visualização geográfica, *reporting*, entre outras. Todas as funcionalidades anteriores fazem parte da solução SpagoBI.

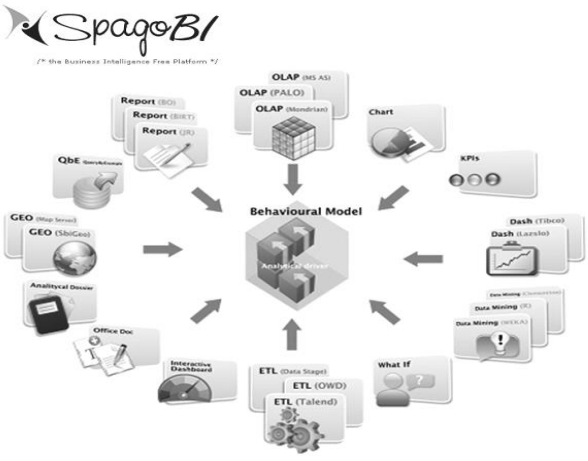


Figura 10 - Conjunto de funcionalidades que formam o SpagoBI (fonte: http://www.stratebi.com/pt/spagobi)

Cada uma das funcionalidades do SpagoBI podem ser utilizadas de forma independente, isto é, a ferramenta permite, por exemplo, a criação de gráficos que podem ser utilizados em

dashboards, criadas também através da ferramenta, ou podem ser integrados noutras aplicações indepententes, desenvolvidas com qualquer tipo de tecnologia (.NET, Java, Python, etc.). O SpagoBI utiliza muito o conceito de documento, isto é, se quisermos criar uma *dashboard* com três gráficos, criamos três documentos, cada um com o gráfico desejado, e por fim, criamos outro documento do tipo *dashboard*, e neste incluímos os primeiros três. Este conceito de documento aplica o conceito de *divide to conquer*, facilitando a resolução de problemas complexos através da resolução de mini-problemas, e tanto ou mais importante, promove a reutilização de documentos/esforço.

Será apresentado, mais à frente deste documento, um caso de estudo onde o SpagoBI foi a tecnologia adoptada pelo ministério da saúde italiano para o desenvolvimento de uma solução de BI.

B. Pentaho

O Pentaho é uma ferramenta disponibilizada de duas formas diferentes. Existe uma versão gratuita e outra paga. A versão gratuita é suportada por uma larga comunidade de utilizadores, sendo esta a principal desvantagem para a versão paga, dado que não existe suporte técnico especializado para as organizações que decidirem adoptar a versão gratuita. No entanto a quantidade de material de apoio e as muitas perguntas/respostas no fórum da comunidade ajudam no esclarecimento de problemas e/ou dúvidas dos *developers*.

A Figura 11 mostra uma *demo* de um portal desenvolvido com a ferramenta Pentaho.

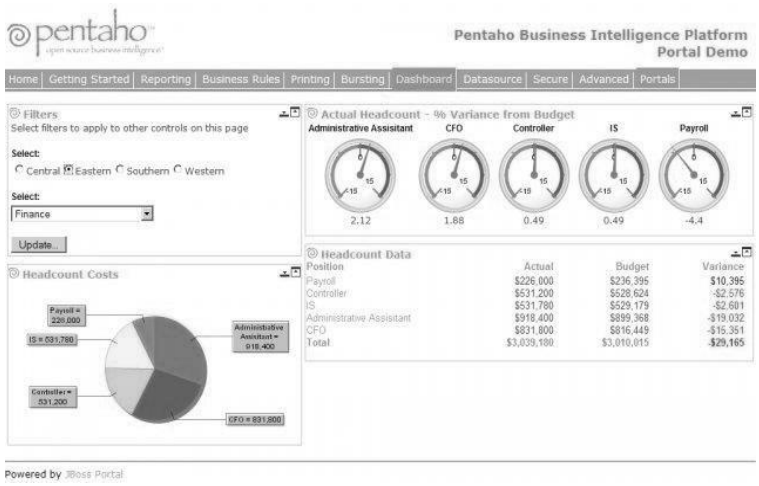


Figura 11 - Exemplo de portal desenvolvido com a ferramenta Pentaho (fonte: http://sourceforge.net/projects/pentaho/)

O conjunto de funcionalidades disponibilizadas pelo Pentaho não são muito diferentes daquelas que o SpagoBI oferece, as principais são: OLAP, gráficos, KPI's, *dashboarding*, *data mining*, ETL, visualização geográfica, *reporting*, entre outras.

A solução Pentaho tem o maior número de utilizadores, comparativamente com outras soluções integradas, isto principalmente por ter sido das primeiras soluções integradas introduzidas no mercado. O Pentaho Data Integration, contribuiu muito para a elevada cota de mercado que o Pentaho alcançou, pois antes da consciencialização, por parte das organizações, da importância da *business intelligence*, já existia a necessidade de extrair, tratar e carregar dados entre fontes de dados diferentes. A capacidade que a comunidade e a própria empresa têm tido para lançar novas versões todos os anos, contribuiu também para a estabilização da posição do Pentaho como líder de mercado. Sendo uma ferramenta poderosa no que toca às funcionalidades que disponibiliza, dada a fama que a mesma possui e a quantidade de ajuda *online* existente, esta apresenta-se como forte candidata à escolha como ferramenta responsável por soluções de BI.

Outras Soluções

Todas as ferramentas apresentadas até ao momento deixam-nos uma garantia: não é por falta de quantidade que as soluções gratuitas não são adoptadas. As soluções anteriores são as que maior destaque têm, de entre as que existem no mercado, no entanto, não são as únicas. Existem muitas outras soluções que disponibilizam uma ou mais funcionalidades que permitem fazer BI, mas dada a quantidade de ferramentas existentes, não foi possível apresentar todas. Fica, no entanto, uma breve referência sobre outras que merecem referência.

A. Rapid Miner

O Rapid Miner tem uma versão gratuita que é responsável por fazer *data mining*. As fontes de dados origem podem ser: ficheiros CSV e Excel.

B. Jedox Palo

A empresa Jedox, disponibiliza uma versão gratuita da sua solução Palo (sendo a principal funcionalidade desta ferramenta a realização de consultas OLAP). O principal destaque do Palo é a sua interface muito parecida com o Excel. O utilizador faz consultas OLAP em páginas muito parecidas com as folhas de cálculo do Excel, podendo manipular directamente os dados, tal como no Excel.

C. SAP Lumira

O Lumira é uma solução apresentada pela gigante alemã SAP. Esta ferramenta permite a criação de gráficos e *dashboards*. O principal senão desta ferramenta é que ela só é gratuita para uso

pessoal e sem fins lucrativos. Para poder utilizar-se para o desenvolvimento de *software* com fins comerciais é necessário adquirir uma licença por utilizador final.

5. Análise Comparativa entre as Ferramentas Estudadas

Nesta secção será feita uma análise comparativa sobre as funcionalidades que as ferramentas até agora apresentadas disponibilizam. Através da comparação das funcionalidades disponibilizadas consegue-se ter uma ideia de quais as ferramentas mais indicadas para fazer BI, desde a simples realização de relatórios até ao desenvolvimento de uma solução de BI que envolva fazer ETL, *dashboarding*, olap, etc.

O Jaspersoft e o BIRT foram as ferramentas apresentadas que permitem fazer *reporting*. Ambas têm editores muito parecidos, tanto na maneira de manuseamento dos mesmos como na forma como se comportam. O Jaspersoft tem uma comunidade de utilizadores maior, mas isso não faz com que tenha um número, consideravelmente maior, de documentos de apoio e suporte. Tanto esta ferramenta como o BIRT têm muito apoio e documentação. O BIRT ganha vantagem no que toca à edição e importação documentos já existentes, pois a dificuldade com que se alteram documentos já existentes é a principal critica apontada ao Jaspersoft. O Jaspersoft ganha ao BIRT com o seu editor gráfico. O editor do BIRT é criticado por sofrer de alguns bugs, fazendo com que o utilizador tenha por vezes de reiniciar o mesmo. Estas duas ferramentas são muito idênticas, têm muita qualidade,e prova disso é serem incluídas noutras soluções de BI para a realização de *reporting* (o SpagoBI utiliza ambos os motores para a criação de *reports*, já o Pentaho utiliza o Jaspersoft como base para o seu motor de *reporting*).

O Talend e o Pentaho Data Integration foram as ferramentas apresentadas que permitem fazer ETL. Oferecem praticamente o mesmo número de funcionalidades. Os componentes a utilizar nos *jobs* são também muito idênticos e a forma de criar processos de ETL é feito de forma muito semelhante. O Pentaho Data Integration tem um leque maior de utilizadores, muito devido ao tempo em que está no mercado. Existe muita documentação para ambas as ferramentas, e isto faz com que o Talend, mesmo sendo uma ferramenta mais recente, não perca para o Pentaho Data Integration nesta área. Estas duas ferramentas são de elevada qualidade, e tal como o Jaspersoft e o BIRT, são incluídas noutras soluções de BI (o SpagoBI e o Jaspersoft ETL utilizam o Talend como motor responsável pelo ETL, e o Pentaho agrega na sua solução o Pentaho Data Integration).

O SpagoBI e o Pentaho são as soluções integradas apresentadas que permitem realizar um conjunto alargado de funcionalidades. Estas duas ferramentas apresentam praticamente o mesmo número de funcionalidades, contudo o SpagoBI ganha uma ligeira vantagem, devido ao

facto de permitir a realização de *self-service* BI, em que o utilizador pode criar os seus próprios relatórios a partir dos modelos de dados disponibilizados. No que toca ao apoio técnico o SpagoBI ganha com larga vantagem ao Pentaho, pois a versão gratuita do Pentaho apenas tem suporte nos forums da comunidade, já o SpagoBI tem suporte técnico especializado. Mesmo o facto do suporte e a consultoria serem serviços pagos, a organização que adoptar o SpagoBI terá maior segurança.

Outro aspecto importante a ter em conta é a facilidade de uso. O SpagoBI permite que utilizadores menos experientes consigam gerar documentos de forma bem mais rápida, e isto conduz a outra questão importante de analise que é a curva de aprendizagem de uma ferramenta. O SpagoBI tem uma menor curva de aprendizagem, devido à sua simplicidade de utilização e pelo facto das configurações por omissão não precisarem de ser muito modificadas para colocar a ferramenta a correr e a gerar documentos. O SpagoBI ganha noutro aspecto ao Pentaho que é a usabilidade dos gráficos disponibilizados. Estes são atraentes e bonitos, ao contrário daqueles que vêm configurados por omissão no Pentaho. Neste, se for pretendido ter o mesmo motor de gráficos que o SpagoBI, é necessária a instalação e configuração de um *pluggin*. Ambas as ferramentas podem ser embutidas noutras aplicações (de BI ou não). O SpagoBI tem vantagem de ter uma API (*Application Programming Interface*) responsável por esta integração, o que faz com que a colocação de um gráfico numa página web seja feita de uma forma mais rápida e simples. A grande desvantagem do SpagoBI face ao Pentaho é na documentação e suporte *online*. O Pentaho tem muita ajuda nos seus fóruns, ao contrário do SpagoBI, o que pode fazer com que algumas dúvidas e/ou esclarecimentos possam não ser esclarecidos, sem suporte e consultoria técnica dada pela empresa responsável pelo SpagoBI. Ambas as ferramentas são poderosas, no entanto o SpagoBI ganha alguma vantagem por ser mais simples e intuitivo.

Após uma análise relativamente ao estado da arte de ferramentas de BI *open source*, pretendemos comparar todas as ferramentas gratuitas analisadas anteriormente. A Tabela 1 apresenta uma comparação entre todas as ferramentas analisadas, tendo em conta as funcionalidades de cada uma. A tabela foi preenchida com as funcionalidades experimentadas e outras recolhidas de várias fontes consultadas.

As primeiras quatro ferramentas que aparecem na tabela não são consideradas *suites*. Por este facto, estas apresentam menos funcionalidades que as outras. A ferramenta Jaspersoft, tal como o BIRT, são ferramentas de *reporting*. Logo, as funcionalidades que apresentam são essencialmente ligadas a esta área, com destaque para a ferramenta Jaspersoft que também permite os utilizadores fazerem ETL. O Talend e o Pentaho Data Integration são ferramentas de ETL, logo, só disponibilizam esta funcionalidade.

Tabela 1 – Comparação de funcionalidades

Funcionalidades das Ferramentas									
Funcionalidades	Jaspersoft	BIRT	Talend	Pentaho Data Integration	SpagoBI	Pentaho	Rapid Miner	Jedox Palo	SAP LUMIRA
Reporting	✔	✗	✗	✗	✔	✔	✔	✔	✗
Ad-Hoc Reporting	✗	✗	✗	✗	✗	✔	✗	✗	✗
Dashboards	✗	✗	✗	✗	✔	✔	✔	✔	✔
Data Mining	✗	✗	✗	✗	✔	✔	✔	✔	✗
ETL	✔	✗	✔	✔	✔	✔	✗	✔	✗
Exportação de Dados	✔	✔	✗	✗	✔	✔	✔	✔	✔
Gráficos	✗	✗	✗	✗	✔	✔	✗	✗	✔
KPI's	✗	✗	✗	✗	✔	✔	✔	✔	✗
OLAP	✗	✗	✗	✗	✔	✔	✔	✔	✗

As últimas cinco ferramentas são do tipo *suite*. Assim, é normal apresentarem muitas mais funcionalidades que as primeiras quatro. O Rapid Miner, o Jedox Palo e o SAP Lumira apresentam menos funcionalidades que o SpagoBI e o Pentaho, mas o maior problema destas ferramentas é o facto de terem licenças gratuitas para uso individual, o que pode provocar algumas limitações à utilização a nível empresarial (por exemplo: descentralização de dados pelos utilizadores de cada sistema). O SpagoBI e o Pentaho são consideradas as melhores ferramentas porque disponibilizam todas as funcionalidades (à excepção do SpagoBI no que toca a relatórios *ad-hoc*) e são ferramentas relativamente fáceis de aprender e utilizar. Estas últimas ferramentas também têm a grande vantagem de poderem ser utilizadas para uso comercial, sendo que o SpagoBI ainda vais mais longe, pois é completamente *open source*. O SpagoBI e o Pentaho, pelos motivos já referidos, ganham destaque em relação a todas as outras soluções, e devido a isso, os casos de estudo seguintes utilizam estas duas ferramentas.

6. Casos de Estudo

Conhecidas algumas ferramentas, assim como as suas funcionalidades; vantagens e desvantagens, chega a hora de apresentar dois casos de estudo que utilizam duas das ferramentas apresentadas. Os casos de estudo seguintes foram escolhidos, em detrimento de outros, por apresentarem soluções interessantes aos problemas a que respondem, e por serem utilizados como cartão-de-visita pelas empresas responsáveis por estas ferramentas.

O primeiro caso de estudo mostra uma plataforma de BI na área da saúde desenvolvida com a ferramenta SpagoBI. O segundo caso de estudo mostra outra plataforma, na área do fabrico, desenvolvida com a ferramenta Pentaho.

Plataforma na Área da Sáude Desenvolvida com o SpagoBI

Este caso de estudo mostra uma plataforma desenvolvida para o ministério da sáude italiano [Cazzin 2007]. Os principais objectivos da plataforma são a monitorização do sistema nacional de saúde e o controlo dos dados clínicos de todos os pacientes italianos.

A Figura 12 mostra uma *dashboard* de controlo de desempenho e operabilidade de servidores do ministério da saúde. Através dos indicadores de KPI consegue-se perceber a utilização da CPU, por parte dos quatro servidores responsáveis por gerir processos de saúde, e depois através do gráfico por baixo destes KPI's conseguem-se identificar os processos, a correr em todos os servidores, mais dispendiosos ao nível da memória utilizada. Para cada processo, são mostradas várias métricas que ajudam a monitorizar o desempenho dos processos e, de modo mais geral, dos servidores.

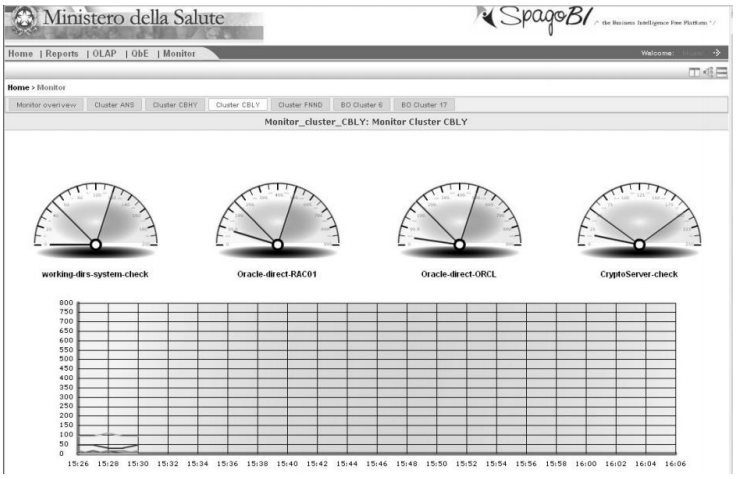


Figura 12 - *Dashboard* de controlo de performance de servidores (fonte: http://thirdnature.net/content/slides/SpagoBI_TDWI_May_07.pdf)

A ferramenta SpagoBI conseguiu implementar com sucesso toda a solução de BI a que se propôs. As questões que mais se destacaram nesta solução foram:

- Construção de *dashboards* analíticas constituídas por múltiplos componentes independentes;
- Navegação entre *dashboards*;
- Inserção de comentários e notas às *dashboards* e dados apresentados pelas mesmas.

Plataforma na Área do Fabrico Desenvolvida com o Pentaho

Este caso de estudo mostra a uma plataforma desenvolvida para a organização AMT (The Association For Manufacturing Technology). O principal objectivo da plataforma é ajudar os membros desta organização a tomar decisões (sobre vendas; contratos com fornecedores; encomenda de produtos, etc.) baseadas em informação económica da indústria de fabrico [Pentaho 2014]. Um requisito da plataforma é a utilização de diversas *dashboards* interactivas que disponibilizam informação em tempo real sobre diversas métricas económicas da indústria.

Através do Pentaho Data Integration foi possível correr consultas sobre dados importantes para os membros da AMT. O Pentaho Data Integration conectou-se a diversas fontes de dados para carregar informação com origem em ficheiros, sites governamentais e canais *web* com dados económicos disponibilizados a partir de *webservices*. A partir dos dados, carregados nas bases de dados referidas, foi possível criar *dashboards* com dados económicos. Uma questão interessante das *dahboards* criadas com o Pentaho é que as mesmas são compatíveis com a tecnologia web HTML5, e com isto foi possível criar dashboards adaptáveis para telemóveis e tablets, aumentando assim o número de formatos e dispositivos em que os utilizadores podem aceder à plataforma.

A Figura 13 mostra uma das *dashboards* da plataforma desenvolvida. Nesta *dashboard* consegue-se perceber que é constituída por indicadores de negócio (como por exemplo: unidades de produto e quota de mercado) e gráficos geográficos. Uma característica importante de realçar é a disponibilização de diversos filtros, estes ajudam os utilizadores a obter mais detalhes sobre informação específica.

Esta plataforma permitiu aos membros a tomar decisões que contribuíram para aumentar as vendas e diminuir as despesas. Os utilizadores também reportaram que com esta plataforma conseguiram melhorar a percepção das flutuações dos mercados, reforçando o apoio à tomada de decisões que permitiram aumentar as suas margens de lucro.

Figura 13 - *Dashboard* com informação económica relevante para tomada de decisões (fonte: http://www.pentaho.com/sites/default/files/uploads/resources/mtinsights_case_study.pdf).

7. Conclusão

A importância da implementação de BI nas organizações catapultou o número de ferramentas que permitem aplicar este conceito. Existem muitas ferramentas disponíveis no mercado, umas que necessitam da compra de licença para serem utilizadas e outras que são completamente gratuitas. As ferramentas pagas têm equipas dedicadas ao seu desenvolvimento e manutenção, adquirindo assim mais maturidade. Porém, o custo das licenças pode chegar aos milhares de euros. A BI não pode ser mais vista como um luxo dentro das organizações. Se existem organizações que não têm disponibilidade para adquirir as licenças das ferramentas pagas, então devem procurar outro meio de implementar BI.

As ferramentas gratuitas e/ou *open-source* surgem para colmatar a falta de recursos financeiros das organizações, permitindo ainda que os *developers* tenham possibilidade de ver, modificar e distribuir o código fonte das mesmas. Existe um elevado número de soluções deste tipo no mercado. Consoante os requisitos de uma plataforma de BI, podem ser adoptadas certas ferramentas gratuitas que respondam a estes requisitos, desde a simples geração de relatórios até soluções mais complexas que envolvam a implementação de *dashboarding*, OLAP, *self-service* BI, etc. A importância de analisar bem as ferramentas a escolher é vital, sendo necessário fazer um levantamento daquilo que se pretende fazer, a importância da utilização fácil e curva de aprendizagem, assim como a quantidade e qualidade de documentação existente e de suporte técnico para a solução de problemas ou o esclarecimento dúvidas. Optar por uma ferramenta que não responda às questões anteriores pode significar um fracasso na escolha da mesma, ou no pior dos casos, um fracasso da solução a desenvolver.

Foram analisadas e comparadas várias ferramentas ao longo deste documento, e é óbvio que todas elas são soluções viáveis para implementar BI. As ferramentas Jaspersoft e BIRT prometem a fácil criação de relatórios, desde relatórios estáticos até relatórios dinâmicos, que permitem ao utilizador filtrar e navegar informação no relatório em si. O Talend e o Pentaho Data Integration são ferramentas poderosíssimas na criação de processos ETL e no tratamento de dados. Estas ferramentas podem, inclusive, ser utilizadas fora do âmbito de aplicações de BI, visto que a extracção, tratamento e carregamento de dados entre fontes de dados diferentes é uma necessidade comum nas organizações actuais. O SpagoBI e o Pentaho são autênticos canivetes suíços que disponibilizam um alargado conjunto de ferramentas para BI. Estas soluções devem ser as primeiras candidatas à escolha de soluções que permitem implementar BI, desde *reporting*, OLAP, ETL, *dashboarding*, *self-service BI*, etc.

As ferramentas gratuitas e *open-source* estão no mercado para ficar. Espera-se que estas evoluam cada vez mais, apresentando-se como alternativas viáveis, e até mesmo concorrentes, às soluções pagas. Estas soluções são viáveis, exequíveis e apresentam qualidade nos resultados obtidos.

Referências

Anselmo, A., "Técnicas de pesquisa", Universidade Católica Portuguesa, 1988

Bonomo, P., *Arquitetura de Data Warehouse – Parte 02*, http://www.imasters.uol.com.br/artigo/11721/bi/arquitetura_de_data_warehouse_parte_02/imprimir, (2 de Março de 2009), 2015.

Cazzin, G., *SpagoBI for Healthcare*, http://thirdnature.net/content/slides/SpagoBI_TDWI_May_07.pdf, (Maio 2007), 2015

Choose, L., *Licenses*, http://choosealicense.com/licenses/, 2015.

Codd, E.F., "Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate", (1993), 2015.

Codd, E.F., "Further Normalization of the Data Base Relational Model". (Presented at Courant Computer Science Symposia Series 6, "Data Base Systems", Prentice-Hall, 1972.

Demarest, M., *The politics of data warehousing*, http://www.uncg.edu/ism/ism611/politics.pdf, (Junho de 1997), 2015.

Gangadharan, R. e Swami, N. "Business intelligence systems: design and implementation strategies", Information Technology Interfaces, 26th International Conference, 2004.

Inmon, W., *The Data Warehouse Budget*, http://www.datawarehouse.inf.br/papers/inmon%20budget-1.pdf, (2000), 2015.

Kimball, R., *The Data Warehouse ETL Toolkit*. *Practical Techniques for Extracting*, *Cleaning*, *Conforming and Delivering Data*, Wiley, 2004.

Lönnqvist A. e Pirttimäkib V., "The Measurement of Business Intelligence", Volume 23, Issue 1, 2006.

Negash, S. "Business Intelligence," Communications of the Association for Information Systems: Vol. 13, Article 15, 2004.

Pentaho, *Powerful and Scalable Embedded Analytics for Manufacturing*, http://www.pentaho.com/sites/default/files/uploads/resources/mtinsights_case_study.pdf, (2014), 2015.

Pereira, R., "BI nas Organizações", Universidade Lusófona de Humanidades e Tecnologias, 2011

Rouse, M., *ad hoc analysis*, http://searchbusinessanalytics.techtarget.com/definition/ad-hoc-analysis, (Julho de 2010), 2015.

Rouse, M., *self-service business intelligence (BI)*, http://searchbusinessanalytics.techtarget.com/definition/self-service-business-intelligence-BI, (Abril de 2012), 2015.

Shilakes, C., *Enterprise Information Portals*. *Enterprise Software Team*, http://www.sagemaker.com/company/downloads/eip/indepth.pdf, (2008), 2015.

Tereso M, "Ferramentas de Business Intelligence Open Source para PMEs", Instituto Politécnico de Coimbra, 2011.

Watson, H.J. e Wixom, H., "The Current State of Business Intelligence", IEEE, 2007.

Wingyan, C., "Discovering business intelligence from online product reviews: A rule-induction framework", Expert Systems with Applications, Volume 40, Issue 1, 2013.

Information management in supply networks: a framework to support risk management decisions

Dario Messina [1,2], Ana Cristina Barros [2],
António Lucas Soares [1,2]

1) Faculty of Engineering, University of Porto, Portugal
dario.messina@fe.up.pt , als@fe.up.pt

2) *INESC TEC,* Portugal
acbarros@inescporto.pt

Abstract

This thesis aims at contributing with new research at the intersection of the research fields of information management and supply chain risk management. Information is a fundamental element for an effective risk management in supply networks. In spite of this, research on how to use internal and external information for risk management decisions in supply networks is scarce. Therefore, this thesis main goal is to characterize information models used in supply networks and the supply chain risk management process in order to propose an aggregated model of internal and external information for risk management decision support in supply networks. This work is based on the Design Science paradigm, which will be applied through several research methods: systematic literature review to analyse the current state of the art on information management in supply networks; field research to analyse the influence of different information organisation models in the process of supply chain risk management and to develop an information model based (artifact); and action research to implement the artifact for the definition of risk mitigation strategies in aeronautic supply networks.

Keywords: Information Management, Supply Chain Risk Management, Information Model, Supply Network, Design Science.

1. Research Situation

This thesis is an interdisciplinary work at the intersection between information management and supply chain risk management. I started my Ph.D. in September 2014 in the "Engineering and Industrial Management" Doctoral Program, and taking into account the two areas of this work I think they perfectly fit in a modern and competitive industrial context. Regarding the benefits of my participation at CAPSI Doctoral Consortium please refer to the CV.

2. Context and Motivations

During the past decades Original Equipment Manufacturers (OEM's) have been increasingly focusing on their core competences (overall design, architecture, integration, and final assembly and delivery to end customers), though becoming large scale integrators by the coordination of a supply network [Ro et al., 2008]. Consequently, OEM's have today a higher integration with first-tier suppliers, but a looser control over second-tier suppliers, mainly due to lack of contractual relationship and the difficulty of strategic alignment between OEM and second-tier suppliers [Fawcett and Magnan, 2002]. Therefore, first-tier suppliers have assumed a critical

role in the supply chain, since they have to coordinate second-tier suppliers and help OEM's managing supply chain risk. This leads to the development of new supply chain capabilities in terms of supply chain risk management. These changes in the role of supply chain players calls for more research on how supply chains risk are being managed, and how information should be managed among supply chain partners to support risk management related decision-making. Consequently, this thesis aims at addressing this research gap and its research design is outlined in the following sections. Section 3 reviews the literature on information management in supply networks, supply chain integration, and risk management. Section 4 defines the problem underlying this thesis while in section 5 research questions, objectives and research methodologies are presented. Section 6 shows the actual status of this thesis work. The last section provides the main conclusions and points future research work.

3. Fundamentals and Related Work

The literature review of this thesis is structured in accordance with the main topics representing the pillars of this work. We start analysing how information is managed in supply networks, then we describe supply chain integration with a focus particularly on the risk management process, in order to identify the best strategies for supply networks risk mitigation.

Information Management in Supply Networks

A supply network is defined as a collection of interconnected organisations with the primary purpose of procure, use and transform resources in order to provide goods and services [Harland et al., 2004]. An important aspect for the management of supply networks is to establish how the information is acquired, organised and used among members, towards a proper risk management of the supply network [Hung et al., 2011]. Also, a correct management of the information flow is critical for increasing visibility over the network [Barratt and Oke, 2007] in order to support decision-making process and mitigate the risk [Yu and Goh, 2014]. Even though previous research has confirmed the strategic importance of information and several authors were interested in developing information models to proper manage this information flow [Chi, 2010], there is still a need to explore in depth this field in order to analyse how different information models influence the risk management decision process of supply networks.

Supply Chain Integration, and Risk Management

Due to the global and turbulent environment in which firms compete, several researchers have highlighted the need for a closer relationship between manufacturers and their partners as a way to improve supply chain performance [Childerhouse and Towill, 2011]. Even if across industries the level of implementation results differ, supply chain integration seems a good alternative to the verticalisation of operations [Figueiredo et al., 2008]. In order to ensure supply chain integration is necessary that supply chain partners share and collaboratively manage processes [Flynn et al., 2010]. According to the literature, in an integrated supply chain, the information is the key element that firms have to share with their partners [Ragatz et al., 2002]. Within the supply chain, the typical information shared is related with material requirements, demand forecasts and inventory levels (Chandra et al., 2007, Omar et al., 2010, Datta and Christopher, 2011, Hung et al.,2011). To ensure the effectiveness of long-term collaborations and to generate trust among partners, these information exchanges should be held within processes with high level of visibility, which should desirably enable effective communications and exchange of real-time data [Chengalur-Smith et al., 2012].

Companies also value the importance of establishing a risk management process to identify, measure, mitigate, and control risks in their supply network (Elsalmi and Hachicha, 2014). In order to be able to identify these risks, companies are starting to work with their network partners towards creating supply networks visibility (Nooraie and Parast, 2015). By sharing

what sometimes is sensitive and proprietary information with their network partners, companies aim at aligning their common objectives though ensuring the efficient management of the whole chain. Still, there is a need to create visibility over the specific information that will enable companies to identify and act upon the risks and opportunities of their supply network. In fact, in the global and digital context in which firms operate, information assumes a distinctive role, as it supports effective decision-making process. In this thesis we propose a framework for the organization of information among a supply network that enables the identification of strategies for supply network risk mitigation.

4. Research Problem definition

The aeronautic sector has shown a trend in the past decade in providing new solutions related to the management of new product development (Strategy&, 2015 (a), (b)). This sector is characterized by increasing costs in developing new products, and consequently there is the need to transfer the risk downstream the value chain. In this context, supply chain is characterized by a geographical dispersion of partners, and an increasing need of communication and coordination that is possible to achieve sharing vital information [Pereira, 2009] and through the adoption of ICT solutions.

Considering the scarcely research on how to use internal and external information for risk management decisions in supply networks, and the relevance that supply chain risk management and mitigation plays for practitioner (Global supply chain institute report, 2014), this research thesis will characterize the information models in supply network in order to develop and implement an information model, our artifact, that support strategic decision making in supply network in order to assess the risk along the whole chain, and provide a practical solution for decisions related to the configuration of the network. In particular, this work will be based on the specific needs of the aeronautic sector.

5. Objectives, Methodology and Research Methods

This thesis presents three research questions:

RQ1: What models exist for managing strategic information in supply networks?

RQ2: How information organization models influence the process of supply chain risk management?

RQ3: What are the specifics that characterise an information organisation model for supply chain risk management and mitigation?

RQ4: How the developed model support strategic decision-making related to aeronautic network configuration?

The research questions led to formulate the following goals:

- To characterize information management models used in supply networks;
- To describe the risk management process of supply networks and the role that first-tier and second-tier suppliers have in this process;
- To propose a multi-view and aggregated model of internal and external information that allow to bring to the surface the main causes of risk, and to identify the strategies to mitigate it.

5.1. Design Science Research Paradigm

The centrality of design activity has been recognised in several disciplines, and particularly it results even more crucial in the areas of Information Systems. Information systems for its own nature represent the meeting point between Information Technology (IT) and human interface, and for this reason it is characterised by challenging problems related to design activity in order to provide unique and creative solutions [Hevner and Chatterjee, 2010]. Based on the seminal book of Herbert Simon "*Sciences of the Artificial*", a pool of scholars in information systems sets the basis of a new paradigm, called Design Science Research Paradigm.

This thesis work is based on Design Science Research Paradigm, and the schema in fig. 1 represent the structure underlying it. The ultimate goal of design science is to address research in order to develop artifacts that, once tested, are able to provide innovative solutions for specific business needs [Holmstrom *et al.,* 2009].

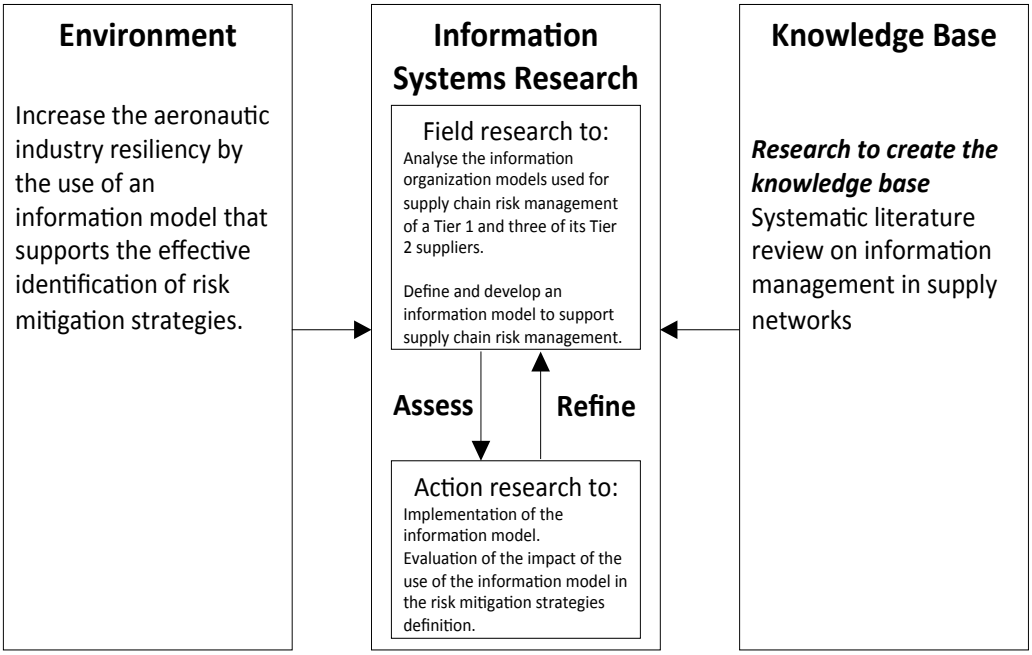


Figure 1 - Design Science Research Framework. Based on Havner et al.[2004]

Starting from the "environment" block we want to identify the needs on the aeronautic sector, then we adopt systematic literature review to built the foundation of the "knowledge base" block, and finally through the adoption of field research and action research, belonging to the "information systems research" block, we want to develop, implement and test an information model to support strategic decision in supply network. The purpose of design science is to develop new knowledge to provide to practitioner in order to design solutions for specific field problems [van Aken, 2005].

Due to the exploratory nature of this research a portfolio of qualitative research methods are used: systematic literature review to analyse the current state of the art on information management in supply networks; field research to analyse the supply chain risk management process of a first-tier and three of its second-tiers in supply networks; and action research to implement and test an information model for the evaluation of the impact of the use of the information model in the risk mitigation strategies in aeronautic supply networks. The different methodologies here are applied in order to fulfil the three different section of the design

research framework presented in fig.1. The following explains briefly the application of each of these methods.

Systematic Literature Review

The systematic literature review is the methodology that allows to build the theoretical foundation, and to ensure rigor in this thesis project. Considering the strategic relevance of information for supply chain management a systematic literature review will be carried out in order to discuss and synthesize the literature on information management in the context of supply networks. This research method is considered appropriate because it allows to identify the most relevant scientific contribution to a field or question [Becheikh et al., 2006], by adopting a replicable, scientific and transparent research process [Tranfield et al., 2003], In order to assure the quality of the review process, search terms will be selected from the literature, a detailed review protocol will be developed, and a team of four researchers will be involved in the paper selection and analysis phases.

Field Research

After describing how information is managed in supply networks, it is important to investigate how supply chain risk management is influenced by the information organisation models adopted, in order to identify the specifics necessary to develop our own information organisation model for supply chain risk management and mitigation. Considering the exploratory nature of this work, field research was adopted. A highly complex supply chains will be selected for this study, namely aircraft maker industry. Data collection will be carried out by means of semi-structured interviews at first-tier and three of its second-tier suppliers belonging to the network. All the interviews will be recorded, transcribed and then coded with the support of specific software (MAXQDA).

Action Research

The results from the two research works above will contribute for the design of an action research study. In particular the action research will act as *trait d'union* between the business needs provided in the field research and the applicable knowledge resulting from the systematic literature review, in order to implement and test the relevance of the developed artifact and the rigor underlying it.

Action research is appropriate to investigate the development of solutions for actual problems [Eriksson and Kovalainen, 2008] where researchers and operations managers collaboratively contributed to the development of solutions [Gummesson, 2010]. In this thesis, action research allows to implement our information model for risk management of supply networks, and to observe how the artifact is used in a manufacturing environment.

6. Thesis Status

A comprehensive analysis of the literature was made both in the field of Information Management and Supply Networks to identify the information management models adopted.

Information management models

At the moment, preliminary results exist only from the Systematic Literature Review. In particular the results of this methodology can be resumed in Table 1:

	STAGES			
Chi (2010)	Problem domain identification	Acquisition	Modelling	Ontological information representation

Haug (2013)	Acquisition	Representation	Evaluation	
Maçada *et al*., (2013)	Definition	Acquisition	Distribution	Usage

Table 1: Information Model in Supply Network

After the analysis of models obtained, the one proposed by Maçada *et al.*, (2013) was considered the most complete in the sample, for its characteristics. Even so, this model doesn't seem exhaustive, we found it lacking of some additional stage between the acquisition and the distribution stages. Analysing other generic information model in the literature, two of them presents some additional stages, namely in Choo (2002) and Detlor (2010). Choo's model is characterised by the following six stages: identification of information needs, acquisition, organisation and storage, development of product and services information, distribution, and usage; also Detlor's model uses six stages, that are: creation, acquisition, organisation, storage, distribution, and usage.

Merging the information from the three models we develop our own characterised by the following six stages: identification of information needs, acquisition, organisation, storage, distribution, and usage. Regarding the other methodologies, we plan to conduct interviews in Spain and Portugal in order to create the base for field research, and so provide all the elements needed to move to the third and last methodology, the action research.

7. Conclusions and Expected contributions

The topic of information management in supply chains has been extensively studied in the past, though most researchers focus on the dyad OEM - first-tier or the dyad first-tier - second-tier.

This paper sheds lights on how the risk management process is currently carried out in a complex multi-tier supply network and the role supply chain partners have in this process. Furthermore, this thesis contributes to the supply chain management field with a multi- view and aggregated model of internal and external information for decision support in supply networks in aeronautic sector. The implementation and pilot use of this model in a real setting will confirm the relevance of this work for practitioners and develop further managerial insights on the implementation of such models in practice.

Acknowledgements

This work is financed by the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project CMUP-ERI/TPE/0011/2013 of the CMU Portugal Program.

8. References

Barratt, M., Oke, A. "Antecedents of supply chain visibility in retail supply chains: a resource-based theory perspective." *Journal of operations management*, *25*(6), (2007): 1217-1233.

Becheikh, N., Landry, R., Amara, N. "Lessons from innovation empirical studies in the manufacturing sector: A systematic review of the literature from 1993–2003." *Technovation*, *26*(5), (2006): 644–664.

Chandra, C., Grabis, J., & Tumanyan, A. "Problem taxonomy: a step towards effective information sharing in supply chain management." *International Journal of Production Research*, *45*(11), (2007), 2507-2544.

Chengalur-Smith, I., Duchessi, P., Gil-Garcia, J. R. "Information sharing and business systems leveraging in supply chains: An empirical investigation of one web-based application." *Information & Management*, *49*(1), (2012): 58-67.

Chi, Y. L. "Rule-based ontological knowledge base for monitoring partners across supply networks." *Expert Systems with Applications*, *37*(2), (2010): 1400-1407.

Childerhouse, P., Towill, D. R. "Arcs of supply chain integration." *International journal of production research*, *49*(24), (2011): 7441-7468.

Choo, C.W. *Information management for the intelligent organization: the art of scanning the environment*. Information Today, Inc., 2002.

Datta, P.P. and Christopher, M.G. "Information sharing and coordination mechanisms for managing uncertainty in supply chains: A simulation study", *International Journal of Production Research*, Vol. 49 No. 3, (2011), pp. 765–803.

Detlor, B. Information management. *International Journal of Information Management, 30*, (2010) 103-108.

Elmsalmi, M., & Hachicha, W. (2014, May). Risk mitigation strategies according to the supply actors' objectives through MACTOR method. In *Advanced Logistics and Transport (ICALT), 2014 International Conference on* (pp. 362-367). IEEE.

Eriksson, P., & Kovalainen, A. *Qualitative methods in business research*. Sage. 2008

Fawcett, S. E., Magnan, G.M. "The rhetoric and reality of supply chain integration." International Journal of Physical Distribution & Logistics Management 32.5 (2002): 339-361.

Figueiredo, P., Silveira, G., Sbragia, R. "Risk sharing partnerships with suppliers: the case of Embraer." *Journal of Technology Management & Innovation*, *3*(1), (2008): 27-37.

Flynn, B. B., Huo, B., Zhao, X. "The impact of supply chain integration on performance: A contingency and configuration approach." *Journal of Operations Management*, *28*(1), (2010): 58–71.

Global supply chain institute report, 2014: http://globalsupplychaininstitute.utk.edu/publications/documents/Risk.pdf

Gummesson, E., "Qualitative Methods in Management Research", Sage, 2000.

Harland, C., Zheng, J., Johnsen, T., Lamming, R. "A Conceptual Model for Researching the Creation and Operation of Supply Networks". *British Journal of Management*, *15*(1), (2004): 1–21.

Hevner, A., Chatterjee, S. Design Science Research in Information Systems Good. In *Design Research in Information Systems* (pp. 9–22). Springer Science+Business Media. 2010

Hevner, A., March, S. T., Park, J., & Ram, S. Design science in information systems research. *MIS quarterly*, *28*(1), (2004): 75-105.

Holmström, Jan, Mikko Ketokivi, and Ari-Pekka Hameri. "Bridging practice and theory: a design science approach." *Decision Sciences* 40.1 (2009): 65-87.

Hung, Wei-Hsi, Chin-Fu Ho, Jau-Jeng Jou, Yi-Ming Tai. "Sharing information strategically in a supply chain: antecedents, content and impact." *International Journal of Logistics: Research and Applications* 14, no. 2 (2011): 111-133.

Maçada, A. C. G., Costa, J. C., Oliveira, M., Curado, C. "Information management and knowledge sharing in supply chains operating in Brazil." *International Journal of Automotive Technology and Management, 13*(1), (2013): 18–35.

Nooraie, S. V., & Parast, M. M. (2015). A multi-objective approach to supply chain risk management: Integrating visibility with supply and demand risk. *International Journal of Production Economics*, *161*, 192-200.

Omar, R., Ramayah, T., Lo, M.-C., Sang, T.Y. and Siron, R. "Information sharing, information quality and usage of information technology (IT) tools in Malaysian organizations", *African Journal of Business Management*, Vol. 4 No. 12, (2010), pp. 2486–2499.

Pereira, J.V. "The new supply chain's frontier: Information management." *International Journal of Information Management* 29.5 (2009): 372-379.

Ragatz, G. L., Handfield, R. B., Petersen, K. J. "Benefits associated with supplier integration into new product development under conditions of technology uncertainty." Journal of Business Research, 55(5), (2002): 389-400.

Ro, Y. K., Liker, J.K., Fixson, S.K. "Evolving models of supplier involvement in design: the deterioration of the Japanese model in US auto." Engineering Management, IEEE Transactions on 55.2 (2008): 359- 377.

Strategy&, 2015 (a): http://www.strategyand.pwc.com/perspectives/2015-aerospace-defense-trends

Strategy&, 2015 (b): http://www.strategyand.pwc.com/perspectives/2015-aviation-trends

Tranfield, D., Denyer, D., Smart, P. "Towards a methodology for developing evidence in formed management knowledge by means of systematic review." British journal of management 14.3 (2003): 207-222.

Van Aken, J. E. "Management research as a design science: Articulating the research products of mode 2 knowledge production in management." *British Journal of Management*, *16*, (2005): 19–36.

Yu, M. C., Goh, M. "A multi-objective approach to supply chain visibility and risk." *European Journal of Operational Research* 233, no. 1 (2014): 125-130.

Information systems design and user satisfaction: A banking case study

Bruno Padinha[1], Henrique O'Neill [2]

1) ISCTE-IUL Instituto Universitário de Lisboa, Portugal

bpadinha@gmail.com

2) ISCTE-IUL Instituto Universitário de Lisboa, Portugal

henrique.oneill@iscte.pt

Abstract

With the growing, pervasive use of information systems (IS) by most organizations, particularly companies in technology-intensive businesses like banking, questions arise as to the effective success of those systems, how to measure it and, most of all, what influences it. User satisfaction, even if not an end goal of IS organizations, is frequently used as a measure of that success and it has been empirically shown to correlate with other more direct measures.

The research reported here looked at what influences user satisfaction, particularly what characteristics of information systems and their development processes tend to result in higher user satisfaction, in the specific case study of a Portuguese bank, which conducts a comprehensive survey every year to measure the satisfaction of its thousands of IS users regarding its numerous information systems.

Additional primary data was collected to characterize the information systems and their development processes, which, once structured and standardized alongside the user satisfaction survey results provided the basis for a quantitative analysis of the relationships between IS success, measured by user satisfaction, and its antecedents, using correlation and multiple linear regression techniques.

The significant relationships are analyzed and the conceptual model's results discussed in light of prior theory on the subject. Several implications for the practice and research communities are finally identified.

Keywords: information systems success; information systems design; user satisfaction; banking information systems

1 Introduction

It is quite common nowadays for companies and governmental organizations to use information systems (IS)[1] to automate activities and support their business activities. These systems enable organizations to become more effective and efficient in delivering products and services to their customers and have become therefore intrinsic to the businesses themselves, as is the case of many high-tech enabled industries like banking and telecommunications, which could not exist in their modern form without extensive use of information systems. Banks and other financial services institutions are a case in point, spending between 5% and 9% of their operating income in information technology (IT) (Gopalan et al. 2012). A direct contribution of IT assets and capabilities to the increase of productivity and general performance of businesses across many sectors of the economy has been reported (Bharadwaj 2000; Farrell et al. 2002). For example, a study of public banks in the US concluded that "a firm's IT capability is crucial for its competitive advantage" (Lin 2007).

As a result, bank employees are frequent users of information systems, which give them access to customers, products and services information, and allow them to process transactions of all kinds. The systems are usually heterogeneous, based on different technologies, with various user interfaces and functionalities and differing integration levels of the captured and processed information (Heidmann 2010).

Because system use is usually compulsory for users, it happens that they exhibit varying levels of satisfaction regarding each system, according to how they perceive it or not as an efficient, effective tool for accomplishing their tasks and responsibilities. In general, this user satisfaction is derived both from how each information system supports them in the fulfillment of their day-to-day activities, as well as, in a broader sense, from how easy the system is to learn and use, how quickly and responsively it performs the needed calculations and other information processing tasks and how effective the information systems department is at developing, operating and maintaining the organizations' information systems, among other factors (Al-Maskari et al. 2010).

For the information systems department of such a large and complex organization as the typical bank, the user satisfaction is very important to measure and monitor, not only because users are their immediate,

[1] The expression Information System, as applied here, means a specific application of Information Technology to some strict set of functional needs of a company or organization such as the Bank, for a certain business process or function. Throughout the text, groups of these information systems are sometimes referred to as an application, the term that the (studied) Bank's IT management employs when communicating to the users to measure their satisfaction. Other authors deal with the ambiguity of the expression with a broader definition of Information System, such as "*[…] either some aspect of an application of information technology (IT), one individual application, a group of applications (including those of an entire organization), or an application of one type of IT.*" (Seddon 1997)

internal customers, but also because it has been shown that it is usually an acceptable and objectively-measurable proxy for the success of the IS organization (Sedera et al. 2005).

If user satisfaction measures the success of the information systems of an organization, a natural question arises, then, of what can management do to increase it in a significant and sustainable way?

It is immediately apparent that for the most part, user satisfaction cannot be changed directly, since it is based on a subjective perception, but rather it has to be influenced via the more objective characteristics of the information systems and the IS organization that provides them, such as their technical, their development processes and the organizational and governance models of the entities that manage the information systems of an organization.

Many of the decisions in these areas are, understandably, also not trivial. Management has to analyze and decide upon information systems' topics like those and hope to make the right decisions regarding their influence on the ultimate success of the information systems, which will be patent in user satisfaction as well as in actual benefits accrued from their usage.

The research reported here comprised the study of a specific case, that of a Portuguese bank, which has numerous information systems and whose management is keen on conducting a yearly appraisal of the satisfaction of its users (just referred to as "*the Bank*" for confidentiality).

Throughout the years, questions have arisen as the Bank observes the ups and downs of user satisfaction with each of its main information systems and attempts to identify their root causes, in the hope of reinforcing the positive ones and eliminating or at least minimizing the negative ones. While the Bank's IS management attempts to identify and isolate the reasons that lead to the evolution of user satisfaction, there had never been a systematic and objective search for its antecedents.

That was an important factor driving the execution of this study, since the increasing challenges of providing more IT and IS services, with a budget that is regularly expected to decrease, require a comprehensive understanding of what leads to users satisfaction and therefore, ultimately, to IS success.

This research looked at the case study of this bank in light of the existing knowledge in the field of Information Systems Research, in particular the findings regarding user satisfaction, information systems adoption and success and its antecedents. The main objectives were, in summary, to increase the understanding of the relationship between satisfaction and the characteristics of information systems and their development processes in the specific case of this Bank. With that in mind, the problem was formulated as follows: *Which characteristics of information systems and their development processes influence user satisfaction?*

2 Methods

Research design

Given the stated problem and research objectives already presented above, the approach chosen was that of a case study, which has benefited from the long-term relationship of the author as a consultant with the Portuguese bank that was the object of the study. This research initiative was conceived as descriptive, leaning essentially on the collection and analysis of quantitative data.

Focusing on a specific case study allowed for a deep analysis of the details of the Bank's information systems, eliciting rich information on their characteristics, the Bank's development processes and the ultimate user satisfaction with the applications, as well as ensuring that information was interpreted in the correct context, an important tenet of the case study method.

Object of the study

The Bank is a medium-sized financial institution operating in the Portuguese Market since the early 20th century, with net assets of 13 billion euros and close to 700 branches at the end of 2013.

The group includes the core banking operating units as well as specialized companies operating in related businesses. In addition, there are a number of corporate and shared service centers providing centralized support functions like marketing, legal and IT.

The Information Systems and Information Technology shared service center fulfills the role of an IS department of the Bank and other group companies, providing the hardware, software and services that enable the use of information technology by the business units. That center was therefore the main source of data for this study, and is henceforth referred to as "the Bank's IS Department" or simply "the IS Department".

Theoretical and conceptual framework

The relevant literature was reviewed, starting with the seminal papers on the key topics of information systems success, user satisfaction, technology acceptance and software development, then proceeding to include their references and subsequent references, as well as other research and documentation found via keyword search for specific topics. Several important theoretical and conceptual constructs emerged from this review, which informed the subsequent design and field work.

The idea for this research came from the practice of the Bank's IS Department to measure the users' satisfaction every year with a comprehensive survey. This prompted the question of what can management do to increase that satisfaction, which could be informed by an analysis of processual determinants that were likely to have a positive or negative influence in the user's satisfaction. Satisfaction, therefore, took the place of dependent variable and a quest ensued to identify suitable independent variables to test. Even though user satisfaction is a perception, therefore a subjective measure, the research strived to collect and analyze real, objective measures like training and testing approaches.

In order to design a conceptual model for this research consistent with the theory and which can be applied to the case study, a preliminary interview took place with the Bank's IS Department leadership for identifying which data was available to operationalize variables that corresponded to the concepts and constructs in the literature, particularly the D&M model (DeLone et al. 1992; DeLone et al. 2003) and its more recent taxonomy of IS success determinants (Petter et al. 2013). This lead to the identification and subsequent collection and standardization of data related to the main variables and constructs, which are grouped according to those models and associated research in Figure **1**. Based on the collected data, the simpler conceptual model for this study was then designed, as presented in Figure **2**.

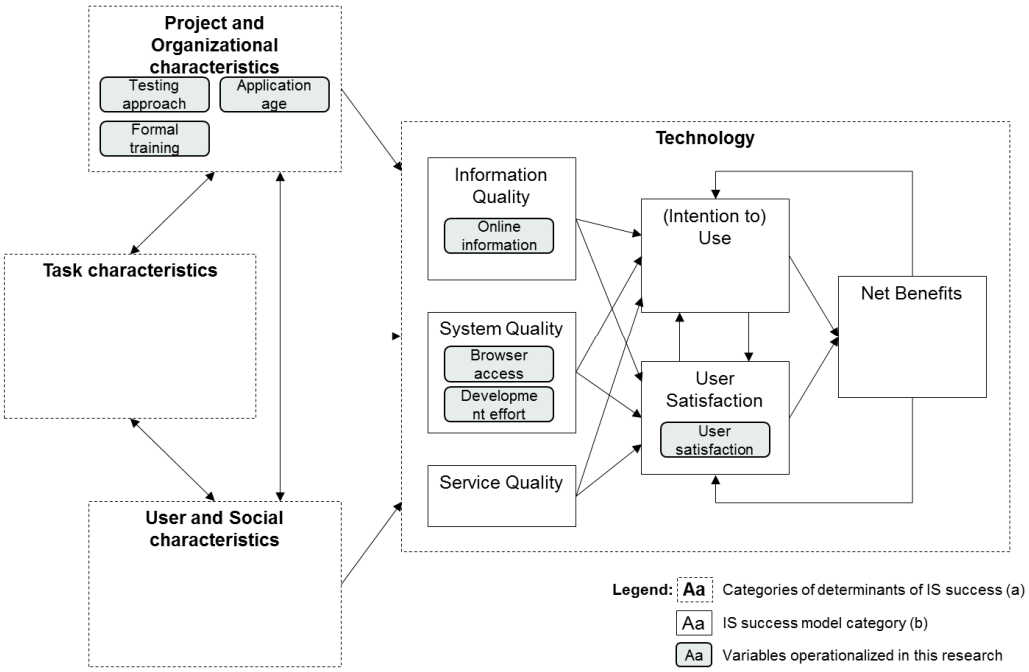


Figure 1 Collected variables shown in light of theoretical research and models (sources: (a) (Petter et al. 2013), (b) (DeLone et al. 2003))

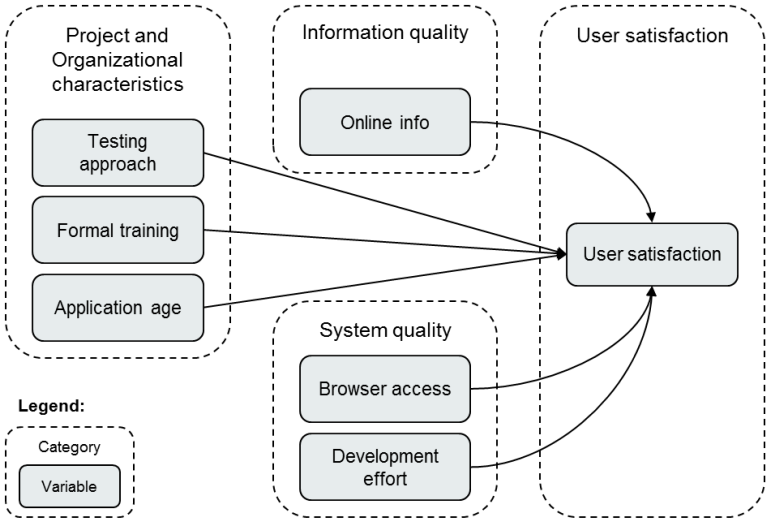


Figure 2 Conceptual model

Data collection

Once the conceptual basis was established, research then proceeded with the collection of data related to the user satisfaction as well as the characteristics of the Bank's information systems and their development processes which could serve to operationalize the dependent and independent variables in the research model.

The collected data covers a time span of five years, from 2008 to 2012, which existed thanks to the Bank's commitment to regularly monitor its IS business processes and annually measure the corresponding user satisfaction.

The Bank's IS Department conducts three yearly satisfaction surveys: an IS user satisfaction survey, a management IS satisfaction survey and an internal IS employee satisfaction survey. All of these are issued electronically to the Bank's employees, with the responses being collected anonymously and analyzed by an external entity. The first of those surveys was used as the source of information for operationalizing the dependent variable in this research. It has consistently high response rates, above 40%.

The set of responses to those surveys over a five year period (2008 through 2012) was used as data for the dependent variable, The user satisfaction surveys considerably resemble some of the most widely used instruments for measuring satisfaction found in the literature review (Bailey et al. 1983; Doll et al. 1988) and which have been empirically validated (Melone 1990). They have three main sections, covering (i) the user's satisfaction with the Bank's information systems, (ii) the IS Department's provision of services and (iii) the importance of those systems and services to the user. The data collection for the Satisfaction variable focused on the first of these sections.

Although some changes are made every year to the surveys, mainly to add questions related to newly introduced systems or services and to remove those which are no longer in use, the majority of the items related to the main banking applications have been regularly included in the surveys for at least a majority of the years included in the research. Each question in each questionnaire was associated, whenever possible, to a main application in the information systems inventory already collected, and coded numerically.

The user satisfaction survey questions ask the respondents' to rate their satisfaction with each main application developed and operated by the Bank's IS department. The responses are recorded in the form of a 4-point Likert-type item, a commonly used approach for measuring satisfaction (Bailey et al. 1983; Carifio et al. 2007; Melone 1990). The employed scale has an even number of points which prevents indecisive responses, for which the users are asked to select the "*Don't know*" option instead.

Although the same Likert-type scale was used in all the questionnaires, the coding of the responses was not exactly the same in each year. Therefore, carefully recoding of the survey responses was undertaken in a spreadsheet program and then all the responses were consolidated into a single satisfaction variable source file, carefully mapping each question to the appropriate application.

While Likert-type scales comprised of multiple items are commonly used for measuring user satisfaction, the use of single-item Likert-scales has also been published, as was the treatment of such variables as continuous, in fact widely used in many instruments to measure user satisfaction (Hayes 1998), and several authors back their use, if carefully applied, as continuous or, at least, ordinal variables (Botelho et al. 2011; Göb et al. 2007).

The operationalization of the variable for user satisfaction regarding each application in each year was achieved by averaging the valid responses from all users to the relevant question in that year's questionnaire. The median was also calculated as an additional measure of satisfaction, but later discarded since it proved to be too limited in range for many analyses and did not yield significantly different results for the remaining ones.

While the dependent variable data came directly from those yearly user satisfaction surveys, the operationalization of the independent variables required the collection of and manipulation of data from many sources in order for it to be in a format that was compatible with the user satisfaction data, for which the approach will now be described.

An obstacle that had to be overcome in the data collection process was the level of detail at which the information systems in the Bank are analyzed and recorded: each of the Bank's main applications is comprised of one or more underlying systems, each of them fulfilling a specific role. The Bank's IS Department management provided an exhaustive inventory of these underlying systems, including key

characteristics (names and short functional descriptions), vendors, types of technologies used, launch dates, how they were developed and to which main application they belong to.

While in the user satisfaction surveys' questions, global application names are referred to, many other sources of information record the data at the underlying system level of detail. Since satisfaction was the driver for this research, in the cases in which data was collected at the system level of detail, it was aggregated in some way for all the systems which are associated with a specific application.

The IS Department regularly executes large-scale projects, which develop new features or even replace entire existing systems and applications altogether, as well as smaller-scale initiatives dealing with features requested by the end-users, and other smaller application maintenance initiatives.

Information was gathered on the projects and requests that were executed throughout the period in analysis, namely which ones contributed to the development of each system and application, how many person-days were employed, when did the activities take place and some of the methodological approaches undertaken.

Although some effort was spent trying to elicit additional information regarding these projects, sadly no other structured data could be obtained about the software development life-cycle, since the main change that was introduced at the Bank in this period, which was the level 2 CMMI certification of the development processes, was achieved precisely in 2007, prior to the projects under analysis taking place, which means that all of the projects for which data has been collected were implemented already following the new process model.

In total, 96 main projects, 142 user requests and 2,020 maintenance initiatives that took place from 2008 to 2012 were analyzed and either associated with the systems and applications that they dealt with or discarded if they did not directly affect the applications being studied. Each project was individually mapped to each of the Bank's applications' underlying systems and then finally aggregated for each of the main applications.

After all the information was collected and processed, tables summarizing and relating the various items were prepared and submitted for validation by the management of the IS Department of the Bank, in order to have an additional measure of confidence on the collected and processed data.

By combining the information on the applications and systems, on one hand, and projects, requests and initiatives, on the other, the independent variables where operationalized, grouped according to their ability to measure key concepts from the literature: (1) Project and Organizational characteristics and (2) application characteristics.

In some cases, dummy variables were coded to represent the presence or absence in an application of certain characteristics in its underlying systems. For example, there were two alternative dummy codings

related to the execution of formal acceptance tests for each underlying system associated with a certain application, the first one following a more conservative approach (1 only if formal acceptance tests took place for all underlying systems) and the second one a laxer approach (1 if any system went through testing). As described later, in the Results section, a correlation significance test led to the selection of the more robust of these two approaches.

For convenience in some analysis, person-years were derived by accumulating 12 months of 21 days of 8 hours of work.

The final case file was then built by combining the independent variables with the appropriate lines in the application satisfaction case file.

Analysis of variable behavior and inter-relationships

After all the information was collected, validated, standardized and consolidated into a single case file, analysis started with descriptive measures of the dependent and independent variables: central tendency measures and dispersion measures. The evolution of satisfaction and of the independent variables over time was thus analyzed, as a whole, as well as per main application.

The next step focused on evaluating the correlation between the candidate independent variables and the dependent variable, User Satisfaction, after which some variables were considered for discarding and others were selected to be applied for operationalizing the conceptual model (Figure **2**). This was required due to the already mentioned fact that there were several possible approaches to coding some of the literature concepts from the collected data. The final variables used to operationalize the conceptual model are shown in Figure **3**.

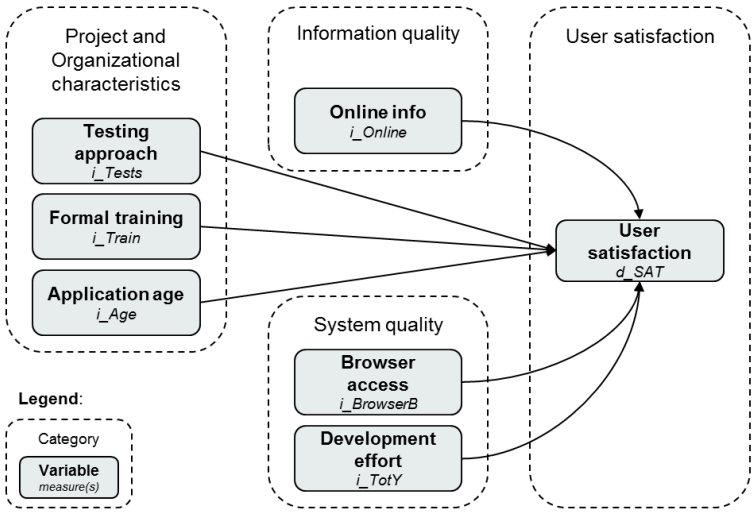


Figure 3 Variables used to operationalized the conceptual model

A number of linear regression models were then evaluated, taking into account the usual requirements on the data for this kind of analysis (Laureano 2011; Pestana et al. 2008):

Several regression models were evaluated, each focusing on testing the relationship between User Satisfaction and one or more of the groups of variables in the conceptual model. Due to the number of cases containing all the relevant variables, in some instances only a few variables from each group could be tested, since regression analysis requires at least 10-15 cases per dependent variable (Stevens 1996).

The actually calculated models are listed in Table **1**. Model #1 was created to test the influence in User Satisfaction of variables from the Project and Organizational (P&O) characteristics group, Model #2 analyzed the influence in User Satisfaction of Information Quality (IQ) and System Quality (SQ) and Model #3 looked after the influence in User Satisfaction of all three variable groups (P&O, IQ and SQ). If any non-significant variable was identified in any of the main models, a variant model, without that variable, was also created (#2b and #3b).

Table 1 Linear regression models

Model	**Dependent variable**	**Independent variables**
#1	Satisfaction (*d_SAT*)	Testing approach (*i_Tests*) Formal training (*i_Train*) Application age (*i_Age*)
#2	Satisfaction (*d_SAT*)	Online information (*i_Online*) Browser access (*i_BrowserB*) Development effort (*i_TotY*)
#2b	Satisfaction (*d_SAT*)	Online information (*i_Online*) Browser access (*i_BrowserB*)
#3	Satisfaction (*d_SAT*)	Application age (*i_Age*) Online information (*i_Online*) Browser access (*i_BrowserB*) Development effort (*i_TotY*)
#3b	Satisfaction (*d_SAT*)	Application age (*i_Age*) Online information (*i_Online*) Development effort (*i_TotY*)

After running each model, its significance was checked, as was the strength of its explanatory ability, then each variable's significance was checked as well and only then were the variables coefficients analyzed, in order to understand the strength and direction of the influence of each independent variable in the dependent variable.

3 Results

A correlation analysis performed between the dependent variable and each of the independent variables, summarized in Table **2**, provides a preliminary evaluation of the relationships between the variables and the selection of those that shall be included in the subsequent analysis.

Table 2 Analysis of correlation with Satisfaction – results and decisions

Concept	Variable code	N	Pearson Correlation	Sig.	Notes
Development effort	i_TotY	54	–0.126~	0.364	
Acceptance testing	i_Tests	57	0.213~	0.112	
Formal training	i_Train	57	–0.241+	0.071	
Browser-based access	i_BrowserA	57	–0.230+	0.086	Discarded
	i_BrowserB	57	–0.340**	0.010	
Online information update	i_Online	57	0.411**	0.001	
Application maturity	i_AgeA	57	0.376**	0.004	

Correlation significance level: *** < 0.001, ** < 0.01, * < 0.05, + < 0.1, ~ = not significant

It emerges that several variables are significantly correlated with user satisfaction: the type of access to the application, whether the information in the information system is updated online or not, how old the underlying systems are and, with a weaker significance, whether formal classroom-training took place.

Some other candidate independent variables proven not to be significantly correlated with the dependent variable but were still kept for subsequent analysis, to account for possible suppressor effects (Pandey et al. 2010).

The only discarded variable at this stage was one of the two alternate codings of the related to the type of user interface access provided to the users of the applications (*i_BrowserA* and *i_BrowserB*), since it became apparent from this analysis that the later, laxer approach, seemed to yield a stronger relationship with satisfaction of the user than the the former, more conservative approach.

The following step dealt with quantifying the eventual influence that each independent variable may have upon User Satisfaction, via linear multiple regression models.

The first model, Model #1, focused on variables in the conceptual model (cf. Figure **3**) which are associated with Project and Organizational characteristics. Its results are summarized in Table **3**.

Table 3 Results of Model 1 – Influence of Project and Organizational variables in Satisfaction

Model	Dep. variable	Independent variables	Standardized coefficients	Non-std. coefficients	T	Adj. R^2	F-test	Std. error
#1	User satisfaction (*d_SAT*)	Constant	–	3.190 (se .084)	37.843***	20.2%	F(3:57)= 5.729**	.148
		Testing Approach (*i_Tests*)	.207	.156 (se .092)	1.699+			
		Formal training (*i_Train*)	-.282	-.102 (se .044)	–2.336*			
		Application age (*i_Age*)	.353	.015 (se .005)	2.931**			

Correlation significance level: *** < 0.001, ** < 0.01, * < 0.05, + < 0.1, ~ = not significant
se = standard error

The model appears to have a good fit (F statistic is significant), i.e., it is better at explaining variance in User Satisfaction than pure guesses. In fact, the Adjusted R^2 shows 20.2% of the variance in user satisfaction is accounted for by that of the model's dependent variables.

All the independent variables are significant (i.e. their variance effectively predicts some of the variance in User Satisfaction). In particular, the coefficients show there is a +.156 increment of Satisfaction for each unit increment in the Testing Approach variable (Formal Training and Application Age remaining constant). Similarly, there is a +.015 increment of User Satisfaction for each additional year of Application Age, all other variables remaining constant. Finally, the independent variable Formal Training has a negative influence on User Satisfaction, which decreases by -.102 for each unit decrease of that independent variable, *et ceteris paribus*.

Application Age is the variable with the strongest influence in User Satisfaction, with a change of one standard deviation in this variable resulting in a change of +.353 standard deviations in User Satisfaction. The second strongest influencer is Formal Training (-.282 standard deviations), followed by Testing Approach (+.207 standard deviations).

A second model, Model #2, was developed to focus on the variables in the conceptual model (cf. Figure **3**) which are associated with Information Quality and System Quality. The results are summarized in Table 4.

This model appears to also have a good fit, accounting for 18.2% of the variance in User Satisfaction. However, one variable, Development Effort, is not significant, which renders the entire model not significant. As explained in section 3.7, in these cases, a variant model was also estimated, removing the non-significant variable. The results are shown as Model #2b in Table 4.

This variant model also has an even better fit, accounting for 23.2% of the variance in User Satisfaction. All its independent variables are significant, which means their variance effectively predicts some of the variance in User Satisfaction. There is a +.150 increment of User Satisfaction for each unit increment in the Online Information variable, Browser Access remaining the same. Conversely, the Browser Access variable has a negative influence on User Satisfaction, with each unit increment of the former leading to a change of -.101 in the later, *et ceteris paribus.*

Table 4 Results of Model 2 – Influence of Information and System Quality variables in Satisfaction

Model	Dep. variable	Independent variables	Standardized coefficients	Non-std. coefficients	T	Adj. R^2	F-test	Std. error
#2	User satisfaction (*d_SAT*)	Constant	–	3.348 (se .043)	77.808***	18.2%	F(3:47)= 4.405**	.150
		Online information (*i_Online*)	.366	.150 (se .055)	2.178**			
		Browser access (*i_BrowserB*)	-.294	-.098 (se .045)	–2.189*			
		Development effort (*i_TotY*)	-.128	-.004 (se .004)	– .946~			
#2b	User satisfaction (*d_SAT*)	Constant	–	3.348 (se .043)	96.597***	23.2%	F(2:57)= 9.474***	.146
		Online information (*i_Online*)	.381	.150 (se .046)	3.242**			
		Browser access (*i_BrowserB*)	-.303	-.101 (se .039)	–2.574*			

Correlation significance level: *** < 0.001, ** < 0.01, * < 0.05, + < 0.1, ~ = not significant
se = standard error

Online Information is the variable with the strongest influence in User Satisfaction, with a change of one standard deviation in this variable resulting in a change of +.381 standard deviations in User Satisfaction, whereas Formal Training accounts for -.303 standard deviations.

A third and final model, Model #3, was developed to include variables in the conceptual model (cf. Figure **3**) which belong to all main groups. Due to the previously mentioned limitation in the number of variables that can be included in linear regression models like these, only the Application Age variable from the first group was used, since Model #1 had shown it to be the strongest influencer of user satisfaction. The results are summarized in Table 5.

Table 5 Results of Model 3 – Influence of P&O, IQ and SQ variables in Satisfaction

Model	Dep. variable	Independent variables	Standardized coefficients	Non-std. coefficients	T	Adj. R^2	F-test	Std. error
#3	User satisfaction (*d_SAT*)	Constant	–	3.170 (se .046)	68.522***	51.5%	F(4:47)= 13.200***	.115
		Application age (*i_Age*)	.719	.029 (se .005)	5.524***			
		Online information (*i_Online*)	.456	.187 (se .043)	4.344***			
		Browser access (*i_BrowserB*)	-.068	-.023 (se .037)	-.613~			
		Development effort (*i_TotY*)	-.459	-.013 (se .004)	–3.819***			
#3b	User satisfaction (*d_SAT*)	Constant	–	3.150 (se .033)	94.819***	52.2%	F(3:47)= 17.731***	.115
		Application age (*i_Age*)	.748	.030 (se .005)	6.227***			
		Online information (*i_Online*)	.459	.188 (se .043)	4.409***			
		Development effort (*i_TotY*)	-.466	-.014 (se .003)	–3.924***			

Correlation significance level: *** < 0.001, ** < 0.01, * < 0.05, + < 0.1, ~ = not significant
se = standard error

The model appears to have a good fit, with an Adjusted R^2 showing that 51.5% of the variance in User Satisfaction can be accounted for by that of the dependent variables included in it. However, here too, one variable (Browser Access) is not significant, which renders the model itself not significant. Once more, a variant model was estimated, removing that variable. Its results are shown as Model #3b in Table 5.

This variant model also has a good fit, accounting for 52.2% of the variance in User Satisfaction. All its independent variables are significant (i.e. their variance effectively predicts more than half of the variance in User Satisfaction). In particular, there is a +.188 increment of User Satisfaction for each unit increment in the Online Information variable, Development Effort and Application Age remaining constant. Similarly, there is a +.030 increment of User Satisfaction for each additional year of Application Age, all other variables remaining constant. Finally, the independent variable Development Effort has a negative influence on User Satisfaction, which decreases by -.014 for each unit decrease of that independent variable, *et ceteris paribus*.

Application Age has the strongest influence in User Satisfaction, with a change of one standard deviation in this variable resulting in a change of +.748 standard deviations in User Satisfaction. The second strongest

influencer of User Satisfaction in this model is Development Effort (-.466 standard deviations), followed closely by Online Information (.459 standard deviations).

Summary of the results

In summary, as shown in Figure **4**, Model #1, which deals with project and organizational characteristics, suggests that, all other things remaining equal:

- The higher the age of the underlying systems, the higher the satisfaction that users derive from the system;
- Similarly, the higher the involvement of end-users in acceptance testing, the higher the user satisfaction that is ultimately derived from the system once it goes into production; and
- Conversely, as requirements for formal, class-based training increase, user satisfaction decreases and vice-versa.

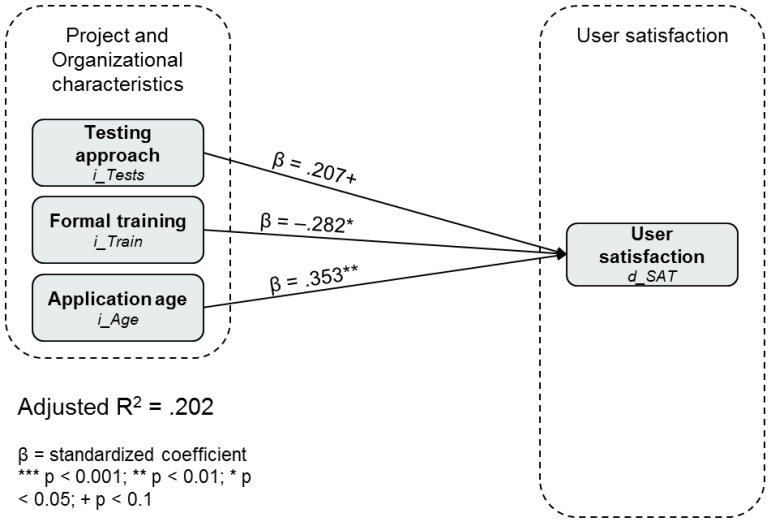


Figure 4 Multiple linear regression findings for Model #1

As summarized in Figure **5**, Model #2b, which looks at information and system quality, adds that, all else being equal:

- The higher the online availability of information within the system, the higher the user satisfaction; and
- On the other hand, the more access to the information system is achieved via a web browser instead of through specific client or terminal applications, the lower the user satisfaction.

Finally, as illustrated in Figure **6**, Model #3b, reinforces the above notions for application age and online information and adds that, *et ceteris paribus*:

- The information systems with a higher total effort employed in their development and maintenance tend to have lower user satisfaction

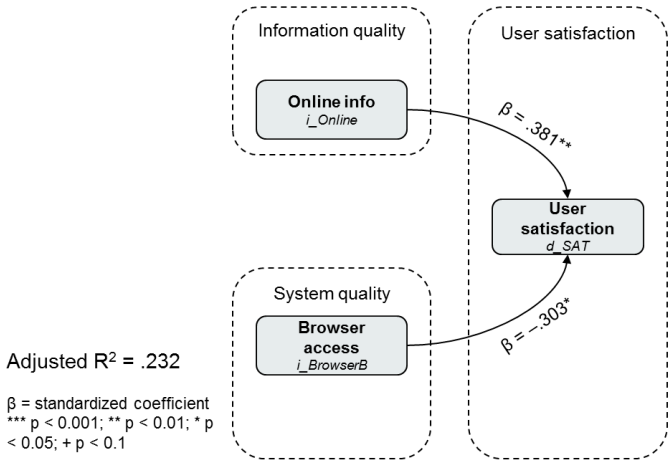


Figure 5 Multiple linear regression findings for Model #2b

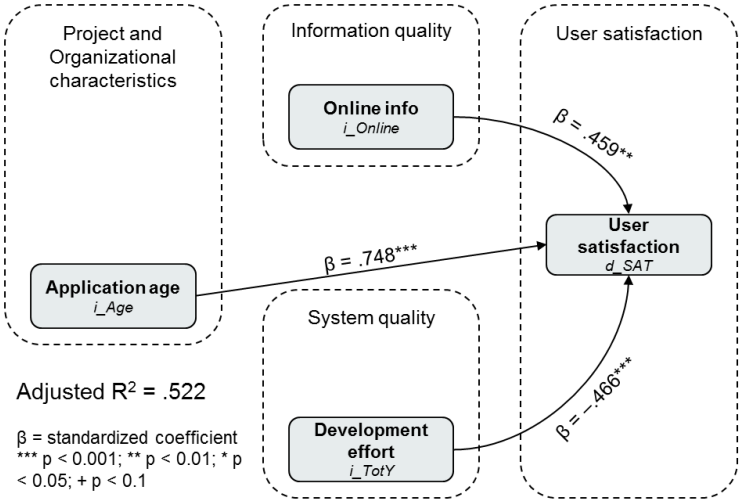


Figure 6 Multiple linear regression findings for Model #3b

4 Discussion

This study aimed to deepen the understanding of the antecedents of user satisfaction with information systems, a proxy for IS success, particularly focusing on the characteristics of said information systems and their development processes that may have influence in user satisfaction. The previously presented results will now discussed in light of those objectives, grouped according to the organization of the theoretical and conceptual research models in Figure 1 and Figure 2.

Project and organizational characteristics as antecedents of user satisfaction

Within the scope of this case study, the most relevant project/organizational characteristic that appears to be a strong predictor of user satisfaction is age (years since earliest underlying system of an application was introduced at the Bank). If there is a causal relationship between the two, it is likely to be in the sense that age influences satisfaction (since age, by its very nature, cannot be influenced by another variable), consistent with the findings of previous research related to the learning curve and the maturity effect in the life-cycle of most information systems (Sedera et al. 2004).

Another positive influencer of user satisfaction in this group is one of the aspects related to involvement of the users in the development of information systems, the testing approach. This corroborates the notion that the more end-users are involved at all stages of an information system's development, the more likely is the success of that information system, measured by the user's satisfaction, also consistent with prior research (Guimaraes et al. 1996).

The other aspect of end-user involvement, which deals with the enrollment of the users in formal, class-based training prior to usage of the system commences, has proven to be negatively correlated with user satisfaction. While one would intuitively expect more training to lead to higher user satisfaction (Guimaraes et al. 1996), there is also considerable support for the opposite in the relevant literature, mostly due to it the being contingent on factors like technical complexity and task inter-dependency (Sharma et al. 2007).

Technology: information and system quality measures as antecedents of user satisfaction

The measure of Information Quality employed in this research, Online Information (provision by an information system of access to online, up-to-date information), has been shown by more than one of the regression models to positively influence user satisfaction. This relationship is in line with several previously published research results which have shown that information quality is a key antecedent of user satisfaction and IS success (DeLone et al. 2003; Gable et al. 2008; Gorla et al. 2010).

Conversely, both System Quality measures have proven to be negative predictors of user satisfaction in this case study, all other parameters being equal.

The first of those is the provision of a user interface via a web browser, which may lead to positing that the other user interfaces are preferred by users to the detriment of those designed to be accessed via a web browser. While interpreting these results, one has to bear in mind that the negative nature of this relationship is solely due to the arbitrary choice of coding of this characteristic of the information systems in a dummy variable. Had the choice been to code, not the characteristic of "Browser-based access" but

rather the characteristic of "Dedicated-client access", then the influence of the variable would have resulted in the opposite sign (i.e., positive).

Regarding the second measure, Development effort (total number of person-years employed in development and maintenance of the system), it may seem counter-intuitive that it is related to lower user satisfaction. Nevertheless, it may be that the causal relationship is in the inverse sense, i.e., that applications for which user satisfaction is low tend to have more development and maintenance work as an attempt to solve the issues that are leading to lower user satisfaction. Some anecdotal cases (e.g. New Local System) concur to this idea, since performance issues led to a significant effort overrun in an attempt to solve them.

Implications for research

The reported findings support the notion that user satisfaction in this case study is a proxy of Information Systems success, since it was shown to be significantly influenced by several variables which can be seen as its antecedents. This is consistent with previous research (Iivari 2005; Petter et al. 2008).

Among those antecedents, the maturity of the information system stands out as a relevant predictor of user satisfaction, as does the online nature of the information within it. These conclusions lead to the proposal of two hypotheses as suggestions for future research.

> Hypothesis 1 (H1): The maturity of an information system positively influences its users' satisfaction.
>
> Hypothesis 2 (H2): Users exhibit more satisfaction towards online updated information systems than towards those with batch-based update of information.

Conversely, support was found for the idea that training is not in itself a stand-alone measure for increasing user-satisfaction. This suggests that understanding of the moderators of training's influence on IS success in the same or a similar scope as that of this case study would provide an interesting research path.

Finally, evidence was found that, at least in this case study, browser-based user interfaces are generally associated with lower user satisfaction. This could be a manifestation of the effect of some of the shortcomings of this kind of user interfaces (e.g. response times) which have been shown to decrease user satisfaction (Hoxmeier et al. 2000), and which could perhaps be confirmed by further research.

Implications for practice

The most natural application of this study's conclusions to the practice of Information Systems management is by the actual Bank which was studied, given that, even if the results cannot be generalized, they are, by their nature, fully applicable to the subject.

It is apparent from the conclusions that there may be a case to be made for decreasing the rate of introduction of new systems at the Bank, increasing the existing trend to have online information systems and dedicating more resources to the crafting of the applications in a way that less formal user training is required.

Further, the Bank could look to eliciting more detailed information on the nature of the effort employed in the development and maintenance of its information systems, since this may prove useful to understand the direction of an eventual causal inverse relationship between person-years spent on an application and user satisfaction.

These conclusions may also apply to other similar organizations, since many of them are supportive of previous research and general practice, but naturally careful examination of the context would be required *a priori*.

Limitations

The use of the Bank's user satisfaction surveys as a source for the user satisfaction data, even if a comprehensive and invaluable one, instead of an instrument that was purposely designed to support this research, has limited the ability to further explore other dimensions of the users' perceptions in addition to their effective satisfaction with the Bank's information systems.

In addition, the quantitative nature of the analysis techniques in the research design limited the ability to explore in-depth some subjects which require a more subjective approach or that cannot be structured to fit the dependent variable as operationalized here.

5 References

Al-Maskari, A., and Sanderson, M. 2010. "A review of factors influencing user satisfaction in information retrieval," Journal of the American Society for Information Science and Technology (61:5), pp 859-868.

Bailey, J. E., and Pearson, S. W. 1983. "Development of a Tool for Measuring and Analyzing Computer User Satisfaction," Management Science (29:5) May, pp 530-545.

Bharadwaj, A. S. 2000. "A resource-based perspective on information technology capability and firm performance: an empirical investigation," MIS quarterly), pp 169-196.

Botelho, M. d. C., Calapez, T., and Ramos, M. 2011. "The effect of shape: comparing different presentations of response," Proc. 58th World Statistical Congress).

Carifio, J., and Perla, R. J. 2007. "Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes," Journal of Social Sciences (3:3), pp 106-116.

DeLone, W. H., and McLean, E. R. 1992. "Information Systems Success: The Quest for the Dependent Variable," Information Systems Research (3:1), pp 60-95.

DeLone, W. H., and McLean, E. R. 2003. "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," Journal of Management Information Systems (19:4), pp 9-30.

Doll, W. J., and Torkzadeh, G. 1988. "The Measurement of End-user Computing Satisfaction," MIS quarterly (12:2), pp 259-274.

Farrell, D., Mendonça, L., Nevens, M., Manyica, J., Lal, S., Roberts, R., and others 2002. "How IT Enables Productivity Growth - The US experience across three sectors in the 1990s," McKinsey Global Institute, San Francisco.

Gable, G. G., Sedera, D., and Chan, T. 2008. "Re-conceptualizing information system success: the IS-impact measurement model," Journal of the Association for Information Systems (9:7), pp 377-408.

Gopalan, S., Jain, G., Kalani, G., and Tan, J. 2012. "Breakthrough IT banking," in McKinsey on Business Technology, M. a. Company (ed.), McKinsey & Company: India.

Gorla, N., Somers, T. M., and Wong, B. 2010. "Organizational impact of system quality, information quality, and service quality," The Journal of Strategic Information Systems (19:3), pp 207-228.

Guimaraes, T., Yoon, Y., and Clevenson, A. 1996. "Factors important to expert systems success a field test," Information & Management (30:3), pp 119-130.

Göb, R., McCollin, C., and Ramalhoto, M. F. 2007. "Ordinal methodology in the analysis of Likert scales," Quality & Quantity (41:5), pp 601-626.

Hayes, B. E. 1998. Measuring Customer Satisfaction: Survey Design, Use, and Statistical Analysis Methods, (ASQ Quality Press.

Heidmann, M. 2010. "Overhauling banks' IT systems," in McKinsey Quarterly, McKinsey & Company: Frankfurt, Germany.

Hoxmeier, J. A., and DiCesare, C. Year. "System response time and user satisfaction: An experimental study of browser-based applications," Proceedings of the Association of Information Systems Americas Conference, Citeseer2000, pp. 10-13.

Iivari, J. 2005. "An empirical test of the DeLone-McLean model of information system success," SIGMIS Database (36:2), pp 8-27.

Laureano, R. M. S. 2011. Teste de hipóteses com o SPSS: O meu manual de consulta rápida, (Edições Sílabo: Lisbon, Portugal.

Lin, B.-W. 2007. "Information technology capability and value creation: Evidence from the US banking industry," Technology in Society (29:1) 1//, pp 93-106.

Melone, N. P. 1990. "A theoretical Assessment of the User-Satisfaction Construct in Information Systems Research," Management Science (36:1), pp 76-91.

Pandey, S., and Elliott, W. 2010. "Suppressor variables in social work research: Ways to identify in multiple regression models," Journal of the Society for Social Work and Research (1:1), pp 28-40.

Pestana, M. H., and Gageiro, J. N. 2008. Análise de dados para ciências sociais: a complementaridade do SPSS, (5th edition ed.) Edições Sílabo: Lisbon, Portugal.

Petter, S., DeLone, W., and McLean, E. 2008. "Measuring information systems success: models, dimensions, measures, and interrelationships," European Journal of Information Systems (17:3), pp 236-263.

Petter, S., DeLone, W., and McLean, E. R. 2013. "Information Systems Success: The Quest for the Independent Variables," Journal of Management Information Systems (29:4), pp 7-62.

Sedera, D., and Gable, G. G. Year. "A factor and structural equation analysis of the enterprise systems success measurement model," Association for Information Systems2004.

Sedera, D., and Tan, F. Year. "User Satisfaction: An Overarching Measure of Enterprise System Success," Pacific Asia Conference on Information Systems, Association for Information Systems, Bangkok, Thailand, 2005, pp. 963-976.

Sharma, R., and Yetton, P. 2007. "The contingent effects of training, technical complexity, and task interdependence on successful information systems implementation," Mis Quarterly), pp 219-238.

Stevens, J. 1996. Applied multivariate statistics for the social sciences, (Lawrence Erlbaum: Mahwah, New Jersey.

Medição de usabilidade e eficiência de dashboards: framework desenvolvida usando os princípios de desenho de Business Intelligence e Interação Homem-Máquina

Miguel Oliveira[1], Elsa Cardoso[2], Marina Santana[3]

1) ISCTE – Instituto Universitário de Lisboa, Portugal

mhdoa@iscte.pt

2) ISCTE – Instituto Universitário de Lisboa e INESC-ID, Portugal

elsa.cardoso@iscte.pt

3) Novabase Business Solutions, Portugal

marina.santana@novabase.pt

Resumo

A usabilidade e eficiência no processo de decisão são aspetos essenciais em interfaces de *Business Intelligence* (BI). Algumas organizações afirmam construir *dashboards* com estas características, mas sem forma efetiva de comprová-lo. Sendo uma temática pouco investigada em BI, é contudo largamente estudada na área científica de Interação Homem-Máquina (IHM). Este artigo apresenta o projeto de desenho e desenvolvimento de uma *framework* para medição da usabilidade e eficiência de *dashboards* utilizando os princípios de desenho de BI e IHM, com adaptação da ISO/IEC 9241-11. Este projeto académico foi realizado na empresa Portuguesa Novabase, na competência de *Business Analytics* (BA). Para a validação da *framework* recorreu-se a profissionais da Novabase nas duas áreas, a professores de IHM e a estudantes com conhecimentos nas duas áreas, observando-se diferenças significativas de usabilidade consoante o número de erros introduzidos e tipo de avaliadores.

Palavras chave: *business intelligence*, usabilidade, eficiência, *dashboards*

1. Introdução

Os *dashboards* são atualmente uma ferramenta indispensável, sendo utilizadas pelos gestores para monitorizar a performance e decidir o rumo das organizações. Um *dashboard* é uma aplicação de *Business Intelligence* (BI), consolidada num único écran, que permite a visualização rápida da informação mais relevante para atingir um ou mais objectivos [Few 2013]. Um *dashboard* é uma ferramenta de comunicação que permite retirar conhecimento e conclusões de uma forma eficiente, apenas com um simples olhar.

Inicialmente os *dashboards* destinavam-se exclusivamente a pessoas com funções de chefia. Mas hoje em dia, os *dashboards* – devido às suas vantagens – são utilizados em vários níveis, permitindo coordenar o esforço dos colaboradores no cumprimento de objetivos de gestão na ótica da visão da organização [Eckerson 2010]. Contudo, nem sempre estes resultados são alcançados por falta de usabilidade e eficiência de algumas interfaces. Às vezes, os utilizadores rejeitam-nas por não cumprirem os requisitos ou por não serem fáceis de utilizar [Fonseca et al. 2012] e [White 2005], com consequente perda de valor dos projetos [Rudin 2007], e perda de eficiência na comunicação e desempenho dos negócios [Few 2006]. Todavia, seguindo as heurísticas de desenho, é possível desenvolver interfaces com maior usabilidade [Few 2006], [Fonseca et al. 2012] e [Nielsen 1993]. A área de Interação Homem-Máquina (IHM) estuda a usabilidade e eficiência de interfaces, podendo-se adaptar metodologias e métodos de avaliação para BI, identificando problemas existentes e possíveis soluções.

Durante este estudo questionou-se "Como medir a usabilidade e a eficiência de uma interface em *Business Intelligence*?" Desenvolveu-se uma *framework* adaptada à ISO/IEC 9241-11 para medir a usabilidade de um *dashboard*, utilizando heurísticas de desenho, aplicável em contexto organizacional, pretendendo-se que seja fácil e prática. Pretende-se testar quatro hipóteses relacionadas com a análise da usabilidade percepcionada em função da existência e do tipo de erros, do tipo de avaliador e do tipo de boas práticas.

Este artigo está estruturado em seis secções. Na secção 2 discutem-se as boas práticas de avaliação de usabilidade em BI e IHM. A *framework* desenvolvida para medir a usabilidade e eficiência de *dashboards* é apresentada na secção 3. O processo de validação da *framework* e a análise dos resultados são apresentados nas secções 4 e 5, respectivamente. Por fim, na secção 6 apresentam-se as conclusões e reflete-se sobre o trabalho futuro.

2. Avaliação de usabilidade

A definição de usabilidade não é consensual na literatura, como é demonstrado na Tabela 1. A usabilidade depende do propósito e dos requisitos dos *stakeholders*. Contudo, produtos com usabilidade permitem poupar dinheiro, aumentar a qualidade e a satisfação do cliente [Fonseca et al. 2012]. A usabilidade deve ser considerada durante o processo de desenvolvimento, requerendo o envolvimento e participação dos *stakeholders* em todas as fases.

Boas práticas de desenho de *dashboards* em BI

Few [2010] afirma que o design do *dashboard* não é o único factor a considerar para este ser aceite. Mesmo que o *dashboard* seja concebido usando as boas práticas de desenho, se não

satisfizer os requisitos pré-estabelecidos, o cliente rejeita-o. Few [2006] refere que um bom design destaca a informação essencial para a monitorização (*data-pixels*), retira a ênfase aos elementos auxiliares para mostrar a informação (*non-data pixels*) e possui simplicidade, mostrando-se claro, organizado e simples.

	BI	**IHM**		
	Few 2006	**ISO/IEC 9241-11 1998**	**Nielsen 1993**	**Norman 1988**
O que é usabilidade?	Agradabilidade estética; deteção de informação fora da norma; facilidade de uso; simplicidade no desenho; suportar as tarefas do utilizador	Eficácia; eficiência; satisfação	Eficiência; erros; facilidade de aprendizagem, de lembrança; satisfação	Facilidade de aprendizagem; eficácia no cumprimento do requisito (intenção vs. ação; perceção vs. interpretação)
Como atingi-la / medi-la?	Dar sentido e uso efetivo à informação; garantir agradabilidade estética; manter a consistência para uma melhor interpretação; dar uma visão global; testar o desenho com os utilizadores	Definição dos requisitos de usabilidade, indicadores para medir cada critério, metas para cumprir os indicadores e avaliar o desempenho face à meta	Definição dos requisitos de usabilidade, indicadores para medir cada critério, metas para cumprir os indicadores e avaliar o desempenho face à meta	Definição dum bom modelo concetual e uma filosofia baseada no utilizador para reduzir as discrepâncias entre os ciclos

Tabela 1 – Comparação usabilidade BI vs. IHM

Few [2006] identifica seis fatores que conjugados conferem usabilidade a um *dashboard*: (1) deve ser agradável ao utilizador, (2) deve promover a simplicidade na comunicação e exposição da informação; (3) deve ser fácil de usar, recorrendo a meios conhecidos pelo utilizador, (4) garantindo a consistência dos mesmos, exceto quando se pretende destacar certa informação; (5) deve fornecer informação necessária para detetar fenómenos fora da norma, auxiliando a tarefa do utilizador; e (6) o desenho deve ser mostrado aos utilizadores para ser aceite e validado.

Boas práticas de desenho de interfaces em IHM

Nielsen [1994] refere existirem diversos conjuntos de heurísticas utilizáveis no desenho de interfaces, por exemplo, os princípios de Norman [1988] e as heurísticas de Molich e Nielsen [1990]. Todavia, segundo Nielsen [1993] alguns conjuntos são apenas formalismos ou demasiado específicos e complexos ou extensivos para ser usados.

Fonseca et al. [2012] e Nielsen [1993] afirmam que uma boa interface possui duas características: eficácia no cumprimento das tarefas pré-definidas e atratividade para o utilizador, sublinhando que os elementos exibidos devem ser claros e explícitos ("*less is more*"), e a interface, eficiente e flexível para os utilizadores recorrentes.

Para agilizar e simplificar o desenvolvimento, Nielsen [1993] recomenda o uso de protótipos facilitadores da criação do modelo conceptual[1] e da obtenção de *feedback* em fases iniciais.

A norma ISO/IEC 9241-11 [1998] define usabilidade através de três dimensões: eficácia (exatidão e plenitude), eficiência (recursos necessários) e satisfação (atitudes e conforto). Utilizando uma *framework* composta por três elementos, avalia-a considerando o cumprimento das metas estabelecidas pelos requisitos. Consequentemente, identifica-se o contexto de uso e o produto tendo em conta o utilizador, o equipamento e a envolvência do sistema no meio. Deve ainda detalhar-se o resultado esperado, global ou especifico (funcionalidade/tarefa), bem como escolher por cada requisito, pelo menos, três indicadores operacionais, um por dimensão, e o seu valor mínimo aceitável [Jokela et al. 2003].

Nielsen [1993] considera que usabilidade é uma componente da aceitação do sistema, medindo o grau de cumprimento dos requisitos estipulados pelos *stakeholders*, e a usabilidade, o desempenho dos mesmos. O autor propõe medir a usabilidade através de cinco fatores: (1) eficiência (produtividade); (2) capacidade de prevenção de erros (entre os requisitos e as ações tomadas); (3) aprendizagem (desempenho numa tarefa, utilizador inexperiente vs. experiente); (4) memorização (facilidade de reutilização sem reaprendizagem integral); e (5) satisfação (agradabilidade).

Norman [2002] decompõe usabilidade em facilidade de uso e eficácia no cumprimento da meta. O autor através de um modelo, as setes fases da ação, explica o funcionamento das ações dos seres humanos, recorrendo a dois ciclos interligados (execução e avaliação). Identificado o requisito, surge a necessidade de o satisfazer, iniciando-se, então, o ciclo de execução onde são planeadas as ações necessárias e a sua execução. Por oposição, o ciclo de avaliação inicia-se com a perceção de algo no meio que é interpretado e avaliado, daí resultando novos requisitos a serem satisfeitos. Assim, a falta de usabilidade corresponde a discrepâncias nos ciclos, i.e., a intenção de agir, diverge da ação tomada e a perceção do meio difere da sua avaliação.

[1] Modelo concetual ou de designer corresponde à compreensão que este possui do sistema, da sua organização e função [Fonseca et al. 2012] e [Johnson e Henderson 2002].

Outros estudos

Embora existam estudos que medem a usabilidade de uma interface ou avaliam a aplicação das boas práticas num *dashboard*, não foi encontrado nenhum estudo aplicado a BI que junte as duas componentes (i.e., BI e IHM). Green e Pearson [2006] e Joo, Lin, e Lu [2011] adaptaram a ISO/IEC 9241-11 para medir a usabilidade, respectivamente, de um *website* de comércio electrónico e de uma biblioteca, recorrendo a utilizadores que avaliaram subjetivamente o cumprimento dos requisitos. Lourenço [2013] desenvolveu uma *framework* que analisa o uso dos princípios de perceção visual no desenho de *dashboards* através de um conjunto de heurísticas.

3. *Framework* para medição da usabilidade e eficiência de *dashboards*

A *framework* para medição da usabilidade e eficiência de *dashboards* (com o acrónimo FMUED) foi desenvolvida no âmbito de uma dissertação de mestrado em Informática e Gestão do ISCTE – Instituto Universitário de Lisboa (ISCTE-IUL), em contexto empresarial na Novabase. O trabalho foi supervisionado por um *steering committee* composto por um orientador académico e dois consultores seniores de BI da Novabase. Com este trabalho, a empresa pretende disseminar internamente as boas práticas de desenho de *dashboards* a todos os colaboradores envolvidos no desenho de interfaces de BI, esperando obter no futuro uma solução que possa ser usada em projetos com clientes. Investigadas as soluções para medir a usabilidade de interfaces, decidiu-se adaptar a ISO/IEC 9241-11 para *dashboards*, dada a sua aceitação pela comunidade. Conceptualmente, a *framework* desenvolvida possui duas componentes: as heurísticas de desenho (organizadas em sete categorias) e o método de avaliação (ver Figura 1).

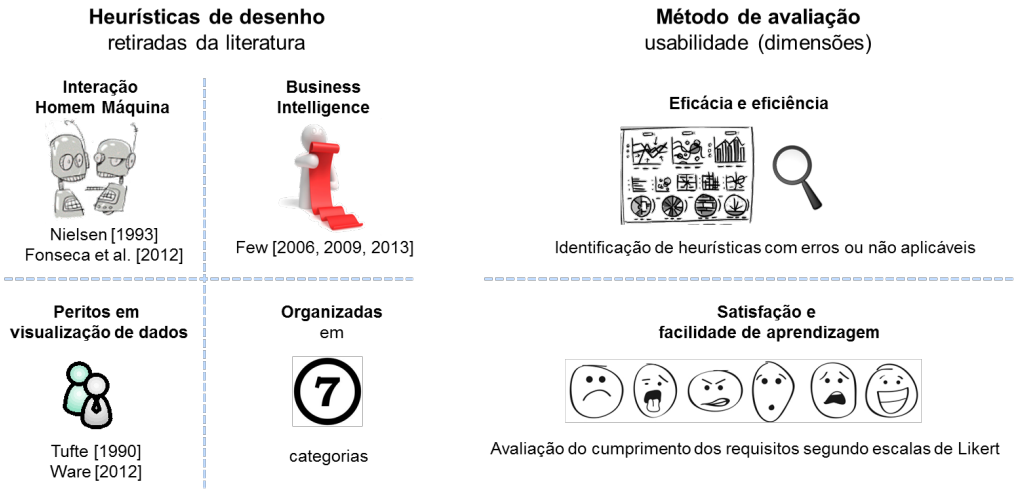


Figura 1 – Modelo conceptual da *framework* FMUED

Few [2013] refere que a definição do propósito do *dashboard* reduz ambiguidades e futuros problemas. Incorporou-se este princípio, na aceção definida por Lakatos e Marconi [2003], i.e., o propósito deve ser claro, não ambíguo e conciso na especificação das suas capacidades, praticável para ser implementado e relevante para o utilizador.

Embora a norma delimite o contexto de uso do produto/interface, este não está ajustado para *dashboards*. O designer deve assim considerar os seguintes aspetos: utilizadores (proficiência em visualização de dados, experiência em *dashboards*, etc.); tarefa (propósito do *dashboard*); equipamento (*software* e *hardware*) e meio (frequência de atualização, departamento do utilizador, ligação a outras ferramentas).

A ISO/IEC 9241-11 operacionaliza os requisitos através da definição de indicadores. Em *dashboards*, o requisito é considerado completo quando é observável ou aceite pelo cliente. Assim, decidiu-se remover a obrigatoriedade de determinar o tipo de requisito, os indicadores subjacentes e as metas para validá-lo, bastando definir os requisitos e categorizá-los de acordo com as dimensões de usabilidade da *framework* Uma simplificação semelhante foi feita por Green e Pearson [2006] e Joo et al. [2011] para facilitar a avaliação.

Segundo a norma, usabilidade é composta por eficácia, eficiência e satisfação. Porém, a facilidade de aprendizagem do *dashboard* é também um fator relevante [Few 2006]. Consequentemente, adotou-se a facilidade de aprendizagem definida por Nielsen [1993], como uma dimensão para definir usabilidade. Sendo assim, as dimensões de usabilidade consideradas na *framework* FMUED são: **eficiência** (produtividade e possíveis análises); **eficácia** (funcionalidades, extração e qualidade da informação); **satisfação** (agradabilidade ao utilizador) e **facilidade de aprendizagem** (capacidade de compreensão).

Os requisitos pertencentes às dimensões de eficácia e eficiência devem ter representação direta e delimitada no *dashboard* (e.g., zonas), contrastando com os outros requisitos que são avaliados globalmente. Para facilitar a extração dos requisitos e aplicar uma abordagem que envolva os utilizadores no design, devem ser usados mock-ups no desenvolvimento de protótipos do *dashboard*.

Embora haja outros métodos para avaliar o desenho da interface, segundo Nielsen [1993] a avaliação por heurísticas é uma solução que requer baixos recursos e permite obter resultados rapidamente. Contudo, exige que os avaliadores tenham alguns conhecimentos em boas práticas de desenho. A *framework* FMUED deve pois ser utilizada por especialistas no desenvolvimento de interfaces de BI.

Categorias	Tipo	Heurísticas
C1 - Considere quais as cores que escolhe	G	Definição de 2 paletes (standard e ênfase)
	G	Número de cores escolhidas (≤ 10)
	G	Uso de cores na natureza ou na roda das cores
	G	Consideração pelo daltonismo, cultura e contexto da cor
	S	Escolha de cores no fundo (baças, não brilhantes e sem gradientes)
C2 - Considere a tipografia que usa	G	Tipo de letra legível
	S	Uso preferencial de texto na horizontal
	S	Não usar todo o texto em letra maiúscula
	S	Uso apenas de dois tipos de letra distintos
C3 - Destaque a informação importante e promova a simplicidade	S	Uso de formas de destaque (cor, forma, movimento, etc.)
	G	Definição de zonas de ênfase no *dashboard*
	G	Organização da informação segundo um processo, desejo do utilizador ou lógica
	S	Colocar casas decimais ou unidades apenas se necessário
	G	Uso de alertas para destacar informação
C4 - Procure a consistência e considere perícia do utilizador	G	Repetição de cores, gráficos, etc.
	S	Utilização do tipo de gráfico mais correto para mostrar a informação
	S	Quebra da consistência apenas com o propósito de destacar
	S	Tipo de análise de acordo com a perícia do utilizador
	G	Permitir o *drill-down* / filtrar data / selecionar específicos *datasets*
C5 - Forneça feedback do sistema e use uma linguagem compreensível pelo utilizador	G	Uso de uma linguagem simples, compreensível e simpática
	G	Fornecer resposta face às ações do utilizador
	G	Mostrar o estado do *dashboard* (última atualização, online ou *frozen*)
	S	Desativar opções após uma ação (sombrear botões, usar opções *drop-down*)
	G	Existência de um botão de ajuda visível
C6 - Reduza o ruído visual ao não enfatizar sem necessidade	S	Limitar o número de *datasets* por gráfico
	S	Utilizar gráficos em 2-D
	S	Uso de linhas com grossura mínima e com cores clara e não saturadas
	S	Uso preferencialmente do espaço em branco
	G	Existência apenas de decorações ou embelezamentos necessários
C7 - Promova o reconhecimento em vez da lembrança	G	Correta definição e delimitação de blocos de informação
	G	Consegue-se ver o *dashboard* sem utilizar o *scroll*
	G	Uso de menus ou ícones para representar ações

Tabela 2 - Heurísticas utilizadas no estudo

Heurísticas de desenho

A *framework* FMUED inclui 32 boas práticas de desenho de interfaces (ver Tabela 2), classificadas em sete categorias (C1 a C7), de forma a facilitar a sua compreensão e evidenciar as áreas de análise. As boas práticas foram compiladas a partir da literatura existente nas áreas de desenho de *dashboards* em BI [Few 2006, 2009 e 2013], desenho de interfaces em IHM [Fonseca et al. 2012] e [Nielsen 1993], e visualização de dados [Tufte 1990] e [Ware 2012]. Para cada heurística foi definido o seu tipo, explicitando o nível da avaliação: global (validação do cumprimento em todo o *dashboard*) ou específico (em uma ou mais zonas do *dashboard*).

Método de avaliação

A ISO/IEC 9241-11 recomenda a avaliação da performance dos indicadores face ao *target*, sem contudo especificar um método de avaliação. A *framework* FMUED requer a avaliação dos requisitos explícitos através da identificação de erros nas heurísticas e o uso de escalas de *likert* para avaliar o cumprimento dos requisitos de facilidade de aprendizagem e satisfação.

Quando é identificado um erro numa boa prática do tipo específico, o avaliador deve indicar as zonas do *dashboard* onde foi mal aplicada. Se o erro for numa boa prática do tipo global, apenas deve assinalá-la. Nas situações em que não há informação suficiente para a avaliação ou quando a boa prática não se aplica ao *dashboard*, deve ser apenas assinalada como não aplicável. A usabilidade na FMUED é medida utilizando as fórmulas da Tabela 3, retirando-se previamente as boas práticas não aplicáveis.

Fórmulas de cálculo Usabilidade

$$Usabilidade\ Total = \frac{\sum Usabilidade\ por\ dimensão}{Número\ de\ dimensões}$$

$$Usabilidade\ Dimensões\ Eficácia\ e\ Eficiencia = \frac{Usabilidade\ por\ requisíto\ explícito}{Número\ de\ requisitos}$$

$$Usabilidade\ por\ Requisíto\ explicíto = \frac{\sum Usabilidade\ por\ heurística}{Nº\ de\ heurísticas\ específicas + Nº\ de\ heurísticas\ globais * NIZ}$$

$$Usabilidade\ por\ Heurística = \begin{cases} Global = NIZ \\ Específica = \frac{NIZ\ com\ erro}{NZI} \end{cases}$$

$$Usabilidade\ Dimensões\ Facilidade\ de\ aprendizagem\ e\ Satisfação = \frac{\sum valor\ da\ escala\ de\ likert}{Valor\ máximo\ da\ escala * Número\ de\ requisitos}$$

Tabela 3 – Fórmulas de cálculo da usabilidade

Legenda: NZI = Número de zonas identificadas

4. Validação da *framework*

Para observar o desempenho do método de avaliação e a aceitação das heurísticas, desenhou-se uma experiência aplicada a um caso de estudo, onde para além do *dashboard* original foram desenhadas três versões suplementares com erros introduzidos. Os erros foram selecionados a partir das heurísticas (Tabela 2), variando apenas nas categorias e no número de erros introduzidos, resultando dois *dashboards* com cinco e um com sete erros, culminando em três situações de avaliação: zero, cinco e sete erros.

Caso de estudo

Optou-se por utilizar no processo de validação da FMUED num caso real, desenvolvido num projeto de Energia da Novabase, no controlo de perdas energéticas da organização que fornece energia em Cabo Verde. O objetivo do *dashboard* é a monitorização das perdas energéticas nas diversas ilhas.

A Figura 2 ilustra a distribuição energética, existindo uma central de produção (P), que produz e distribui pelos postos de transformação (PT) nas diversas ilhas, em média tensão (M.T.), sendo convertida em baixa tensão (B.T.), para abastecer a iluminação pública e os clientes. Contudo, durante a distribuição podem ocorrer dois tipos de perdas: (1) as de origem técnica, dispersão por transferência na rede (11%), e.g., P para o PT; (2) as perdas não técnicas/fraude, devido a ligações à rede, não permitidas, para obtenção de energia sem custo. Assim, fluxos não previstos originam interrupções/sobrecargas na rede resultando na quebra de abastecimento a alguns sectores, afetando a economia, a qualidade de vida dos utentes e a rentabilidade da organização. Por existir exclusividade na distribuição e a resolução ser morosa, os clientes ficam sem energia até à reposição. Através deste *dashboard* pretende-se uma melhor monitorização destas perdas para definição de medidas de contingência de modo a aumentar a qualidade de serviço e a satisfação dos clientes.

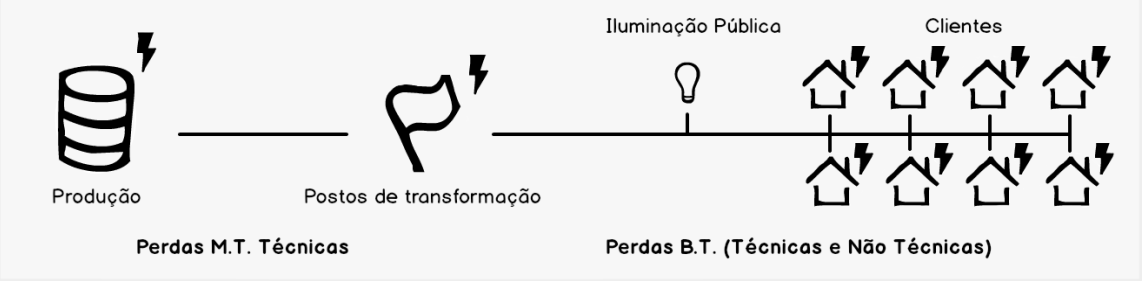


Figura 2 – Ilustração do caso de estudo

Com o intuito de identificar os requisitos, utilizar uma abordagem centrada no utilizador e averiguar o aspeto final do *dashboard* aplicou-se a técnica de *mock-ups*, (ver Figura 3). O *dashboard* final (ver Figura 4) foi implementado usando a ferramenta Microstrategy©.

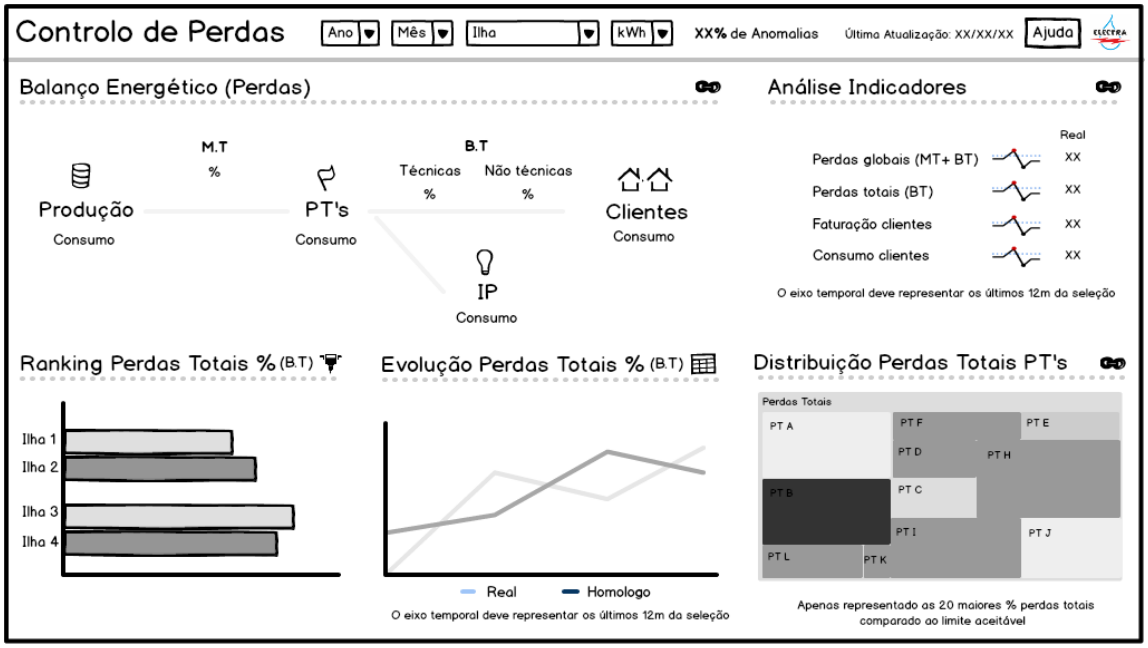


Figura 3 – *Mock-up* do caso de estudo, utilizando Balsamiq Mockups©

Protótipo e questionário

Para facilitar a avaliação da *framework* FMUED foi construído um protótipo em Microsoft Excel©, que inclui: o *dashboard* em avaliação; um temporizador para registar o tempo da avaliação; as boas práticas de desenho; um gráfico com os resultados; e três botões, para observar os requisitos, avaliar os requisitos de facilidade de aprendizagem e satisfação e visualizar um resumo dos erros identificados. A Figura 5 apresenta um exemplo de uma avaliação real. As boas práticas sublinhadas a vermelho simbolizam erros que foram identificados pelo avaliador, afetando a usabilidade do *dashboard*.

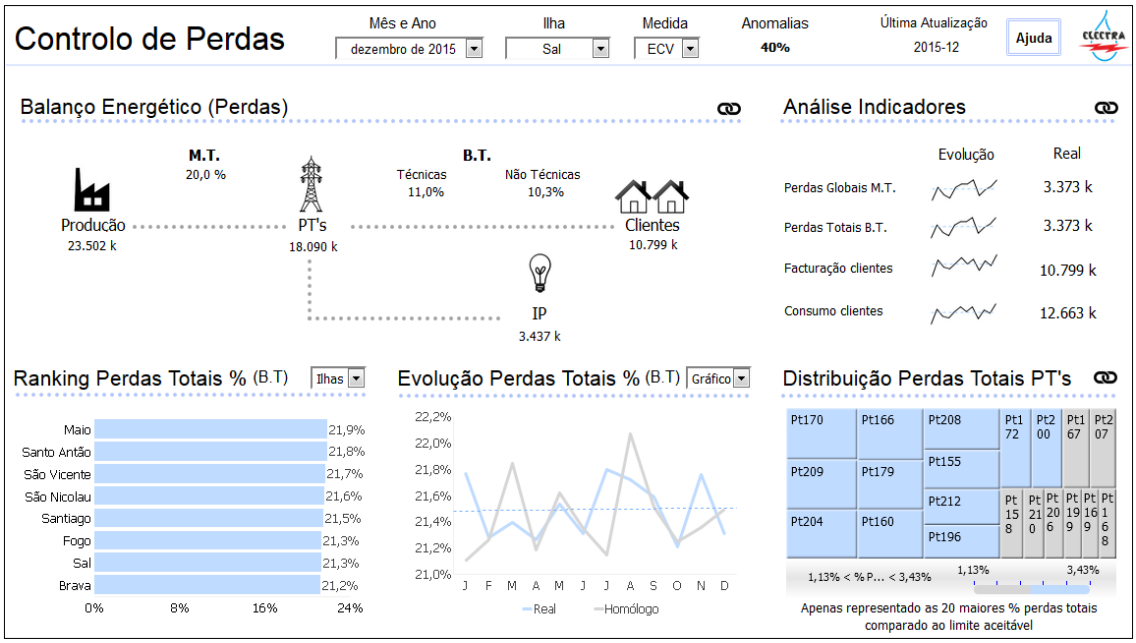


Figura 4 – Implementação do *mock-up*, utilizando Microstrategy©

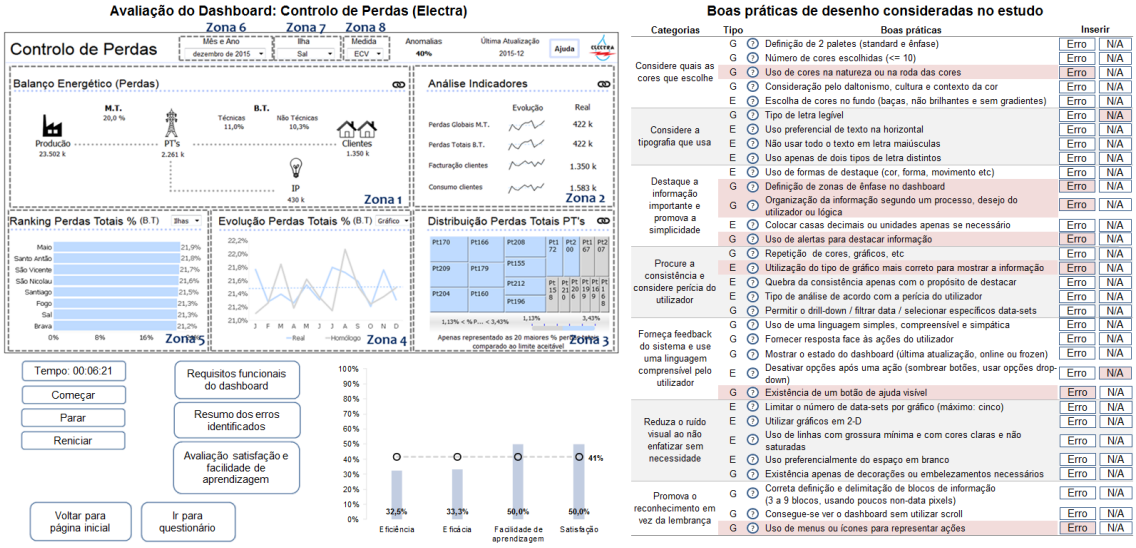


Figura 5 – Exemplo de uma avaliação utilizando o protótipo desenvolvido

De forma a contextualizar a amostra de inquiridos e obter feedback sobre o processo de avaliação, foi incorporado um questionário no protótipo, cujos detalhes são apresentados na Figura 6.

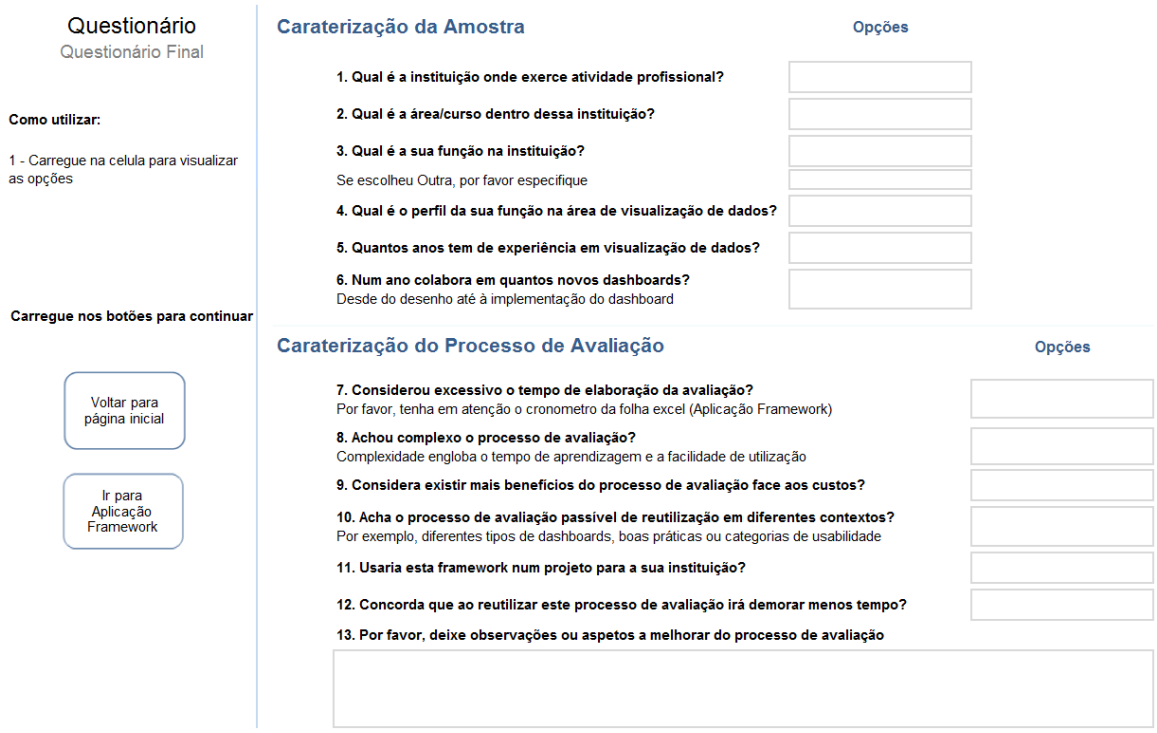


Figura 6 – Questionário embutido sobre a amostra e o processo de avaliação

Metodologia de distribuição do protótipo

A amostra de avaliadores da *framework* FMUED é composta por três perfis: (1) consultores de BI e IHM (competência de *Experience Design*) da Novabase; (2) professores de IHM; e (3) estudantes do mestrado em Engenharia Informática, com unidades curriculares em BI e IHM, e estudantes de licenciatura em Informática e Gestão, apenas com conhecimentos de BI, do ISCTE-IUL.

A amostra foi selecionada por conveniência, através de conhecimentos estabelecidos a nível académico com professores de várias instituições de Ensino Superior. A nível empresarial o protótipo foi enviado para todos os colaboradores da competência de *Business Analytics* envolvidos no desenho de interfaces. O método de distribuição pode ser observado na Figura 7. Cada inquirido selecionado recebeu por email duas versões do *dashboard* (escolhidas através de um gerador aleatório). As respostas foram todas devolvidas por email. A avaliação com os estudantes foi realizada em aulas práticas dedicadas, tendo sido distribuído aleatoriamente duas versões dos *dashboards* a cada estudante. De forma a garantir o anonimato das respostas, foi utilizado um extrator que eliminou a identificação do inquirido.

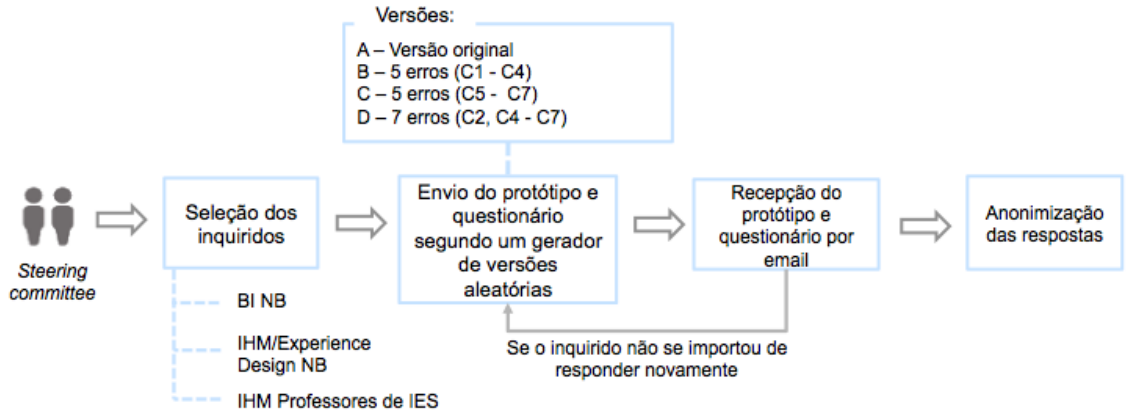


Figura 7 – Metodologia de distribuição do protótipo

Legenda: Categorias (C#) da Tabela 2; NB = Novabase; IES = Instituições do Ensino Superior

5. Análise de resultados

Foram identificados seguintes fatores que podem ter impacto nos resultados obtidos com a experiência:

- Nem todos os inquiridos possuíam igual experiencia no desenho de interfaces;
- Alguns inquiridos tiveram uma explicação prévia (e.g., os estudantes), outros apenas viram um vídeo explicativo;
- Por não haver controlo durante a avaliação, os avaliadores podem ter tomado decisões com dúvidas pendentes;
- Por não existir uma ordem de preenchimento obrigatória, os inquiridos podem ter respondido a uma versão com erros e sem erros posteriormente e vice-versa;
- O cronómetro tem que ser iniciado pelo avaliador, pelo que podem existir desvios face ao verdadeiro tempo decorrido.

Obteve-se uma amostra de 166 inquiridos, dos quais 50 eram profissionais (40 BI e 10 IHM) e 116 estudantes (68 BI e 48 BI + IHM). O tempo médio de avaliação foi aproximadamente de 11 min 30 s. Os estudantes demoram 12 min e os profissionais 11 min.

Constataram-se alterações da usabilidade consoante o número de erros propositados (Figura 8), variando entre os 76,8% e os 67,2% e uma média de 70,5%. Contudo, os resultados variaram consoante o tipo de inquirido (Figura 9). Para os profissionais, verificou-se um decréscimo por incremento de erros. Os estudantes afirmaram que com cinco erros possuía menor usabilidade do que com sete erros. Os estudantes deram menor usabilidade em todas as situações comparativamente aos profissionais.

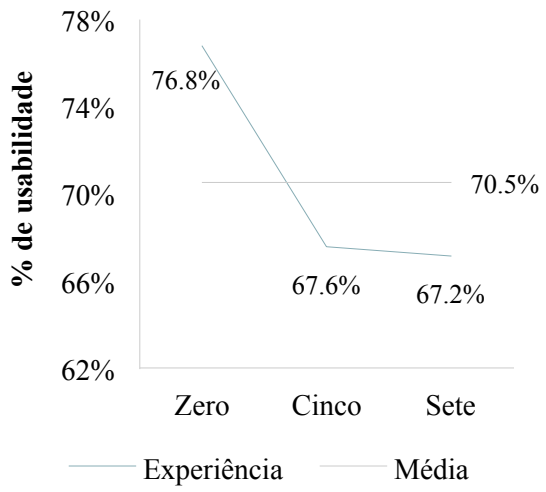


Figura 8 – Usabilidade por número de erros conforme o tipo de inquirido

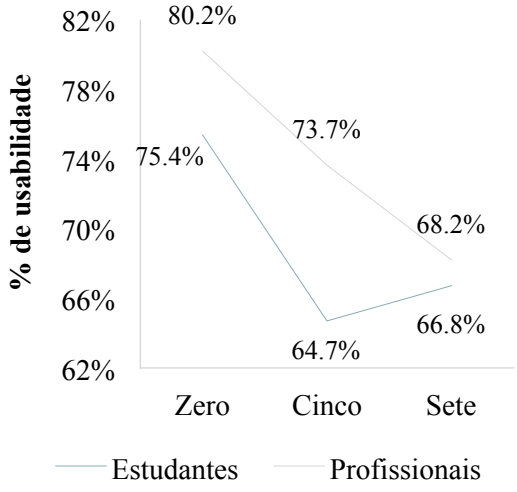


Figura 9 – Usabilidade por número de erros

A discrepância nas versões com cinco e sete erros decorre da diferença do número de erros específicos detetados entre as versões, conjugado com o acréscimo do número de boas práticas que foram consideradas não aplicáveis (Figura 10), e, ainda, por não existirem diferenças substanciais nos valores obtidos nas dimensões de satisfação e facilidade de aprendizagem com o incremento do número de erros (Figura 11).

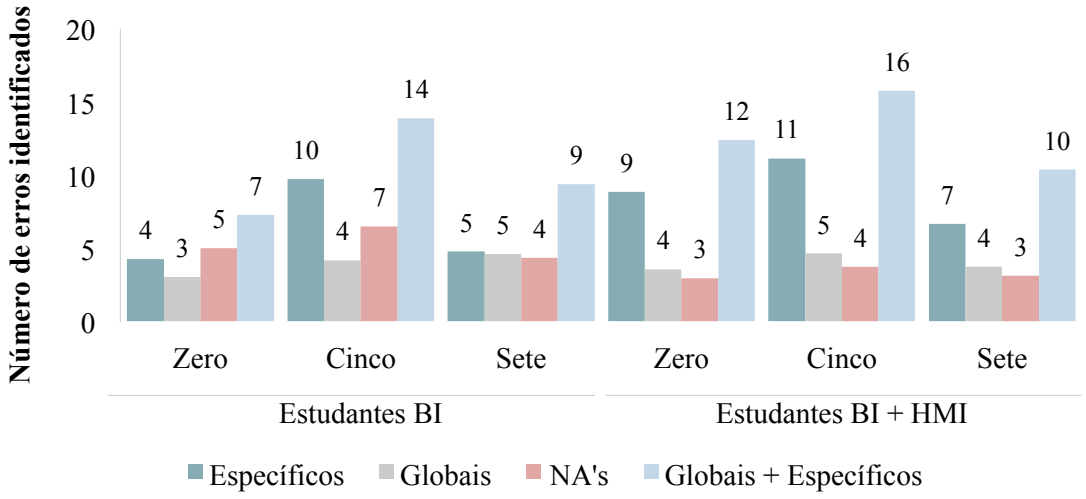


Figura 10 – Número de erros identificados pelos estudantes

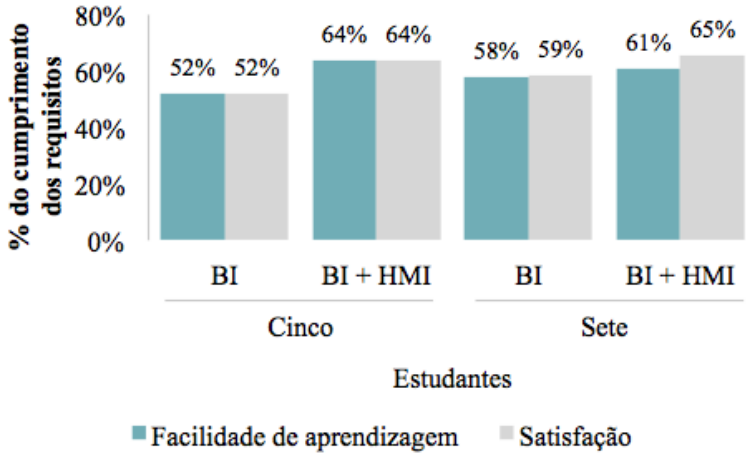


Figura 11 – Opinião em % dos estudantes no cumprimento dos requisitos de satisfação e facilidade de aprendizagem consoante o número de erros introduzidos

Analisando a percentagem de erros detetados (Figura 12), os profissionais, nas versões com cinco erros, apontaram uma percentagem superior média de identificação dos erros globais (69%), face aos estudantes (56%). Porém, nos erros específicos, verificou-se o oposto (35% vs. 55%). Na versão com sete erros, os estudantes (39%) reconheceram uma maior percentagem nos erros globais comparativamente aos profissionais (28%) e igualmente nos específicos (respetivamente, 49% e 44%).

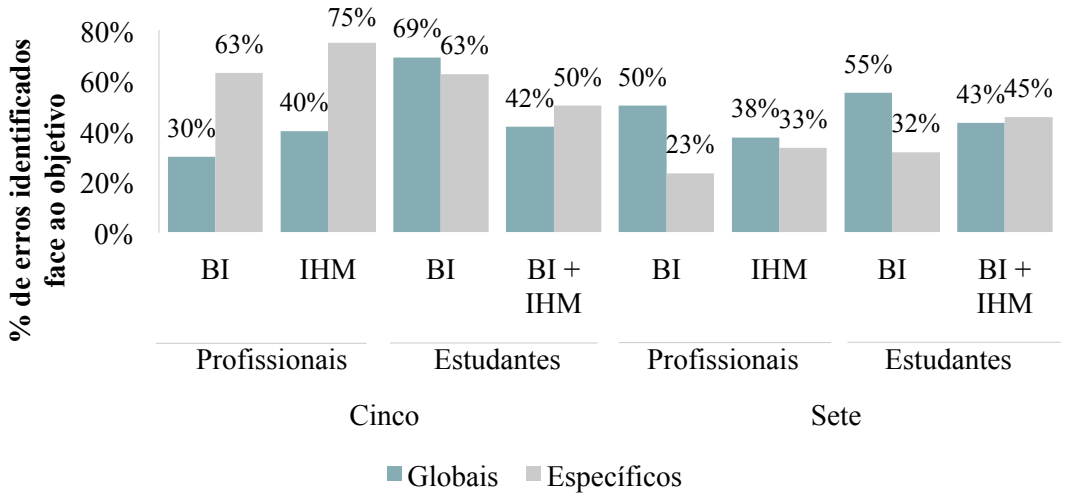


Figura 12 - % de erros identificados face ao objetivo por número de erros, avaliador e área

Relativamente à percentagem individual dos erros inseridos propositadamente, existem discrepâncias por tipo de avaliador.

Na versão B, apurou-se que, exceto no erro do texto em letras maiúsculas, os estudantes identificaram sempre em maior percentagem (Figura 13). Obteve-se uma média de 67%, concluindo-se que cada avaliador em média detetou 3,35 dos 5 erros introduzidos.

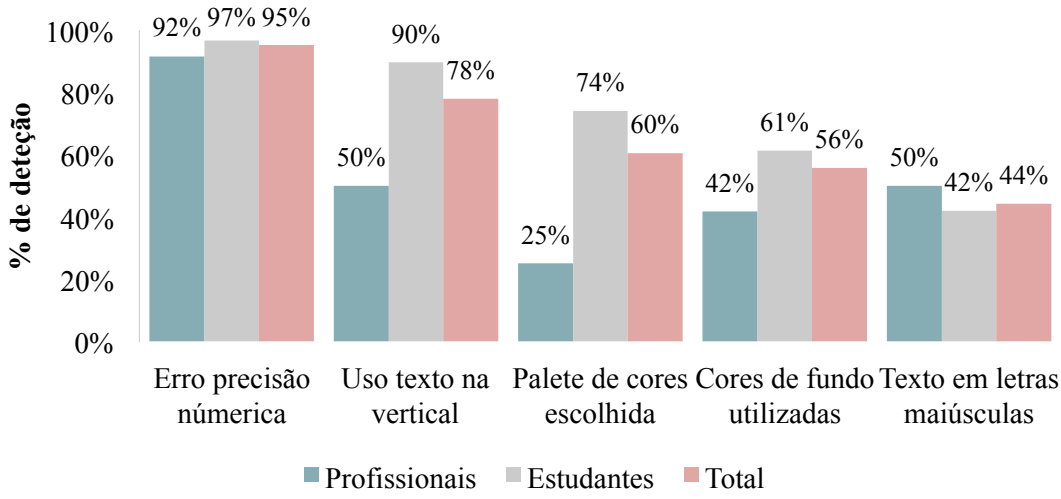


Figura 13 - % de erros detetados na versão B (5 erros)

Na versão C, existiam três casos (sem ajuda visível, aplicação de tabelas, decorações supérfluas) onde os profissionais conseguiram detetar uma maior percentagem (Figura 14). Porém, a utilização da barra de *scroll* e a separação de blocos, foi detetado em maior percentagem pelos estudantes. Concluiu-se que a percentagem total de identificação de erros foi de 41%, tendo cada avaliador detectado 2,05 dos 5 erros colocados.

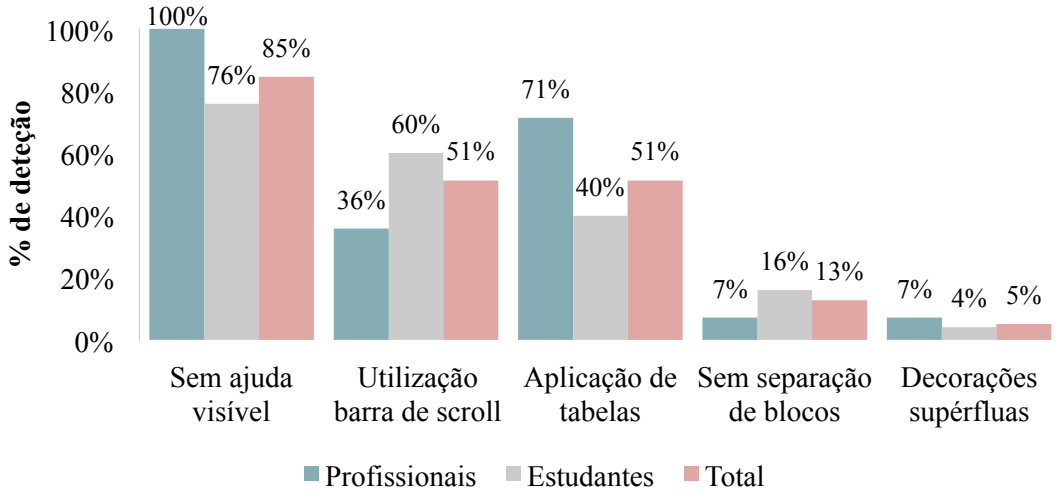


Figura 14 - % de erros detetados na versão C (5 erros)

Por fim, na versão D, os estudantes detetaram uma maior percentagem de erros, exceto no uso do texto na vertical (Figura 15). Verificou-se uma percentagem de 44% no total, significando que cada avaliador detetou 3,09 erros face aos 7 inseridos.

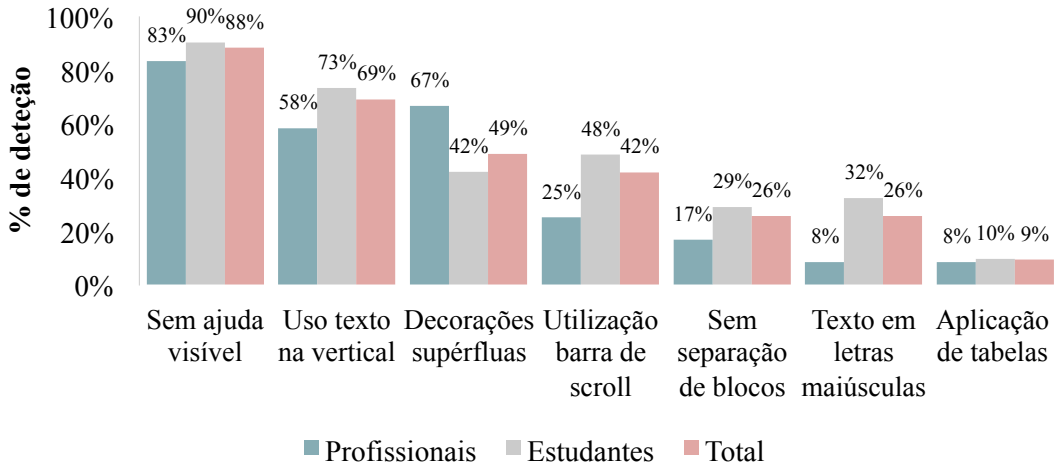


Figura 15 - % de erros detetados na versão D (7 erros)

Relativamente à complexidade do processo de avaliação (Figura 16), 47% dos inquiridos manifestaram opinião negativa e 18% uma opinião neutra sobre a complexidade do processo. Embora o protótipo desenvolvido tenha sido simplificado, ainda é necessário tornar o processo de avaliação mais intuitivo.

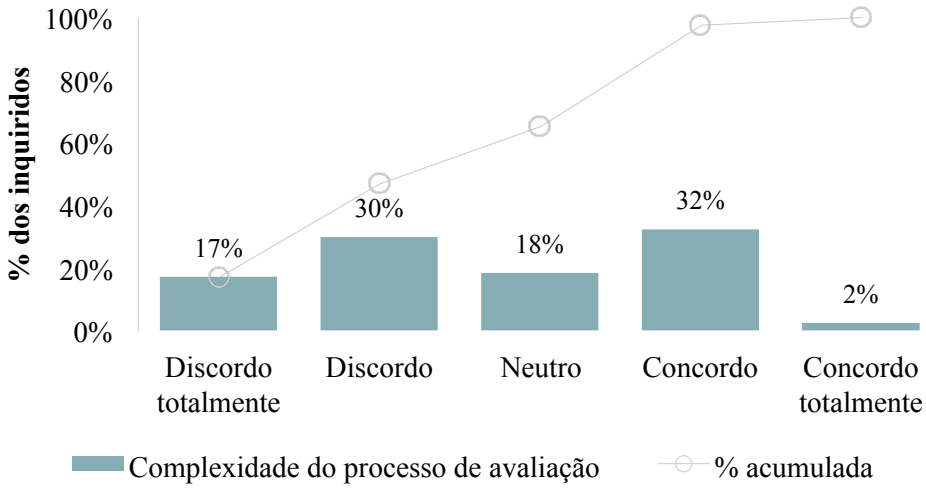


Figura 16 – Opinião sobre complexidade do processo de avaliação

Os profissionais consideram que a avaliação era reutilizável noutros contextos (diferentes tipos de dashboards, boas práticas e dimensões de usabilidade). Em relação à utilização do processo

de avaliação em projetos organizacionais, 70% responderam positivamente, o que é um aspecto bastante relevante (Figura 17).

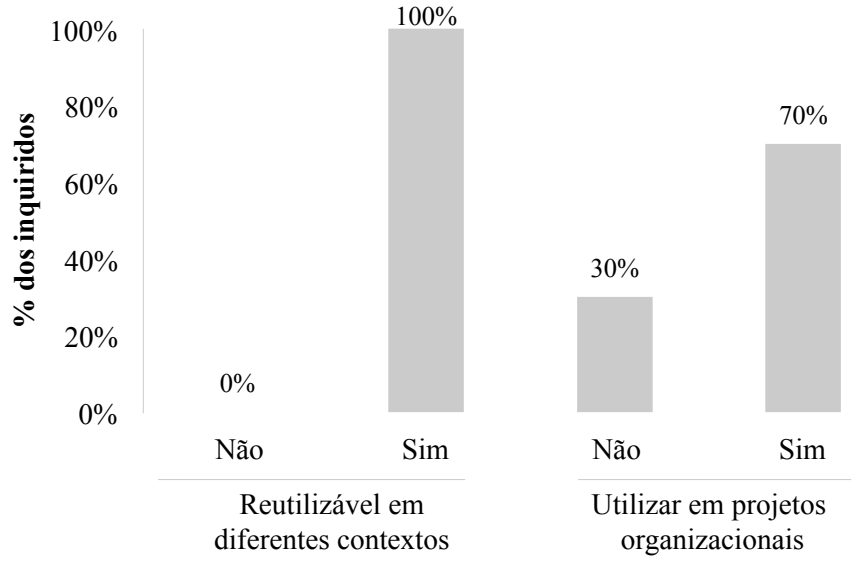


Figura 17 – Opinião dos profissionais sobre reutilização da *framework*

Análise das hipóteses de estudo

Nesta secção, será feita um análise detalhada das quatro hipóteses de estudo identificadas.

H1 - A usabilidade difere significativamente com a existência de erros?

De acordo com Few [2006], Fonseca et al. [2012] e Nielsen [1993] uma interface que não siga as boas práticas de desenho possui menor usabilidade. Assim, pretendeu-se averiguar se efetiva e significativamente se constatavam as diferenças referidas pelos autores.

Com esse fim, realizou-se o teste de Mann-Whitney U[2], obtendo-se um valor de 1595,0, z = -3571 e p = 0,000, a 95% de confiança, verificando-se a existência de diferenças significativas consoante a existência de erros. Também se apurou que os avaliadores deram maiores rankings à versão sem erros, comparativamente com as que tinham erros, 106,10 e 75,36, respectivamente.

A Figura 18 demonstra que o *dashboard* sem erros possui, em termos medianos, maior usabilidade face aos que apresentam erros. Mais de 50% dos avaliadores consideraram que a versão sem erros tinha mais de 82% de usabilidade, diferenciando dos 67,8% obtidos nas outras

[2] Devido ao pressuposto da normalidade não ter sido garantido nos subgrupos, aplicou-se um teste não paramétrico.

versões. Nos marginais, verifica-se apenas a existência de um, diferenciando da versão com erros, onde foram identificados dois.

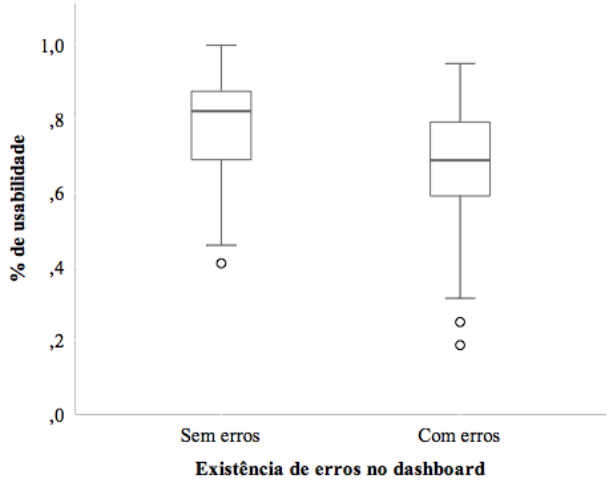


Figura 18 – Usabilidade consoante a existência de erros

H2 - A usabilidade diverge significativamente consoante os erros inseridos?

Para se apurar se as diferenças encontradas na H1 são visíveis consoante o número de erros introduzidos, aplicou-se o teste Kruskal-Wallis, $H_{(2)} = 12{,}766$, $p = 0{,}002$, a 95% de confiança, confirmando a existência das diferenças.

Na Figura 19 observa-se a distribuição para as diferentes situações, verificando-se que a versão com zero erros possui maiores valores nos quartis comparativamente aos restantes. Porém, nas versões com erros, as medianas são similares, possuindo a versão com cinco erros uma menor mediana, embora com uma distribuição menor e com valores superiores face aquela que tem sete erros. Todavia, possui 2 inquiridos marginais, diferenciando-se na versão com sete erros. Relativamente aos rankings, a versão com zero erros teve maior ranking (106,10), seguidamente, as versões com cinco erros (75,73) e, por fim, a com sete erros (74,66).

Para se apurar se as diferenças nas médias eram significativas, aplicou-se *a posteriori* o teste Scheffe[3], confirmando-se apenas diferenças significativas com zero erros face à de cinco (0,92) e à de sete erros (0,96) e vice-versa. O teste evidenciou dois subgrupos homogéneos, com e sem erros.

[3] Embora a ANOVA não seja robusta a violações da normalidade, o teste Scheffe é robusto, sendo aplicável nestas situações [Laureano 2013].

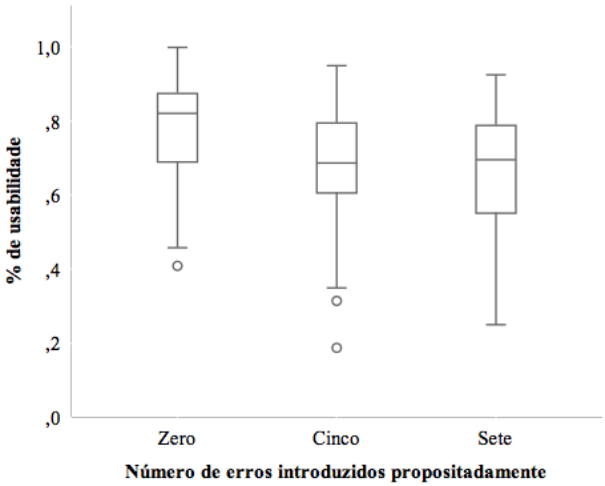


Figura 19 – Usabilidade consoante o número de erros introduzidos

H3 - A usabilidade diverge significativamente dependendo do tipo de avaliador?

Confirmadas as hipóteses H1 e H2 e sendo visíveis diferenças consoante o tipo de avaliador (Figura 9), questionou-se se estas são significativas. Assim, aplicou-se o t-test[4] e confirmou-se existirem significativas diferenças entre os tipos de avaliadores[5], $t_{(115,706)}$, $p = 0,014$, a 95% de confiança. Os profissionais (73,9%), em termos médios, avaliaram com maior usabilidade face aos estudantes (68%).

Na Figura 20 constatam-se as diferenças de usabilidade, tendo os profissionais uma mediana (75,9%), superior aos estudantes (68,5%). Contudo, a distribuição dos valores de usabilidade nos profissionais é inferior à dos estudantes, tendo sido identificados três profissionais considerados marginais e apenas um estudante.

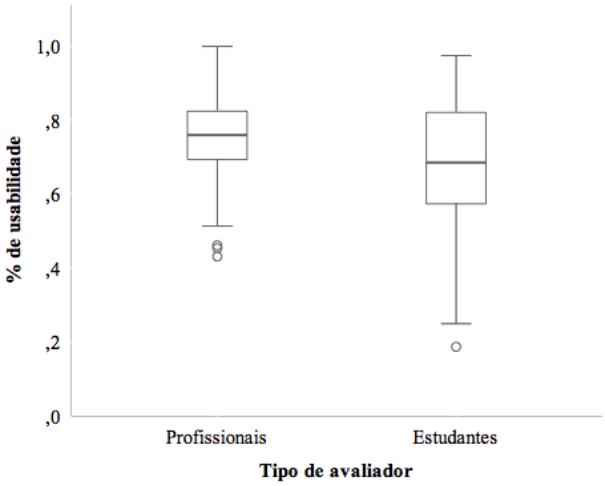


Figura 20 – Usabilidade por tipo de avaliador

[4] Constatou-se normalidade nos subgrupos e inexistência de fortes assimetrias nas distribuições.
[5] Observou-se inexistência de homogeneidade de variâncias, $F(4,204)$, $p = 0,042$.

H4 - Existem diferenças significativas na percentagem de boas práticas globais e específicas?

Durante a elaboração do estudo questionou-se se o tipo de boa prática (específica ou global) tinha influência na sua identificação por parte do avaliador. Com o intuito de confirmar se as diferenças eram significativas, aplicou-se o t-test de amostras emparelhadas[6], $t_{(121)} = 0{,}336$, $p = 0{,}737$, a 95% de confiança.

Confirmou-se que a percentagem de erros identificados específicos (50,6%) e globais (49,2%), face ao objetivo, não são significativamente diferentes, apresentando os erros específicos uma ligeira superioridade face aos globais (1,4%).

Por fim, verificou-se correlação muito baixa, Pearson $_{(122)} = 0{,}177$, sig = 0,051, concluindo-se que maior percentagem de erros específicos identificados não resulta numa maior percentagem de erros globais identificados e vice-versa.

Em síntese, das quatro hipóteses de estudo três foram suportadas: H1 (versão sem erros possuí maior usabilidade), H2 (a usabilidade da versão sem erros difere significativamente da das versões com cinco e sete erros), H3 (estudantes atribuíram uma menor usabilidade face aos profissionais). Apenas a hipótese H4 não foi suportada (o tipo de boa prática não influencia a percentagem de identificação de erros).

Feedback recebido

No questionário enviado (ver Figura 6) foi pedido aos avaliadores *feedback* sobre o processo de avaliação. Apenas os profissionais enviaram observações e aspetos a melhorar, relativos a questões de interação, heurísticas, requisitos de facilidade de aprendizagem e satisfação, e observações sobre os erros. Alguns dos comentários recebidos foram: "A utilização de uma imagem estática do *dashboard* não permite verificar se a interação segue as boas práticas"; "As heurísticas estão demasiado detalhadas, devendo ser removidas algumas e introduzidas outras (e.g., apresentação de informação hierarquicamente); "É dificil avaliar a facilidade de aprendizagem e satisfação por o avaliador não ser o utilizador final"; e "Apenas assinalar o erro não facilita a justificação do motivo, devendo ser possível comentar a decisão e afirmar se deriva de excesso ou defeito".

[6] Validaram-se os pressupostos das amostras emparelhadas.

6. Conclusões e trabalho futuro

O desenvolvimento de interfaces em BI está muitas vezes dissociado das boas práticas aceites em IHM, resultando em produtos finais com uma avaliação negativa por parte dos clientes de projetos de BI. O trabalho apresentado pretende quebrar essa tendência propondo uma *framework* que permite medir a usabilidade e eficiência de *dashboards*, utilizando boas práticas de desenho de BI e IHM, com adaptação da ISO/IEC 9241-11.

A *framework* desenvolvida possui duas componentes: um conjunto de heurísticas de desenho aplicadas a *dashboards* (organizadas em sete categorias, identificadas com base na revisão de literatura) e um método de avaliação. O trabalho foi desenvolvido no âmbito de uma dissertação de mestrado, mas inserido num contexto empresarial.

Foi desenhada uma experiência para validar a *framework*, incluindo a participação de avaliadores profissionais (consultores de BI e de IHM, e professores de IHM) e estudantes. Para tal, a Novabase forneceu um caso de estudo real aplicado ao sector energético em Cabo Verde, com vista ao desenvolvimento de um *dashboard* que permitisse controlar as perdas energéticas não técnicas, associadas a fraudes. Os objetivos principais da experiência consistiram na verificação das diferenças de usabilidade por número de erros introduzidos e na observação da aceitação da *framework* pela equipa de BI da Novabase. A amostra de avaliadores incluiu 166 inquiridos, dos quais 50 profissionais (40 BI e 10 IHM) e 116 estudantes (68 BI e 48 BI + IHM). O tempo médio de avaliação foi aproximadamente de 11 min 30 s, sem diferenças significativas entre estudantes e profissionais.

Os resultados demonstram as diferenças de usabilidade consoante o número de erros introduzidos e o tipo de inquiridos. Três das quatro hipóteses de estudo foram suportadas. Concluiu-se que a usabilidade do *dashboard* sem erros é maior (H1) e que a usabilidade da versão sem erros difere significativamente da das versões com cinco e sete erros (H2). Verificou-se também que os estudantes atribuíram uma menor usabilidade face aos profissionais (H3). Contudo, para os estudantes, um *dashboard* com sete erros não significa menor usabilidade face aos *dashboards* com cinco erros. Relativamente à percentagem de identificação de erros específicos e globais, não foram encontradas diferenças significativas, tendo-se concluído que o tipo de boa prática não influencia a identificação do erro (H4, não suportada). Finalmente, um aspecto a realçar é a elevada percentagem de respostas positivas (70%) por parte dos profissionais quanto à reutilização da *framework* noutros contextos, nomeadamente em projetos com clientes.

Este estudo apresenta algumas limitações, que serão colmatadas com trabalho futuro. Apesar de se ter medido a usabilidade em termos médios, a avaliação das dimensões de satisfação e

facilidade de aprendizagem não produziu resultados satisfatórios, de acordo com o *feedback* recebido pelos profissionais. Será necessário investigar e testar um novo método para validar estes requisitos, recorrendo em particular aos utilizadores finais [Nielsen, 1993]. O protótipo possui fragilidades devido ao software utilizado (Microsoft Excel©) que não permite testar a interatividade dos *dashboards* em diferentes sistemas operativos) e consequentemente a escolha de desenho feita (utilização de uma imagem estática).. Este problema será resolvido na solução final da *framework*, que deverá ser colocada em utilização na empresa patrocinadora do estudo. Atendendo ao feedback recebido acerca da complexidade do processo de avaliação, a solução final deverá contemplar um método mais intuitivo.

De forma a consolidar o trabalho desenvolvido, pretende-se realizar uma nova experiência de validação com uma maior amostra de avaliadores e estudando mais duas hipóteses que não puderam ser verificadas neste estudo: *Existem categorias de heurísticas onde a identificação foi significativamente menor?* e *Existe uma correlação negativa e significativa entre o número de erros detetados e a usabilidade?* Devido ao número limitado de erros selecionados, não foi possível incluir erros de todas as categorias. A ordem de avaliação dos dashboards pode também ter impacto nos resultados finais, pelo que a metodologia deve dividir a amostra em duas partes e prescrever para cada uma a ordem de avaliação, i.e., avaliar primeiro a versão sem erros (ou com menos erros) e vice-versa.

Referências

Eckerson, W. W., *Performance Dashboards: Measuring, Monitoring, and Managing Your Business* (2nd ed.), Wiley, 2010.

Few, S., *Information Dashboard Design: The Effective Visual Communication of Data*, O'Reilly Media, 2006.

Few, S., *Now You See It: Simple Visualization Techniques for Quantitative Analysis*, Analytics Press, 2009.

Few, S., Visual Business Intelligence - Assessing the Effectiveness of a New Dashboard's Design, http://www.perceptualedge.com/blog/?p=672, (20 Dezembro 2014), 2010.

Few, S., *Information Dashboard Design: Displaying Data for At-a-Glance Monitoring* (2nd ed.), Analytics Press, 2013.

Fonseca, M. Campos, P. Gonçalves, D., *Introdução ao Design de Interfaces* (2nd ed.), FCA, 2012.

Green, D. Pearson, J. M., "Development of a web site usability instrument based on ISO 9241-11", *The Journal of Computer Information Systems*, 47*,* 1 (2006), 66–72.

ISO/IEC 9241-11, *Ergonomic requirements for office work with visual display terminals (VDT) - Part 11: Guidance on usability*, 1998.

Johnson, J. Henderson, A., "Conceptual models: begin by designing what to design", *Interactions*, 9, 1 (2002), 25–32.

Jokela, T. Iivari, N. Matero, J. Karukka, M., "The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11", in Proceedings of Latin American conference on Human-computer interaction, ACM, NY, 2003, 53-60.

Joo, S. Lin, S. Lu, K., "A Usability Evaluation Model for Academic Library Websites: Efficiency, Effectiveness and Learnability", *Journal of Library and Information Studies*, 9, 2 (2011), 11–26.

Lakatos, E. M. Marconi, M. de A., *Fundamentos de metodologia científica* (5th ed.), Atlas, 2003.

Laureano, R. M. S., *Testes de Hipóteses com o SPSS* (2nd ed.), Edições Silabo, 2013.

Lourenço, P., *Evaluation of Business Intelligence application design using the principles of Visual Perception*, ISCTE-IUL, Lisboa, 2013.

Molich, R. Nielsen, J., "Improving a human-computer dialogue", *Communications of the ACM*, 33, 3 (1990), 338–348.

Nielsen, J., *Usability Engineering*, Morgan Kaufmann, 1993.

Nielsen, J., "Enhancing the explanatory power of usability heuristics", in *Proceedings of the SIGCHI conference on Human factors in computing systems,* ACM, 1994, 152-158.

Norman, D. A., *The Design of Everyday Things* (2nd ed.), Basic Books, 2002.

Rudin, K., "The Shift to On-Demand Business Intelligence", *Business Intelligence Journal*, 12, 1 (2007), 44–54.

Tufte, E., *Envisioning Information* (2nd ed.), Graphics Press, 1990.

Ware, C., *Information Visualization: Perception for Design* (3rd ed.), Morgan Kaufmann, 2012.

White, C., *Improving Business Intelligence Usability*, http://www.b-eye-network.co.uk/view/1557, (22 Dezembro 2014), 2005.

Metodologia para Análise de Oportunidades de Melhoria do Sistema de Gestão de Transporte de Doentes

Constança Roquette[1], Mónica Duarte Oliveira[1], José Tribolet[2], Lúcia Costa[3], Paulo Marques[3]

1) Centro de Estudos de Gestão, Instituto Superior Técnico, Universidade de Lisboa, Portugal
constanca.roquette@tecnico.ulisboa.pt, monica.oliveira@tecnico.ulisboa.pt

2) Departamento de Engenharia Informática, Instituto Superior Técnico, Universidade de Lisboa, Portugal
jose.tribolet@tecnico.ulisboa.pt

3) Link Consulting, Portugal
lucia.costa@linkconsulting.com, paulo.marques@linkconsulting.com

Resumo

O Sistema de Gestão de Transporte de Doentes (SGTD) é uma plataforma integrada *on-line*, desenvolvida pela Link Consulting, que suporta todos os procedimentos inerentes ao transporte programado de utentes. Atualmente, este sistema é usado na gestão do transporte não urgente de utentes em praticamente toda a rede de cuidados primários do Serviço Nacional de Saúde (SNS) português.
Este estudo reporta o desenvolvimento e implementação de uma metodologia que permite analisar oportunidades de melhoria dos algoritmos do SGTD, contribuindo para a literatura da área de transporte programado de utentes ao explorar alternativas na lógica e parâmetros dos algoritmos de organização de transporte. Estas alternativas têm em conta os objetivos do sistema de transporte, assim como os seus *stakeholders*-chave.
A metodologia proposta foi aplicada ao SGTD, e os resultados mostram de que forma foi gerada informação capaz de apoiar a Link Consulting na análise de melhorias ao sistema que vão ao encontro dos interesses dos múltiplos *stakeholders*.

Palavras-Chave: Sistemas de Informação na Saúde, Sistemas de Transporte, Transporte de Utentes, Análise de *Stakeholders*, Melhorias de Sistemas, Análise de Conflito

1. Introdução

Em Portugal, um dos objetivos fundamentais do Serviço Nacional de Saúde (SNS) consiste na promoção de igualdade de oportunidades de acesso a cuidados de saúde aos cidadãos [Ministério da Saúde 1990]. Neste âmbito, o SNS é responsável por providenciar um serviço de transporte a quem, por insuficiência económica e condição clínica incapacitante, não seja capaz de se deslocar às entidades prestadoras de serviços de saúde pelos seus próprios meios [Ministério da Saúde 2011] [Ministério da Saúde 2012].

De forma a dar resposta a esta necessidade, no âmbito do transporte não urgente de utentes, múltiplas entidades do SNS utilizam o Sistema de Gestão de Transporte de Doentes (SGTD), uma plataforma *on-line* desenvolvida pela Link Consulting que integra todas as atividades e

gestão de procedimentos relativos a este serviço. Atualmente, o SGTD encontra-se implementado em praticamente toda a rede de cuidados primários do SNS.

Este sistema não consiste apenas numa ferramenta de gestão do transporte de utentes, sendo também responsável pela formação de agrupamentos de transporte, pela atribuição de entidades transportadoras aos agrupamentos formados e, por fim, pela contabilização dos custos gerados no transporte. Estas ações são feitas com recurso a três algoritmos heurísticos - (1) Algoritmo de formação de agrupamentos de transporte; (2) Algoritmo de atribuição de entidade transportadora; (3) Algoritmo de contabilização - os quais satisfazem um conjunto de regras e constrangimentos, sobre os quais se baseia a análise do presente trabalho.

O SGTD é utilizado por um conjunto de *stakeholders* diretamente envolvidos no processo de transporte não urgente de utentes, sendo estes: entidades requisitantes, entidades transportadoras, entidades prestadoras e entidades supervisoras. Desta forma, o SGTD apresenta uma arquitetura modular orientada a serviços, uma vez que presta um conjunto de serviços independentes, que comunicam entre si, de modo a criar um *software* de aplicação funcional [Laudon & Laudon 2011]. De forma a melhor compreender as várias fases do processo de organização, planeamento e contabilização do transporte não urgente de utentes, bem como o papel dos vários intervenientes ao longo deste processo, apresenta-se o seguinte diagrama do funcionamento do SGTD (Figura 1).

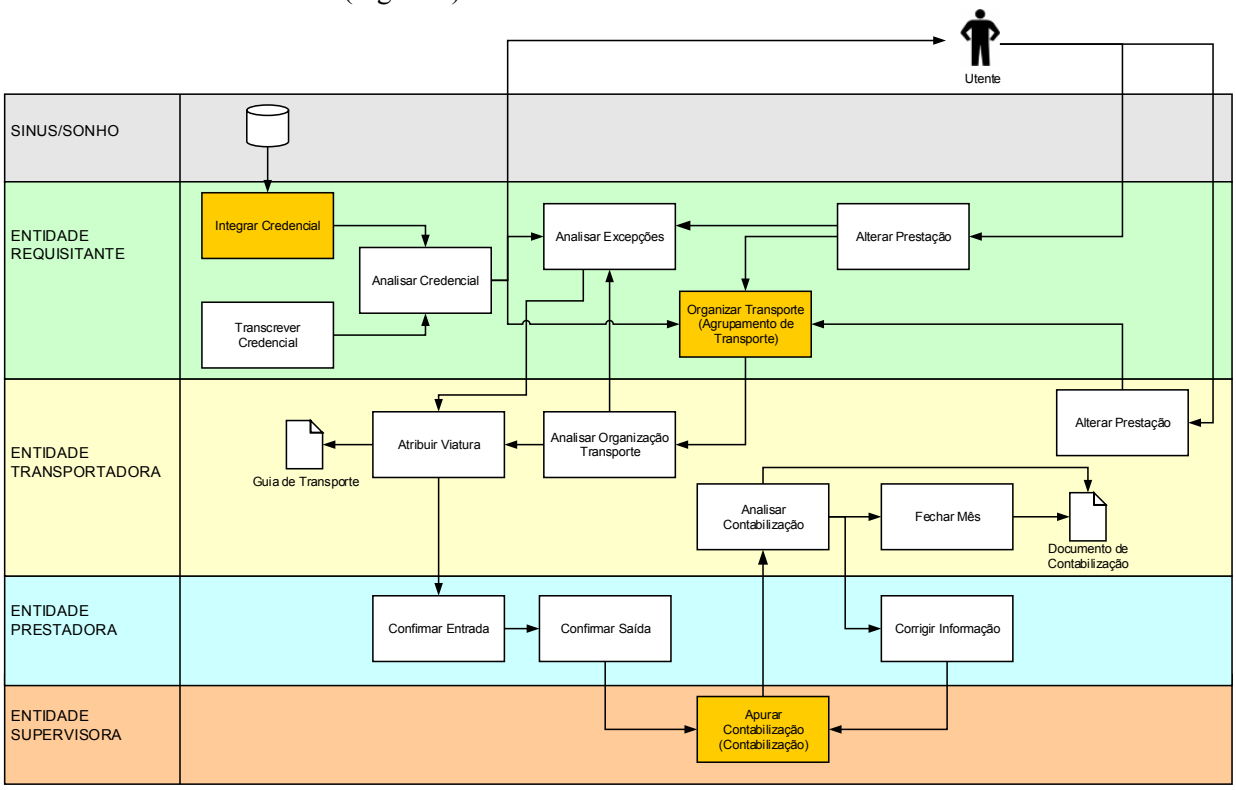


Figura 1 – Diagrama representativo do funcionamento do SGTD [Link Consulting 2010].

Numa primeira análise da Figura 1 é possível compreender que os quatro intervenientes destacados atuam em diferentes fases do transporte programado de utentes, correspondendo a quatro tipos distintos de utilizadores no sistema. Cada um destes utilizadores tem um papel bem definido e autónomo no SGTD, havendo adicionalmente outros intervenientes diretos e indiretos neste sistema.

Dado o contexto socioeconómico em que Portugal atualmente se encontra, há uma pressão para a redução de custos. Adicionalmente, a Link Consulting pretende melhorar o seu produto, estando interessada em explorar alternativas que melhorem o funcionamento dos algoritmos do SGTD. No entanto, a introdução de alterações a um sistema como o SGTD tem associada um conjunto de desafios, uma vez que há vários objetivos a serem prosseguidos no transporte de utentes e que alterações ao sistema têm implicações numa multiplicidade de *stakeholders* com interesses, por vezes, em conflito. Assim, afigura-se pertinente a análise de oportunidades de melhoria a este sistema de transporte, tendo em conta os vários pontos de vista e interesses dos seus múltiplos *stakeholders*. Como tal, este estudo desenvolve uma metodologia para analisar alterações na lógica e parâmetros dos algoritmos do SGTD, de forma a compreender o seu impacto nos múltiplos objetivos e *stakeholders* do sistema e a avaliar a viabilidade da sua implementação. Através desta análise espera-se gerar informação útil que permita apoiar a Link Consulting na análise de melhorias ao SGTD.

Este artigo está organizado nas seguintes secções: **secção 2**, na qual se apresentam as principais conclusões após revisão da literatura na área de transporte programado de utentes; **secção 3**, onde se apresenta a metodologia proposta neste estudo; **secção 4**, na qual se encontra uma síntese dos resultados obtidos ao aplicar a metodologia ao SGTD; e por fim, **secções 5** e **6**, onde se discutem as principais limitações e vantagens da metodologia proposta neste estudo e a conclusão, respetivamente.

2. Revisão Bibliográfica

2.1. Transporte Programado de Utentes

Após identificação do âmbito do SGTD e das suas características globais, realizou-se uma revisão bibliográfica com o intuito de compreender de que forma o transporte programado de utentes é abordado na literatura e de identificar sistemas semelhantes ao SGTD. Esta revisão explorou metodologias existentes que pudessem ser usadas na análise de melhorias ao SGTD. Foi possível concluir que nenhuma metodologia existente satisfazia na íntegra as especificidades do SGTD e permitia responder aos objetivos deste estudo – nomeadamente, nenhuma metodologia tem em consideração o impacto das alterações nos seus múltiplos

stakeholders. A análise da literatura evidenciou que o problema de transporte programado de passageiros é tipicamente abordado como um problema de otimização de rotas [Cordeau & Laporte 2007]. Este tipo de abordagem não se coaduna com este estudo uma vez que o SGTD não é responsável pelo planeamento de rotas de pedidos de transporte, mas, apenas, pela formação dos seus agrupamentos, cabendo às entidades transportadoras a escolha posterior do melhor percurso.

No âmbito dos cuidados de saúde, vários estudos encontrados na literatura pretenderam identificar melhorias em sistemas de transporte programado de utentes com recurso a modelos de simulação [Melachrinoudis et al. 2007] [Hanne et al. 2007] [Bowers et al. 2012]. Através dos resultados obtidos, estes estudos formularam recomendações que, quando implementadas, introduziram significativas melhorias para cada um dos sistemas em análise. O recurso a modelos de simulação revelou ser um meio eficaz para explorar oportunidades de melhoria de um sistema, indo ao encontro dos objetivos deste trabalho. No entanto, nenhuma das metodologias usadas nestes estudos tem em conta as especificidades do SGTD e aborda o impacto das alterações ao sistema nos diferentes *stakeholders* dos sistemas, como se pretende neste trabalho.

Deste modo, ficou clara a necessidade de desenvolver uma metodologia ajustada aos objetivos e *stakeholders* do SGTD. Tal requereu uma pesquisa de métodos e técnicas que suportassem cada etapa constituinte da metodologia a ser proposta. De seguida, apresenta-se uma breve revisão de métodos de análise de *stakeholders* e de abordagens para estruturar objetivos de um sistema.

2.2. Metodologias de Análise de *Stakeholders*

Uma análise de *stakeholders* revela-se útil na identificação e categorização dos vários *stakeholders* de um sistema, permitindo compreender os seus diferentes graus de envolvimento, interesses e relações entre si. Uma metodologia de análise de *stakeholders* deve apresentar, necessariamente, três etapas [Reed et al. 2009]: uma primeira, em que é identificado o sistema em foco e as suas fronteiras (análise de contexto). Uma segunda etapa, onde se aplicam os métodos de análise de *stakeholders*, de forma a identificar os vários intervenientes no sistema, bem como o interesse que nele têm. Nesta fase, é também feita uma categorização dos *stakeholders*, consoante os seus interesses e grau de envolvimento no sistema, bem como uma análise das relações entre estes. Por último, devem ser feitas recomendações acerca de estratégias de comunicação adequadas a cada perfil de *stakeholders* do sistema.

Após serem identificados, os *stakeholders* podem ser categorizados de acordo com a sua localização na organização/sistema: internos, quando operam dentro das fronteiras da

organização/sistema; de interface, quando interagem com o ambiente externo; e externos, sendo aqueles que podem contribuir, competir ou ter algum especial interesse no funcionamento da organização/sistema [Fottler 1990].

Existem diversos métodos que permitem classificar o posicionamento dos *stakeholders* no sistema em que se inserem. Destaca-se o método de mapeamento dos *stakeholders* numa matriz de interesse *versus* influência [Reed et al. 2009], visto tratar-se de uma ferramenta visual que permite a fácil perceção de vários fatores de interesse num contexto de decisão: (1) Grau de envolvimento de cada *stakeholder* no sistema; (2) Possibilidade de encontrar padrões na distribuição dos vários *stakeholders* na matriz; (3) Identificação de potenciais resistências à mudança; (4) Previsão de alianças [Reed et al. 2009].

2.3. Estruturação e Operacionalização dos Objetivos do Sistema

De modo a identificar melhorias num sistema, revela-se necessário estruturar com rigor os seus objetivos, para compreender de que forma estes se relacionam com os vários *stakeholders*.

Existem dois tipos de objetivos: fundamentais e intermédios [Keeney 1992]. Num contexto de decisão, um objetivo fundamental representa uma razão essencial de interesse de um sistema/organização. Objetivos intermédios, por sua vez, têm interesse pela forma como se relacionam com os objetivos fundamentais, influenciando o seu grau de alcance. No entanto, em situações de análise de decisão, ou de avaliação de alternativas, o foco do estudo deve centrar-se nos objetivos fundamentais. Como tal, estes representam o conjunto de objetivos sobre os quais os indicadores de desempenho devem ser definidos [Keeney 1992].

Assim, após identificação dos objetivos fundamentais e intermédios de um sistema e da sua operacionalização através da definição de indicadores de desempenho, torna-se possível estabelecer o foco da análise aquando da simulação de alterações nos seus algoritmos. Desta forma, o impacto destas alterações será medido através do maior, ou menor, grau de sucesso alcançado para cada um dos objetivos e *stakeholders* associados. Estes conceitos serão usados neste estudo na análise da adequação e viabilidade das alterações propostas aos algoritmos do SGTD.

3. Metodologia Proposta

Com o intuito de gerar informação capaz de apoiar a Link Consulting na análise de oportunidades de melhoria nos algoritmos do SGTD, foi desenvolvida uma metodologia ajustada às características e objetivos deste sistema. A metodologia proposta permite analisar o impacto de alterações na lógica funcional e parâmetros dos algoritmos do SGTD, nos múltiplos

objetivos e intervenientes do sistema. Para tal, diferentes métodos e técnicas com base teórica sólida foram combinados, levando à proposta de uma multimetodologia, como definido por [Mingers & Brocklesby 1976]: esta combina diferentes métodos e técnicas que, quando considerados conjuntamente, permitem analisar o sistema como um todo, de forma mais eficaz.

3.1. Visão Global da Metodologia Proposta

A metodologia proposta é composta por várias etapas, como representado na Figura 2, e adota uma abordagem sociotécnica [Phillips & Bana e Costa 2007], combinando tanto processos técnicos, como sociais – duas componentes de igual importância [Cherns 1976]. Estes últimos devem-se ao facto deste trabalho pretender apoiar a decisão da Link Consulting, sendo como tal pertinente ajustar e aferir as diferentes análises de acordo com as necessidades e pontos de vista da empresa.

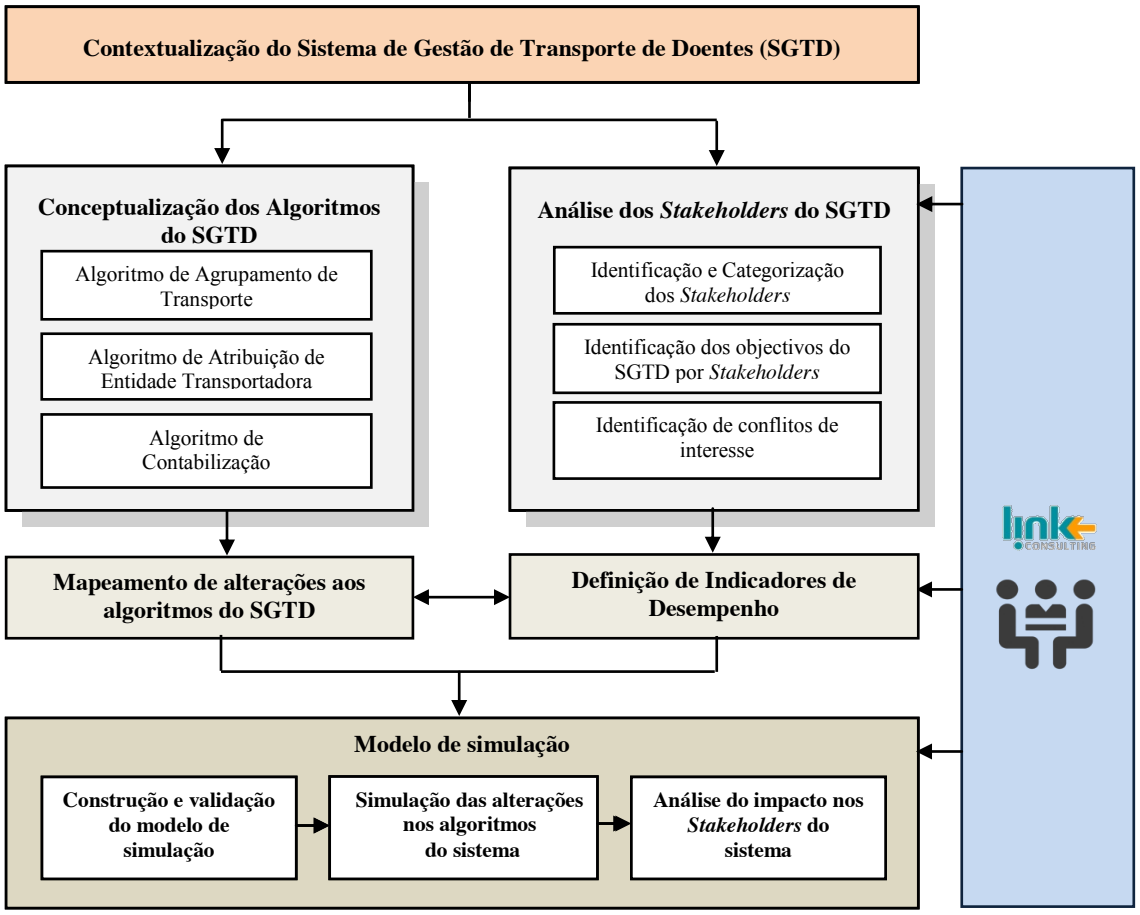


Figura 2 – Desenho metodológico.

Relativamente à componente social da metodologia, esta consistiu num processo contínuo de interação com a equipa de *HealthCare Solutions* da Link Consulting, ao longo das várias fases

do trabalho. Esta interação foi feita sob a forma de reuniões e *workshops* com os vários elementos da equipa, sempre que se julgou pertinente explorar os pontos de vista da empresa e aferir se o trabalho realizado ia ao encontro da análise pretendida. Quanto à componente técnica, um conjunto de métodos e técnicas foram desenvolvidos e implementados, nomeadamente nas fases de análise de *stakeholders*, estruturação de objetivos e construção do modelo de simulação.

Para o desenvolvimento da metodologia, numa primeira etapa foi feita a contextualização do SGTD, de forma a compreender o âmbito do transporte não urgente de utentes em Portugal, os seus intervenientes e o seu funcionamento como plataforma integrada *on-line* de arquitetura modular. Tal permitiu propor uma metodologia que fosse ajustada ao contexto e características e objetivos deste sistema. Em seguida, serão descritas em maior detalhe as componentes da metodologia proposta.

3.2. Conceptualização dos Algoritmos do SGTD

Os algoritmos do SGTD são algoritmos heurísticos uma vez que consistem num conjunto de regras e constrangimentos que se sabe serem capazes de garantir o bom funcionamento do sistema, quer no agrupamento de prestações, atribuição da entidade transportadora ou contabilização dos transportes. Desta forma, não oferecem ao sistema soluções ótimas, sendo no entanto suficientemente boas do ponto de vista prático.

Os três algoritmos do SGTD, implementados na linguagem de programação *Java*, atuam no sistema de modo sequencial e como um ato programado que decorre a cada 20 minutos, de forma a permitir a contínua entrada de novos pedidos de transporte no sistema e a contabilização dos que vão sendo realizados. Após terem sido identificadas as principais regras e princípios de funcionamento dos algoritmos do SGTD, foi possível propor algumas alterações que se julgaram relevantes para análise. Estas serão apresentadas na secção 3.5.

3.3. Análise de *Stakeholders* do SGTD

Após contextualização do SGTD foi possível realizar uma análise dos seus *stakeholders*, através da metodologia proposta por [Reed et al. 2009]. Esta permitiu identificar e categorizar os *stakeholders* do sistema, bem como compreender os seus interesses e inter-relações entre estes.

3.3.1. Identificação e Categorização de *Stakeholders* do SGTD

Entende-se por *stakeholder* todo o indivíduo, conjunto de indivíduos, ou entidades, que são afetados pelas decisões e ações tomadas sobre um sistema e todos aqueles que têm poder para influenciar o resultado que destas advier [Freeman 1984]. Como tal, e de acordo com a

classificação apresentada por [Fottler 1990], identificaram-se e categorizaram-se os *stakeholders* do SGTD, como indicado na Tabela 1.

Stakeholders		Descrição	Papel do SGTD
***Stakeholders* Internos**	**Entidades Requisitantes**	Unidades de cuidados primários e da Rede Nacional de Cuidados Continuados Integrados (RNCCI)	▪ Prescrição do transporte na plataforma do SGTD; ▪ Pagamento do serviço de transporte (princípio do prescritor pagador).
	Entidades Transportadoras	Corporações de Bombeiros e da Cruz Vermelha Portuguesa e outras transportadoras públicas ou privadas desde que legitimadas para tal.	▪ Aceitação/rejeição dos agrupamentos de transporte na plataforma do SGTD e, em caso de aceitação, atribuição de quartel, viatura e motorista; ▪ Realização do ato de transporte dos utentes até à Entidade Prestadora.
	Entidades Prestadoras	Clínicas de Fisioterapia, Clínicas de Hemodiálise, RNCCI, Laboratórios, entre outras.	▪ Registo das entradas e saídas de utentes das suas instalações na plataforma do SGTD; ▪ Prestação dos serviços prescritos pela Entidade Requisitante (consultas, etc..).
***Stakeholders* de Interface**	**Entidade Contratante**	Serviços Partilhados do Ministério da Saúde (SPMS)	▪ Contratualização do sistema com a Link Consulting, fazendo a ponte entre a empresa e as várias Administrações Regionais de Saúde (ARS) e a Administração Central dos Serviços de Saúde (ACSS).
	Entidades Gestoras/ Supervisoras	*ARSs*	▪ Monitorização, gestão e supervisão do sistema; ▪ Resolução de eventuais conflitos; ▪ Realização de acordos com diversas entidades. ▪ Alteração das áreas de atuação e de casos de fidelização.
		ACSS	▪ Supervisão do SGTD.
***Stakeholders* Externos**	**Legisladores**	Ministérios da Saúde, Finanças e Economia; Parlamento.	▪ Criação de legislação que regule, normalize e apoie o bom funcionamento do SGTD.
	Financiadores	ACSS População contribuinte	▪ Financiamento do SGTD.
	Reguladores/ Inspecção	Entidade Reguladora da Saúde (ERS) Inspecção-Geral das Actividades em Saúde (IGAS)	▪ Regulação e garantia de qualidade dos serviços prestados pelas várias entidades intervenientes no SGTD; ▪ Auditoria, inspeção, fiscalização e desenvolvimento de ação disciplinar mediante o incumprimento de requisitos legais estabelecidos no âmbito do SGTD;

Tabela 1 – *Stakeholders* internos, de interface e externos do SGTD.

Como foco deste sistema, identificou-se o utente. Embora a sua ação não tenha praticamente influência no funcionamento do sistema, qualquer alteração que seja implementada poderá afetá-lo, tendo repercussões não só na forma como lhe é prestado o serviço de transporte em si,

mas, também, em caso de atrasos, ou má gestão de agrupamentos, pode levar à falta de comparência aos atos clínicos prescritos.

De forma a complementar esta primeira categorização dos *stakeholders* do SGTD e a compreender os seus diferentes perfis, foi desenvolvido um *workshop* com a equipa de *Healthcare Solutions* da Link Consulting com o intuito de mapear os *stakeholders* do SGTD numa matriz de interesse *versus* influência. Apresenta-se o resultado final na Figura 3.

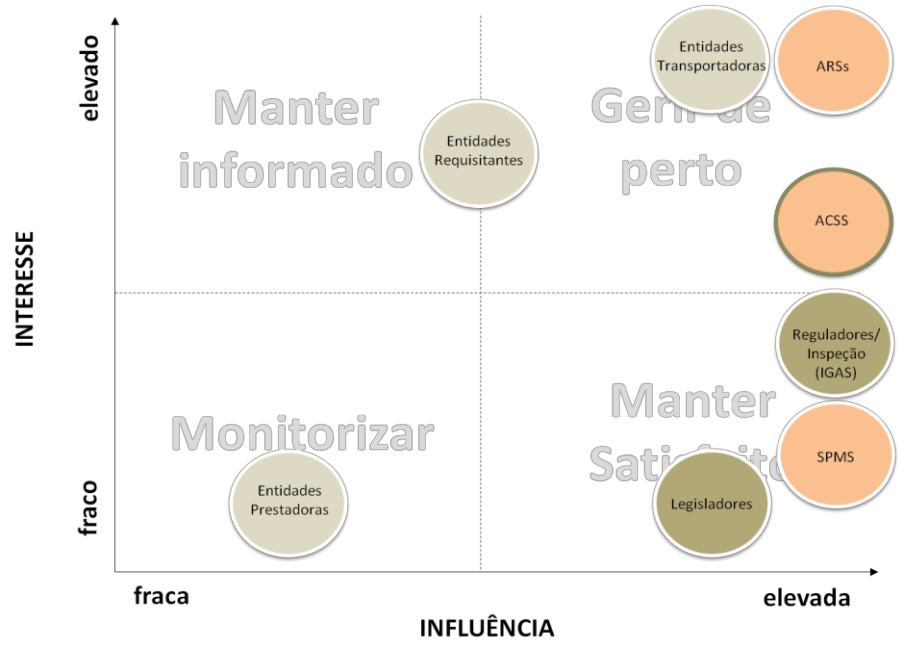


Figura 3 – Matriz de interesse *versus* influência dos *stakeholders* do SGTD.

Através da análise do mapeamento obtido foi possível identificar os *stakeholders*-chave do sistema.

3.3.2. Identificação dos Objetivos do SGTD por *Stakeholder*

Após terem sido identificados e categorizados os *stakeholders* do SGTD, foi feito um levantamento dos objetivos fundamentais e intermédios deste sistema (Figura 4). Posteriormente, foram identificadas as relações entre os objetivos fundamentais do SGTD e os interesses dos seus vários *stakeholders* (Tabela 2). Os objetivos identificados foram discutidos e validados em reunião com a Link Consulting.

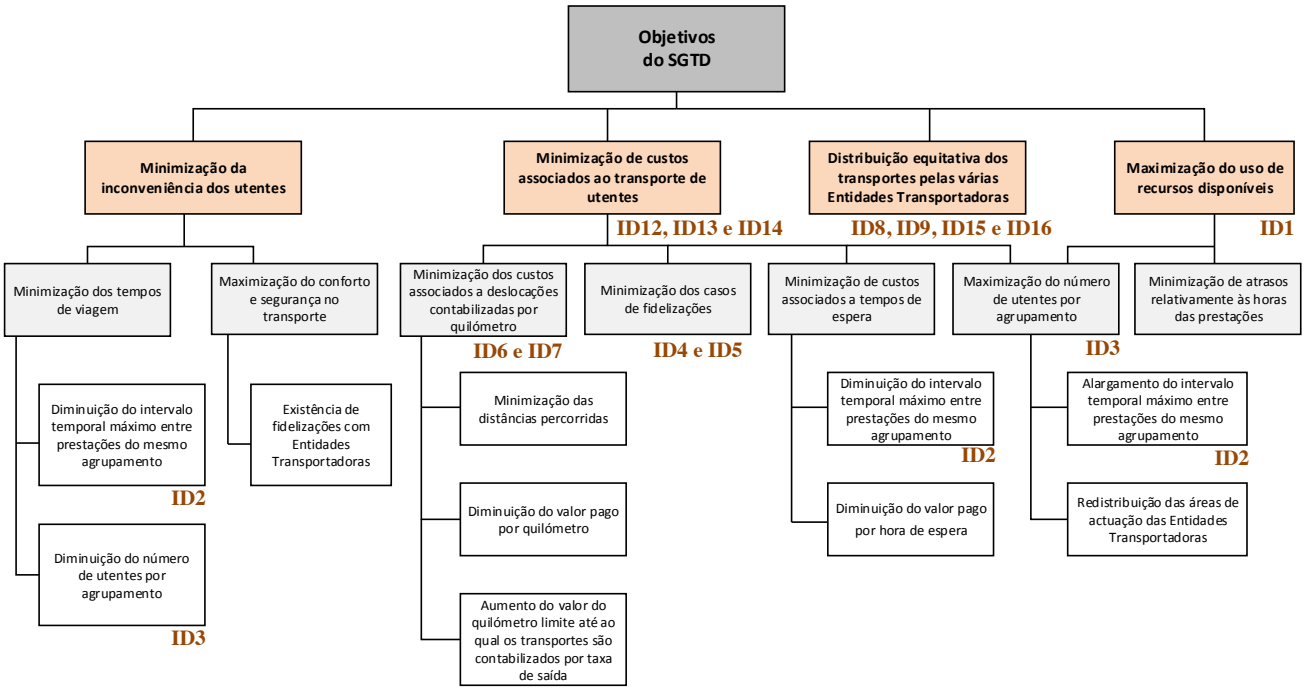


Figura 4 – Objetivos fundamentais do SGTD, respetivos objetivos intermédios e finais e indicadores de desempenho associados (ver correspondência na Tabela 3).

Objetivos fundamentais do SGTD	**Principais *stakeholders* associados**
Minimização da inconveniência dos utentes	Utentes, Entidades Requisitantes e Entidades Gestoras/Supervisoras.
Minimização de custos associados ao transporte de utentes	Entidades Requisitantes e Entidades Gestoras/Supervisoras.
Maximização do uso de recursos disponíveis (veículos, equipas, equipamentos de meios de diagnóstico, etc…)	Entidades Transportadoras, Entidades Prestadoras e Entidades Gestoras/Supervisoras
Distribuição equitativa dos agrupamentos de transporte pelas várias Entidades Transportadoras	Entidades Transportadoras e Entidades Gestoras/Supervisoras.

Tabela 2 – Objetivos fundamentais do SGTD e principais *stakeholders* associados.

Analisando o esquema da Figura 4, é possível identificar algumas situações de conflito entre os objetivos intermédios e finais de alguns objetivos fundamentais do SGTD. Deve ser, portanto, encontrada uma situação de compromisso que permita, por meio de algumas concessões, satisfazer os *stakeholders* cujo interesse maior, por exemplo, é minimizar a inconveniência dos utentes e *stakeholders* cuja preocupação maior recai na minimização dos custos de transporte.

3.4. Definição de Indicadores de Desempenho

Após terem-se identificado os objetivos fundamentais do SGTD, foram definidos indicadores de desempenho (Tabela 3), de forma a tornar mensurável o impacto produzido pelas alterações propostas aos algoritmos do sistema.

ID	Indicadores de Desempenho
ID1	Número de agrupamentos múltiplos
ID2	Intervalo temporal médio entre a primeira e última prestação / Agrupamento
ID3	Nº médio de prestações / Agrupamento
ID4	Custos associados a Fidelizações dentro da área de residência
ID5	Custos associados a Fidelizações fora da área de residência
ID6	Custo total associado a prestações contabilizadas por km
ID7	Custo total associado a prestações contabilizadas por taxa de saída
ID8	Percentagem de Entidades Transportadoras que aumentaram a média das distâncias máximas percorridas
ID9	Percentagem de Entidades Prestadoras que diminuíram a média das distâncias máximas percorridas
ID10	Custos totais associados a tempos de espera
ID11	Tempo de espera médio / Agrupamento
ID12	Percentagem de Agrupamentos de Centros de Saúde (ACeS's) que aumentaram a despesa associada ao transporte
ID13	Percentagem de ACeS's que diminuíram a despesa associada ao transporte
ID14	Custos totais associados ao transporte de utentes
ID15	Percentagem de Entidades Transportadoras que aumentaram a receita associada ao transporte
ID16	Percentagem de Entidades Transportadoras que diminuíram a receita associada ao transporte

Tabela 3 – Indicadores de desempenho definidos para a análise do impacto de alterações no SGTD.

3.5. Mapeamento de Alterações Propostas aos Algoritmos do SGTD

Após contextualização do SGTD, conceptualização dos seus algoritmos e análise dos seus *stakeholders*, foi feito um mapeamento das alterações aos algoritmos cujo impacto no sistema se pretendia analisar (Tabela 4). Estas alterações tiveram por base a procura de alternativas relativamente à lógica funcional e valores de parâmetros dos algoritmos atuais do sistema.

	Em funcionamento	Alterações a simular	
Alterações na lógica de formação de agrupamentos	Intervalos máximos entre prestações do mesmo agrupamento Destinos: ▪ Dentro do concelho: 1h ▪ Fora do concelho: o 2h (< 100km) o 4h (≥ 100km)	▪ Dentro do concelho: 1h ▪ Fora do concelho: o 1h (< 100km) o 2h (≥ 100km)	S9
		▪ Dentro do concelho: 1h ▪ Fora do concelho: o 1h (< 50km) o 2h (≥ 50km)	S10
	Agrupamento de utentes por igual **concelho destino** da prestação	Agrupamento de utentes por igual **freguesia destino** da prestação	S7
		Agrupamento de utentes por igual **Entidade Prestadora**	S6
Alterações na lógica de atribuição de Entidade Transportadora e na distribuição de áreas de atuação	Fidelizações de utentes a Entidades Transportadoras são possíveis e prioritárias na formação de agrupamentos	Inexistência de fidelizações	S4
		Fidelizações apenas a Entidades Transportadoras que atuem na freguesia de origem	S5
	Atribuição da responsabilidade total do transporte das prestações de uma credencial à mesma Entidade Transportadora	Distribuição equitativa das várias prestações de uma credencial pelas Entidades Transportadoras candidatas	S11
	Utentes pertencentes a conjuntos de freguesias diferentes de uma Entidade Transportadora não podem ser agrupados	Permissão de agrupamento de utentes, desde que pertençam à mesma área de atuação, independentemente dos conjuntos de freguesias	S8
Alterações de parâmetros de contabilização	Valor máximo pago pelo SNS por quilómetro percorrido = 0.51 €	Preço por quilómetro = 0,52 €	S1
	Valor da Taxa de Saída = 7.5 €	Valor da Taxa de Saída = 10 €	S2
	Contabilização por taxa de saída até aos 20 km	Contabilização por taxa de saída até aos 25 km	S3

Tabela 4 – Alterações propostas aos algoritmos do SGTD.

3.6. Construção e Validação do Modelo de Simulação

De modo a compreender o impacto das alterações propostas nos múltiplos objetivos e *stakeholders* do sistema, foi construído um modelo de simulação que permitisse prever o comportamento do SGTD quando sujeito às novas condições. Para este fim, optou-se por utilizar a própria aplicação do SGTD suportada pela Link Consulting, de forma a garantir a maior fiabilidade possível dos resultados das simulações. Para tal, utilizou-se um servidor

aplicacional para correr aplicação do SGTD e para aceder a uma base de dados criada, pela Link Consulting, com o intuito de ser submetida aos cenários de simulação definidos.

Quanto à amostra de dados gerada para fins de simulação, e utilizada neste trabalho, esta consistiu num universo de 28538 prestações, das quais 94.3% correspondiam a transporte do tipo múltiplo. Relativamente a casos de fidelização, constituíram ao todo 8.3% das prestações, sendo que 6.7% correspondiam a utentes fidelizados dentro da área da residência e 1.6% a utentes fidelizados fora da área. Quanto a prestações recorrentes, estas correspondiam a 98.6% das prestações, sendo que 45.39% eram referentes a prestações de hemodiálise.

Nesta fase da metodologia foram simuladas 15 simulações, cada qual com duração aproximada de 6 horas, e construídas cerca de 50 *queries* em linguagem PL-SQL, de forma a retirar os valores obtidos para os indicadores de desempenho da base de dados utilizada, através do *software SQL Developer*.

O facto de terem sido usados os algoritmos da própria aplicação do SGTD para realizar as simulações minimizou significativamente possíveis desvios do que aconteceria no caso real, credibilizando os resultados obtidos através do modelo.

4. Resultados

Antes de iniciar as simulações planeadas, obtiveram-se valores de referência, através de uma primeira simulação sem alterações nos algoritmos do SGTD (S0). Inicialmente, foram simuladas alterações individuais sobre a amostra considerada (S1-S11, ver Tabela 4), isto é, cada simulação representou apenas uma única alteração num algoritmo. Posteriormente, após análise dos resultados obtidos, realizaram-se quatro simulações combinadas (A, B, C e D), cada qual representando mais do que uma alteração. Estas últimas permitiram analisar o comportamento do sistema quando sujeito a um conjunto de alterações que alteram significativamente a lógica de funcionamento atual dos algoritmos do SGTD. Os resultados de cada simulação individual podem ser encontrados na Tabela 5. Para cada indicador de desempenho foi calculada a variação percentual relativamente ao seu valor de referência. Desta forma, foi possível analisar que simulações revelaram ter maior impacto nos objetivos e *stakeholders* do SGTD.

De forma a destacar alguns valores e a facilitar a sua posterior análise, utilizou-se o seguinte código de cores, nas tabelas de resultados: **Amarelo**: para realçar valores que revelam maior impacto no objetivo "minimização da inconveniência do utente"; **Azul**: valores que revelam maior impacto no objetivo "minimização dos custos associados ao transporte de utentes" **Verde**: valores que revelam maior impacto no objetivo "distribuição equitativa dos transportes

pelas Entidades Transportadoras candidatas"; **Vermelho**: valores que revelam maior impacto no objetivo "maximização do uso de recursos disponíveis".

Simulações	Agrupamentos				Custos						Equidade					
	(ID1) Nº de Agrupamentos Múltiplos	(ID2) Intervalo temporal médio entre a primeira e última prestação / Agrupamento (minutos)	(ID3) Nº médio de prestações / Agrupamento	(ID11) Tempo de espera médio / Agrupamento (horas)	(ID14) Totais	(ID6) Prestações contabilizadas por quilómetro	(ID7) Prestações contabilizadas por taxa de saída	(ID4) Fidelizações dentro da área de residência	(ID5) Fidelizações fora da área de residência	(ID10) Tempos de espera	(ID8) % ET's que aumentaram a média das distâncias máximas percorridas	(ID9) % ET's que diminuíram a média das distâncias máximas percorridas	(ID12) % ACeS's que aumentaram a despesa associada ao transporte	(ID13) % ACeS's que diminuíram a despesa associada ao transporte	(ID15) % ET's que aumentaram a receita associada ao transporte	(ID16) % ET's que diminuíram a receita associada ao transporte
S0	11952	26,74	2,25	3,37	597170	582545	14625	66719	20973	131730	-	-	-	-	-	-
S1	-	-	-	-	+ 1,48	+ 1,51	0,00	+ 1,42	+ 1,39	- 0,01	-	-	100,00	0,00	100,00	0,00
S2	-	-	-	-	+ 0,82	0,00	+ 33,33	+ 0,09	0,00	0,00	-	-	100,00	0,00	60,22	0,00
S3	-	-	-	-	- 0,30	- 0,59	+ 11,28	0,93	0,00	- 3,37	-	-	0,00	100,00	8,60	22,58
S4	- 4,79	+ 4,10	+ 5,03	+ 0,70	- 3,13	- 2,93	- 11,23	- 100,00	- 100,00	- 3,09	26,88	32,26	0,00	100,00	32,58	34,83
S5	- 4,19	+ 3,22	+ 4,38	+ 0,72	- 2,44	- 2,23	- 10,72	+ 17,20	- 99,37	- 2,54	31,18	27,96	0,00	100,00	32,22	35,56
S6	+ 9,30	- 3,87	- 8,51	- 0,51	+ 5,69	+ 5,71	+ 4,87	+ 10,53	+ 9,20	+ 7,57	26,88	39,78	100,00	0,00	50,00	22,22
S7	+ 7,51	- 0,50	- 6,99	+ 0,42	+ 5,07	+ 5,09	+ 4,51	+ 10,41	+ 9,02	+ 7,27	26,88	37,63	83,33	16,67	46,67	22,22
S8	- 3,24	+ 3,58	+ 3,35	+ 1,51	- 0,91	- 0,66	- 10,72	+ 0,28	- 1,51	- 0,52	26,88	22,58	16,67	66,67	27,78	27,78
S9	- 0,72	- 13,31	+ 0,72	- 0,73	- 0,26	+ 0,01	- 10,97	+ 0,22	- 1,35	- 1,31	20,43	31,18	50,00	50,00	32,22	31,11
S10	- 0,74	-12,80	+ 0,74	- 0,70	- 0,30	- 0,04	- 10,56	+ 0,44	- 1,51	- 1,22	20,43	31,18	50,00	33,33	33,33	26,67
S11	- 14,39	+ 7,02	+ 16,81	+ 2,14	- 4,41	- 3,80	- 21,85	- 4,78	+ 0,78	- 9,80	38,71	24,73	0,00	100,00	29,03	31,11

Tabela 5 – Variações percentuais dos indicadores de desempenho de cada simulação individual relativamente aos valores de referência.

Perante os resultados das simulações individuais, foram planeadas simulações combinadas, envolvendo alterações marginais ou estruturais, incluindo: A (S4+S9+S1); B (S4+S9+S11+S1); C (S6+S4+S11+S9) e D (S6+S11+S9). Os resultados obtidos encontram-se esquematizados na Tabela 6.

Simulações	Agrupamentos				Custos						Equidade					
	(ID1) Nº de Agrupamentos Múltiplos	(ID2) Intervalo temporal médio entre a primeira e última prestação / Agrupamento (minutos)	(ID3) Nº médio de prestações / Agrupamento	(ID11) Tempo de espera médio / Agrupamento (horas)	(ID14) Totais	(ID6) Prestações contabilizadas por quilómetro	(ID7) Prestações contabilizadas por taxa de saída	(ID4) Fidelizações dentro da área de residência	(ID5) Fidelizações fora da área de residência	(ID10) Tempos de espera	(ID8) % transportadoras que aumentaram a média das distâncias máximas percorridas	(ID9) % transportadoras que diminuíram a média das distâncias máximas percorridas	(ID12) % ACeS's que aumentaram a despesa associada ao transporte	(ID13) % ACeS's que diminuíram a despesa associada ao transporte	(ID15) % transportadoras que aumentaram a receita associada ao transporte	(ID16) % transportadoras que diminuíram a receita associada ao transporte
V.R.	11952	26,74	2,25	3,37	597170	582545	14625	66719	20973	131730	-	-	-	-	-	-
A	-10,19	-10,29	+11,35	-0,94	-1,85	-1,25	-25,90	-100,00	-100,00	-9,71	30,11	31,18	33,33	66,67	65,17	34,83
B	-14,27	-7,95	+16,65	-1,00	-5,48	-4,77	-33,69	-100,00	-100,00	-13,80	38,71	29,03	33,33	66,67	61,96	38,04
C	-5,19	-13,95	+5,47	-2,25	-2,12	-1,91	-10,21	-100,00	-100,00	-7,81	36,56	40,86	50,00	50,00	47,31	36,56
D	-0,64	-16,40	+0,65	-0,84	+1,74	+2,05	-10,51	+6,12	+0,53	-1,66	40,22	33,70	83,33	16,67	52,17	29,35

Tabela 6 – Variações percentuais dos indicadores de desempenho de cada simulação combinada relativamente aos valores de referência.

A partir dos resultados obtidos nas simulações de alterações aos algoritmos do SGTD foi possível analisar o seu impacto nos múltiplos objetivos e *stakeholders* do sistema. Esta análise permitiu comprovar conflitos entre objetivos do sistema e, consequentemente, entre interesses dos vários *stakeholders*, para a grande maioria de alterações propostas. As simulações que procederam à alteração dos intervalos horários máximos entre prestações do mesmo agrupamento revelaram resultados merecedores de especial atenção.

Ao testar diferentes intervalos horários máximos entre a primeira e última prestação de um agrupamento (S9 e S10), bem como os limites a partir dos quais estes se aplicam, verificou-se que a imposição de menores intervalos não se repercuta num aumento de custos. De fato, esta alteração mostra potenciar uma melhor reorganização dos agrupamentos, aproximando prestações com horários próximos, o que é inegavelmente vantajoso, tanto para os utentes, como para as transportadoras (constituindo uma melhoria *win-win*). Como tal, o ajuste dos intervalos horários máximos deve ser seriamente considerado como uma melhoria a implementar no sistema.

5. Discussão

Decorrem do trabalho desenvolvido as seguintes considerações: quanto às limitações da metodologia proposta, destaca-se o processo de definição de alguns indicadores de desempenho, nomeadamente os que dizem respeito a valores médios (ID2, ID3 e ID11). Para tirar ilações mais rigorosas sobre o comportamento do sistema, seria pertinente realizar sobre estes valores um tratamento estatístico que permitisse comprovar a sua adequação como indicadores. Uma outra limitação desta metodologia prende-se no facto das simulações terem incidido numa amostra de dados que, por ser restrita e gerada para fins de simulação, tem a desvantagem da perda de perceção do impacto da sazonalidade e de outros possíveis padrões comportamentais do SGTD.

Como principais vantagens, esta metodologia apresenta uma abordagem diferenciadora quanto ao estudo de alternativas ao funcionamento dos algoritmos de um sistema, uma vez que direciona o foco da análise de impacto para os seus *stakeholders*. Da mesma forma, o facto de fomentar uma constante interação com a empresa cuja decisão pretende apoiar, através de uma abordagem sociotécnica, leva a uma maior adequação e ajuste dos métodos às necessidades identificadas por esta, permitindo aferir e validar as análises realizadas, de forma consistente.

Assim, a metodologia desenvolvida permitiu:

- Analisar as componentes técnicas do sistema, nomeadamente o funcionamento dos algoritmos, de modo a identificar e propor a análise de alterações pertinentes;

- Analisar os múltiplos *stakeholders* do sistema, de forma a tecer recomendações sobre estratégias de comunicação a ter com cada perfil identificado;
- Identificar os diferentes objetivos do sistema e relacioná-los com os *stakeholders* cujo interesse pretendem refletir;
- Simular e analisar o impacto de alterações nos algoritmos do sistema nos seus múltiplos objetivos e *stakeholders*;
- Gerar informação que possibilite identificar oportunidades de melhoria do sistema em estudo.

Reforça-se que esta metodologia, pelas suas características e análises que possibilita, pode ser aplicada a outros sistemas que apresentam multiplicidade de objetivos e *stakeholders*, cenário frequente nas organizações/sistemas de hoje em dia. Embora se verifique uma maior sensibilização para a importância de analisar criteriosamente os diferentes objetivos e *stakeholders* de uma organização/sistema, são cada vez maiores as pressões no sentido de contenção de despesa e redução de custos. Não descurando este ponto, há no entanto que não esquecer toda a envolvência e componente social inerente a cada organização/sistema, considerando-se pertinente a utilização de metodologias como a construída neste trabalho.

6. Conclusões

Este estudo desenvolveu uma metodologia capaz de analisar alterações aos algoritmos do SGTD, com foco no impacto nos seus múltiplos *stakeholders*. A metodologia evidenciou gerar informação para apoiar a Link Consulting na análise de oportunidades de melhoria do SGTD.

7. Referências

Bowers, J., Lyons, B., & Mould, G. "Developing a resource allocation model for the Scottish patient transport service" *Operations Research for Health Care*, *1, 4* (2012), 84–94.

Cherns, A., "The Principles of sociotechnical design". *Human Relations*, 29, 8 (1976), 783–792.

Cordeau, & Laporte, "The dial-a-ride problem : models and algorithms". *Annals of Operation Research*, 153, 1 (2007), 29–46.

Fottler, M., *Challenges in health care management: strategic perspectives for managing key stakeholders*. San Francisco, CA: Jossey-Bass, 1990.

Freeman, R. E., *Strategic Manegement: A Stakeholder Approach*, Pitman Publishing, 1984.

Hanne, T., Melo, T., & Nickel, S., "Bringing robustness to patient flow management through optimized patient transports in hospitals", *Interfaces*, 131 (2007).

Keeney, R., *Value-Focused Thinking*. *A Path to Creative Decisionmaking*. Harvard University Press, 1992.

Laudon, K., & Laudon, J., *Management Information Systems*. *Managing the Digital Firm*, Prentice Hall, 2011.

Link Consulting, "Workshop de Apresentação SGTD", Julho de 2010

Melachrinoudis, E., Ilhan, A. B., & Min, H., "A dial-a-ride problem for client transportation in a health-care organization", *Computers & Operations Research*, 34 (2011) 742–759.

Mingers, J., & Brocklesby, J., "Multimethodology: Towards a Framework for Mixing Methodologies", *Omega*, *25*, 5 (1997), 489–509.

Ministério da Saúde, "Despacho n.º 7861/2011", *Diário da República de 31 de Maio*, 2ª Série, nº105, 2011.

Ministério da Saúde, "Lei nº48/90", *Diário da República de 24 de Agosto*, 1ª Série, nº 195, 1990.

Ministério da Saúde, "Portaria n.º 142-B/2012", *Diário da República de 15 de Maio*, 1ª Série, nº 94, 2012.

Phillips e. Bana e Costa, "Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing," *Annals of Operations Research*, 154, 1, (2007), 51–68.

Reed, M. S., Graves, A., Dandy, N., Posthumus, H., Hubacek, K., Morris, J.,Stringer, L. C., "Who's in and why? A typology of stakeholder analysis methods for natural resource management". *Journal of Environmental Management*, 90, 5 (2009), 1933–49.

Mobile ERP: Otimização do processo de inventário em loja

Alexandra Monteiro [1], Carlos Costa [2]

1) ISCTE – Instituto Universitário de Lisboa, Portugal

aramo@iscte.pt

2) ISCTE – Instituto Universitário de Lisboa, Portugal
carlos.costa@iscte.pt

Resumo

A introdução dos ERPs na gestão da cadeia de abastecimento veio trazer grandes melhorias a este processo, no entanto existem ainda problemas que o afetam negativamente. O processo de inventário é apontado como sendo dos principais contribuintes para a problemática. Na revisão da literatura são frequentes os casos de estudo que verificaram diferenças consideráveis entre o *stock* físico e a informação presente no sistema de informação. Neste estudo tem por objetivo melhorar o processo de inventário pelo aumento de mobilidade do ERP da organização. Mais especificamente, da atividade de registo de quantidades. Foi concebida uma solução adaptada ao contexto de estudo ambicionando o menor custo de implementação, formação e posterior manutenção da mesma. A solução consiste no desenvolvimento de uma aplicação móvel integrada com o ERP da organização que permite a utilização da camara do *smartphone* para efetuara leitura de códigos de barras e assim melhorar o registo de quantidades. Desta forma foi possível eliminar alguns passos da atividade de registo de contagens, tornar o processo mais eficiente e impulsionar o aumento de mobilidade do ERP da organização.

Palavras chave: Inventário, Mobile ERP, SCM

1. Introdução

Os sistemas integrados de gestão empresarial (ERPs) são um conceito que suporta as necessidades de uma empresa ao nível do planeamento e controlo de recursos. Este evoluiu de conceitos anteriores como o *material requirement planning* (MRP) e veio provocar alterações no paradigma das organizações [Jacobs 2007]. Este evoluiu de sistemas que desempenhavam um conjunto de tarefas independentes associadas apenas a um determinado departamento para um sistema de gestão que reúne informações e dados de todos os departamentos de uma empresa de forma integrada [Chan 2002].

A introdução dos ERPs veio melhorar significativamente os processos das organizações, nomeadamente a performance dos processos de inventário que suportam a cadeia de abastecimento, também denominada *supply chain* [Rekik Y. 2012]. Estes sistemas facilitam as

tarefas de logística como o planeamento e a previsão de necessidades de stock [Nettstrater, 2015].

Apesar desta melhoria existem ainda problemas que afetam negativamente a precisão e rigor destes processos. Em 2008, Raman e DeHoratius verificaram com base em estudos empíricos realizados em trinta e sete lojas de um retalhista de dimensão considerável que, em apenas 35% de um total de 370000 dos SKU (*Stock Keeping Unit*) o *stock* no sistema de informação coincidia com o stock físico nas lojas. Os restantes 65% apresentavam falhas relacionadas com fatores diversos como, roubos por parte de clientes, erros de reposição, mercadoria danificada não registada e roubos por parte de empregados [DeHoratius 2008].

Dada a componente humana envolvida no processo de inventário surgem também aspetos comportamentais que podem influenciar a performance deste processo. Bruccoleri descreve num estudo de 2014 a influência dos aspetos comportamentais que poderão resultar em falhas de precisão dos inventários. São apontados como principais fatores o *stress*, a motivação e a complexidade do cenário envolvido [Bruccoleri M. 2014].

Muitas organizações têm os seus processos de gestão de inventário automatizados nos quais confiam para tomar decisões críticas. Contudo, se os dados de entrada não forem precisos a capacidade do sistema de fornecer informação confiável é afetada. Para negócios em que existe muita concorrência e margens muito reduzidas, oferecer a maior disponibilidade de produto ao mínimo custo é o objetivo principal e um dos principais fatores para o sucesso da organização [Kang 2005] [Mersereau 2012] [Lee H. O. 2005].

Nos sistemas de inventário automatizados é efetuada uma previsão das necessidades da loja que resulta na criação de requisições de compra para os produtos em falta. Segundo trabalhos realizados pelo Auto-ID Center no MIT verificou-se que em média para 25% das SKUs o *stock* físico difere em seis ou mais unidades do que stock registado no sistema informação [Kang, 2005].

As diferenças de *stock* entre os sistemas de inventário e o stock físico podem resultar numa situação de rutura de stock o que pode significar uma perda considerável não só para a organização como para os seus fornecedores. Segundo Corsten e Gruen perante a ausência de *stock* para um produto 11% dos clientes desiste da compra, 16% atrasam a compra, 21% compram um produto semelhante da mesma marca, 22% compram um produto semelhante de outra marca e 31% compram o produto noutra loja [Corsten 2003]. Toma assim relevância o estudo deste processo. Identificar a origem dos problemas e criar soluções tendo em vista a melhor articulação entre os colaboradores que são peças fundamentais para a obtenção de bons resultados e os sistemas que potenciam a performance dos processos.

Falhas no processo de inventário resultam muitas vezes em problemas que podem ser altamente prejudiciais para a organização, como a rotura de *stock*.

Como resposta a este problema, será proposta a criação de uma solução móvel, integrada com o ERP da organização, para melhorar a performance e fluidez do processo de inventário.

Especificando, este objetivo poderá ser decomposto nos seguintes objetivos específicos:

1. Propor uma solução conceptual que responda aos seguintes requisitos:
    a. Possibilitar a utilização de dipositivos móveis para a interação entre colaboradores e ERP da organização;
    b. Otimizar o processo de inventário;
2. Desenvolvimento de uma aplicação móvel com integração com o ERP da organização;
    a. Estudo da envolvente tecnológica que consiste na avaliação das tecnologias existentes, a escolha da tecnologia a utilizar para o desenvolvimento do protótipo, assim como a arquitetura da solução a desenvolver.

Na próxima secção é apresentada uma breve revisão da literatura. Em seguida é apresentada a solução proposta e descrita a sua implementação. Por fim, são apresentados os resultados preliminares do estudo.

2. Revisão da Literatura

A constituição de um ERP pode ser vista como um sistema aberto relativamente à sua estrutura uma vez que este permite a adaptação do sistema de informação aos processos específicos da organização. Assim, consoante o negócio que suporta, podem ser desenvolvidas melhorias que poderão potenciar a fluidez dos processos [Olson 2013].

Relativamente à componente de *hardware* do processo de inventário tem também havido esforços para a melhoria deste processo. A introdução da tecnologia RFID (*Radio Frequency Identification*) tem tido a atenção da comunidade de investigação nos últimos anos. Esta é apontada como sendo uma tecnologia que pode ajudar as organizações na automatização do processo de identificação e assim eliminar possíveis fontes de erro [McFarlane 2003] [Bensoussan 2007] [Kang 2005] [Hardgrave 2013].

A necessidade de automatização e melhor identificação levou a que a tecnologia RFID fosse considerada para o contexto de SCM (*Supply Chain Management*). No entanto, os códigos de barras continuam a ser amplamente utilizados. O principal motivo está relacionado com o custo financeiro associado ao RFID que é muito superior relativamente aos códigos de barras

[Preradovic 2008]. Também a ausência de Standards que regulem esta tecnologia são um fator a considerar. Foi criada a EPCglobal's com o objetivo de estandardizar a tecnologia no entanto esta ainda não tem o apoio da *International Organization for Standardization* (ISO).

Em contextos de pequenas lojas os requisitos para a implementação do RFID podem inviabilizar a utilização desta tecnologia. Este tipo de contexto com necessidades muito específicas requer soluções adaptadas aos mesmos.

Neste estudo pretende-se investigar aplicar o conceito de Mobile ERP ao processo de inventário efetuado em loja. Deste modo pretende-se tirar proveito da utilização de dispositivos móveis como os *smartphones* para simplificação do processo.

Como resultado do aumento da utilização de *smartphones* o desenvolvimento de produtos e serviços adaptados a estas plataformas veio ganhar expressão [Hsu et al. 2014]. Um exemplo deste facto é a introdução dos dispositivos móveis no contexto dos ERPs.

2.1. ERPs na otimização de processos

O termo *Entreprise Resource Planning* surgiu na década de 1990 pela consultora Gartner Group [Wylie 1990]. Segundo a qual os ERPs são sistemas que possibilitam a integração das várias áreas de negócio de uma organização (vendas, produção, compras, finanças, recursos humanos) pela implementação de módulos que representam cada área. Deste modo, é aumentada a partilha de informação por toda a estrutura da organização centralizando toda a informação numa base de dados comum [Samaranayake 2009] [Kelle 2005].

Segundo Fawcettet os ERPs são considerados impulsionadores de integração e partilha de informação em tempo real [Fawcett 2011]. Estes foram desenvolvidos a partir de sistemas planeamento de necessidades de material, ou MRP (*Material Requirements Planning*) e de sistemas de planeamento de recursos de manufatura (MRP II) para responder à necessidade de integração financeira para estes processos [Jacobs R. W. 2007].

Relativamente à componente operacional os ERPs são apontados como um meio para reduzir o tempo de ciclo e consequentemente num melhor apoio ao cliente e gestão da relação com fornecedores [Hassabelnaby 2012]. Segundo Bendoly a cultura que contextualiza a organização é também um fator determinante para o sucesso da implementação do sistema. Num estudo que compara a implementação de um ERP entre os Estados Unidos e na China verificou-se que a interdependência de tarefas representa um fator determinante no posterior desempenho do sistema [Bendoly 2006]. Verificou-se que quando devidamente alinhado com os processos de

negócios, a implementação do ERP levava à redução dos custos operacionais e melhorava a produtividade [Su 2010].

2.2. Aquisição de dados

O processo de inventário inicia-se com o Registo de Quantidades. O Registo de Quantidades consiste na identificação da matéria-prima e é efetuada com recurso a tecnologia de identificação e recolha de dados. Exemplos deste tipo de tecnologia são o caso do código de barras e mais recentemente utilizado em SCM o RFID [Qian 2012].

Os códigos de barras são um método de identificação amplamente utilizado na indústria para controlo do inventário [Froschle 2009]. Relativamente à tecnologia RFID estes convergem dado que se tratam ambos de tecnologias permitem a identificação de materiais para posterior integração num sistema de informação. No entanto, tecnologicamente apresentam características muito diferentes. O RFID pode ser lido a uma distância de 100 m, têm uma elevada capacidade de leitura (100 tags/s) e permite o armazenamento de dados até 4 kB em tags passivas e 1MB em tags ativas (que possuem bateria) [Qian 2012]. Quanto aos códigos de barras é obrigatório o contanto visual e implica uma distância de centímetros entre o identificador e o código de barras. Apesar disso, o RFID não pode ser considerado um substituto dado que os códigos de barras se adequam melhor a determinados contextos, sendo o seu principal benefício o custo que é consideravelmente baixo relativamente ao RFID [Preradovic 2008].

2.3. Integração de aplicações empresariais

A disponibilização de aplicações integradas com os sistemas empresariais trata-se uma oportunidade, para as empresas, de melhorar e tornar mais eficientes os seus processos. Pela integração, é fortalecida a base de comunicação entre fornecedores e organização tornando-a mais eficiente.

A integração entre as novas soluções estratégicas e os sistemas já existentes é efetuado por meio de um *middleware* [Hasselbring 2000]. Assim podemos identificar três principais níveis a assegurar para uma integração consistente (Figura 1).

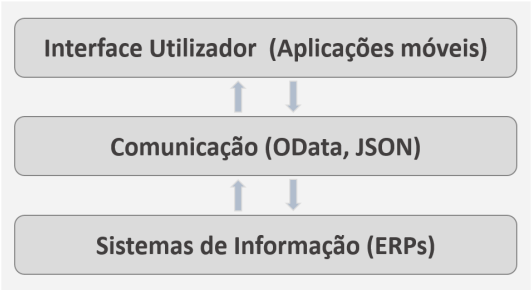


Figura 1 - Níveis de Integração para uma solução móvel integrada com ERP [Sutherland et al. 2002].

A um nível mais elevado, temos a interface com o utilizador. Atualmente, uma grande percentagem dos clientes possui dispositivos móveis, como *smartphones* que suportam aplicações web. Um dispositivo móvel, permite que o colaborador circule pela loja e utilize a aplicação aumentando a mobilidade da solução.

O segundo nível de integração prende-se com a comunicação com o servidor. Esta comunicação pode ser efetuada segundo um conjunto de modelos de comunicação como OData, XML ou JSON.

Por fim, temos os sistemas fonte que alimentam as aplicações como os ERP ou outras aplicações empresariais [Sutherland et al. 2002].

3. Proposta de solução

A solução para a otimização do processo de inventário em loja consiste no aumento de mobilidade do ERP através de tecnologia móvel.

Para uma melhor compreensão do contexto a que a solução se destina começamos por efetuar uma descrição dos requisitos da mesma:

- Lojas que pequena dimensão;
- Que permitam o acesso à internet por Wi-Fi;
- Matéria-prima identificada com tecnologia descodificada por leitura ótica (códigos de barras, QR Code).

Neste caso de estudo a organização a que se destina a solução atua na área das telecomunicações tratando-se os telemóveis de uma das principais matérias-primas deste negócio.

O processo de inventário é composto por várias atividades que são especificadas no diagrama de casos de uso (Figura 2).

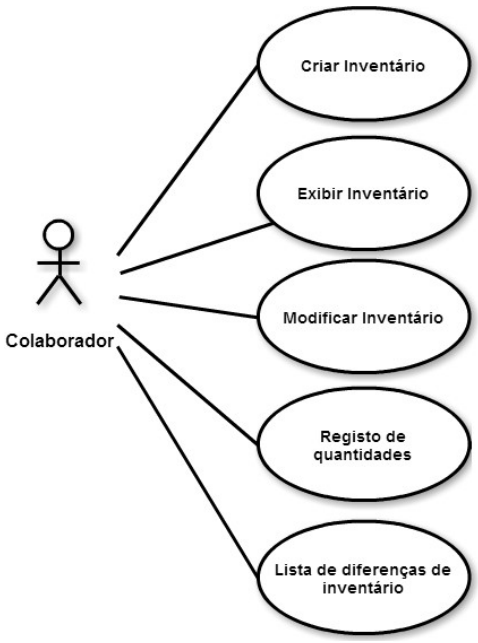


Figura 2 – Diagrama de casos de uso para o processo de inventário.

Este estudo foca-se na implementação atividade de registo de quantidades como primeiro passo para implementação do processo completo de inventário.

Através da aplicação móvel, o colaborador acede ao detalhe do inventário e efetua a contagem da matéria-prima que é automaticamente integrada com o ERP da organização (Figura 3).

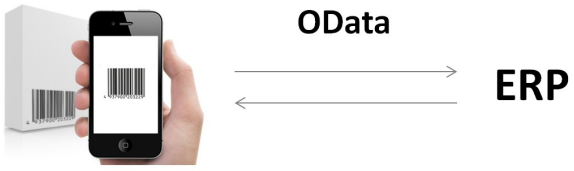


Figura 3 - Esquematização da proposta de solução.

Iterativamente, é lido o identificador do material, a informação é validada no ERP e de seguida atualizada na aplicação. No final na leitura dos materiais em loja, a atividade é concluída e é submetido o registo de quantidades no ERP.

Tecnologicamente, a solução implementada consiste numa aplicação móvel híbrida, integrada com o ERP da organização que utiliza a câmara do dispositivo para efetuar a leitura do identificador do material. Optou-se pela escolha de uma aplicação híbrida pela característica multiplataforma deste tipo de aplicações. É reduzido o custo e tempo de implementação e ao mesmo tempo desenvolvida uma solução que facilmente se adapta a outros sistemas operativos.

Em suma, a solução consiste numa interface simplificada para o utilizador que o encaminha de forma eficiente pelo processo de Registo de Quantidades.

4. Implementação

Para a implementação da solução foi efetuado o estudo da envolvente tecnológica com o objetivo de determinar qual a melhor se adaptava as necessidades do contexto de estudo. Foram abordados os tipos de aplicações móveis e as *frameworks* de desenvolvimento.

Dado o reduzido custo de implementação e necessidade de comunicação com os recursos do *smartphone*, optou-se pela implementação de uma aplicação híbrida. Assim, foi desenvolvida uma componente de interface com o utilizador que é executada num *browser* e utilizadas APIs para a comunicação com a camara do *smartphone* através da qual são lidos os códigos de barras.

Para a escolha da *framework* de desenvolvimento optou-se pelo SAPUI5 por se tratar da *framework* disponibilizada pela SAP, otimizada para a comunicação com o mesmo. A comunicação entre os dois sistemas é efetuada através de serviços OData.

A implementação iniciou-se pela criação de um serviço OData em SAP que recebe o número de inventário e através de um módulo de função retorna a informação sobre o inventário inserido.

Quanto à interface com o utilizador foram desenvolvidos dois ecrãs. O primeiro permite a introdução da chave do documento inventário que é composta por número de inventário e exercício (Figura 4).

Figura 4 - Tela 1.

O segundo ecrã apresenta o detalhe do inventário. Nomeadamente, os itens do mesmo que consistem num material, descrição e quantidade registada (Figura 5).

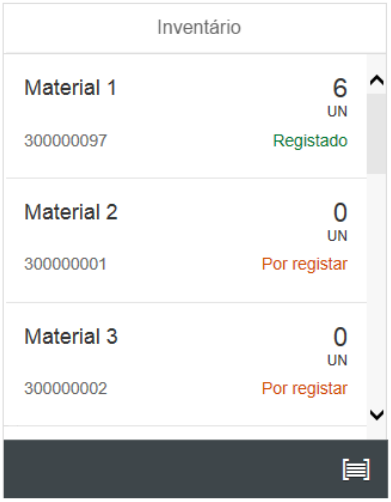


Figura 5 - Tela 2.

Para efetuar a leitura dos códigos de barras é selecionada uma linha que ativa automaticamente o modo de leitura de código de barras. O material é lido, verificado em SAP e incrementado o valor do registo no ecrã. Depois de lidos todos os materiais, o registo de quantidades é integrado em SAP e é finalizada a atividade.

5. Avaliação preliminar

Neste ponto será efetuada uma análise comparativa entre o fluxo do processo antes e depois da sua otimização com o objetivo de identificar os benefícios da implementação da solução concetual proposta.

A atividade de registo de contagem antes da otimização implicava os seguintes passos:

1. Deslocação dos materiais para junto do dispositivo de leitura ótica que está ligado por cabo ao computador;
2. Leitura dos códigos de barras para um ficheiro;
3. Importação do ficheiro e integração de dados no ERP.

Após a otimização do processo foi eliminado o ponto 1 e reformulados os pontos 2 e 3. Ou seja, deixou de ser necessário a deslocação dos materiais para junto do computador dado que o dispositivo móvel permite a deslocação até junto dos mesmos. Através da aplicação o colaborador seleciona o material que pretende contar e inicia a leitura dos códigos de barras com o dispositivo móvel. Em caso de erro a verificação é efetuada automaticamente o que possibilita o tratamento do mesmo sem que seja necessário a repetição dos pontos 1 e 2 (Figura 6).

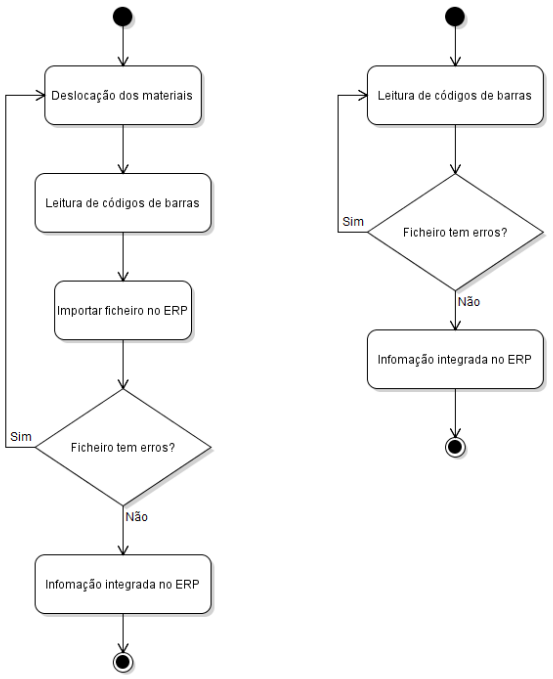


Figura 6 - Comparação do processo de registo de contagem. À direita fluxo para o processo antes da otimização, à esquerda o fluxo para o processo já otimizado.

Outra vantagem a apresentar é a adoção de um ambiente simplificado, com baixo custo de adaptação. Não só pelo *software* mas também pelo hardware utilizado com que os colaboradores já estão familiarizados o que reduz a necessidade de formação para a utilização de ERPs. Ao contrário dos *smartphones* os leitores de códigos de barras são específicos para uma funcionalidade o que limita a utilização dos mesmos para outros fins. Nesse sentido, a utilização dos *smartphones* também se apresenta como uma vantagem. Posteriormente o conceito poderá ser aplicado a outros processos da organização que impliquem mobilidade sem a necessidade de aquisição de novo *hardware*.

Quanto à manutenção da aplicação foi projetada para ter o menor custo, tendo-se optado por uma aplicação híbrida em que grande parte do código pode ser reutilizada por qualquer sistema operativo móvel.

Na opinião dos utilizadores a aplicação veio simplificar a atividade de registo de quantidades. É fácil de utilizar, intuitiva e permite a execução da atividade de forma rápida. A melhorar pode ser considerada a utilização de um dispositivo auxiliar, acoplado no *smartphone* com o objetivo de melhorar a leitura dos códigos de barras.

6. Conclusão

A mobilidade dos ERPs veio dotar as organizações de vantagens significativas ao nível da otimização de processos. A possibilidade de aceder em tempo real ao ERP da organização permite a reação imediata às adversidades encontradas. O desenvolvimento dos dispositivos móveis como os *smartphones* combinado com a disponibilidade de acesso à internet de qualidade veio possibilitar o aumento de produtividade, redução de custos e aumento de satisfação dos clientes.

Neste estudo procurou-se responder à necessidade de otimização do processo de inventário no contexto específico de lojas de dimensão reduzida. Pela análise do contexto, foi possível identificar pontos a melhorar no processo e conceber uma solução para a problemática.

Em síntese, foi possível compreender que os dispositivos móveis como os *smartphones* podem ser utilizados como ferramentas para a otimização de processos. Com um custo de implementação reduzido foi possível desenhar e desenvolver uma solução enquadrada no contexto atual que possibilitou a otimização do processo de inventário.

Futuramente este conceito poderá estender-se a outros processos da organização tirando partido da multifuncionalidade dos *smartphones*.

7. Referências

Bendoly, E., Bachrach, D.G., Wang, H., Zhang, S., “ERP in the minds of supervisors”, *International Journal of Operations & Production Management, 26*, 5, (2006), 558 - 578.

Bensoussan, A., Çakanyildirim, M., “Economic Evaluation of Systems that Expedite Inventory Information”, *Production and Operations Management, 16*, 3 (2007), 360–368.

Bruccoleri M., Canella, S., La Porta, G., “Inventory record inaccuracy in supply chains: the role of workers’ behavior”, *International Journal of Physical Distribution & Logistics Management*, 44, 10 (2014), 796 - 819.

Chan, J., Burns, N., “Benchmarking manufacturing planning and control (MPC) systems: an empirical study of Hong Kong supply chains”, *Benchmarking: An International Journal, 9*, 3 (2002), 256–277.

Corsten, D., Gruen, T., “Desperately seeking shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks”, *International Journal of Retail & Distribution Management, 31*, (2003), 605–617.

DeHoratius, N., Raman, A., “Inventory Record Inaccuracy: An Empirical Analysis”, *Management Science, 54*, 4 (2008), 627-641.

Fawcett, S. E., Wallin, C., Allred, C., Fawcett, A.M., Magnan, G.M., “Information technology as an enabler of supply chain collaboration: a dynamic-capabilities perspective”, *Journal of Supply Chain Management, 47*, 1 (2011), 38-59.

Froschle, H. -K., Gonzales-Barron, U., McDonell, K., Ward, S, “Investigation of the potential use of e-tracking and tracing of poultry using linear and 2D barcodes”, *Computers and Electronics in Agriculture*, 66, 2 (2009), 126–132.

Hardgrave, B. C., Aloysius, J. A., Goyal, S.C., “RFID-Enabled Visibility and Retail Inventory Record Inaccuracy: Experiments in the Field”, *Production and Operations Management*, 22, 4 (2013), 843–856.

Hassabelnaby, H. R., Hwang, W., Vonderembse, M. A., “The impact of ERP implementation on organizational capabilities and firm performance.”, *Benchmarking: An International Journal*, 19, 4/5 (2012), 618 - 633.

Hasselbring, W., “Information System Integration”, *Communications of the ACM*, 43, 6 (2000), 32-38 .

Holder, E. G., “Native Execution of Mobile Applications Across Platforms”, *Bachelor of Science Honors Thesis*. Blacksburg, Virginia. (2014)

Hsu, C.; Lin, J. C., “What drives purchase intention for paid mobile apps? – An expectation confirmation model with perceived value”, *Electronic Commerce Research and Applications*, (2014)

Jacobs, F., Weston, J., “Enterprise resource planning (ERP): a brief history”, *Journal of Operations Management*, 25, 2 (2007), 357–363.

Kang, Y., Gershwin, S. B., “Information inaccuracy in inventory systems: stock loss and stockout.”, *IIE Transactions.* 37 (2005), 843–859.

Kelle, P., Akbulut, A., “The role of ERP tools in supply chain information sharing, cooperation, and cost optimization”, *International Journal of Production Economics*, 93 (2005), 41–52.

Lee, H.L., Ozer, O., “Unlocking the value of RFID”, *Working Paper, Standford University, Graduate School of Business*, (2005)

McFarlane, D., Sheffi, Y., “The impact of automatic identification on supply chain operations.”, *International Journal of Logistics Management*, 14, 1 (2003), 1-17.

Mersereau, A. J., “Information-Sensitive Replenishment when Inventory Records Are Inaccurate.”, *Production and Operations Management Society*, 22, 4 (2012), 792–810.

Nettstrater, A., Geißen, T., Witthaut, M., Ebel, D., Schoneboom, J., “Logistics Software Systems and Functions: An Overview of ERP, WMS, TMS and SCM Systems.”, Springer International Publishing, (2015).

OData. (s.d.). *Overview*. Obtido em 04 de Janeiro de 2015, de Documentation - OData Version 2.0: http://www.odata.org/documentation/odata-version-2-0/overview

Olson, D. L., Chae. K., Sheu C., “Relative impact of different ERP forms on manufacturing organisations: an exploratory analysis of a global manufacturing survey”, *International Journal of Production Research*, 51, 5 (2013), 1520–1534.

Preradovic, S., Balbin, I., Karmakar, N. C., Swiegers, G., “A Novel Chipless RFID System Based on Planar Multiresonators for Barcode Replacement”, *International Conference on RFID*, (2008), 289-296

Qian, J., Yang, X., Wu, X., Zhao, L., Fan, B., Xing, B., “A traceability system incorporating 2D barcode and RFID technology for wheat flour mills”, *Computers and Electronics in Agriculture*, *89*, (2012), 76–85.

Rekik Y., Sahin E., “Exploring inventory systems sensitive to shrinkage – analysis of a periodic review inventory under a service level constraint”, *International Journal of Production Research*,, 50, (2012), 3529–3546.

Samaranayake, P., “Business process integration, automation, and optimization in ERP: Integrated approach”, *Business Process Management Journal*, 15, 4 (2009), 504 - 526.

SAP AG. (s.d.). *UI Development Toolkit for HTML5 (SAPUI5)*. Obtido em 04 de Janeiro de 2015, de Description: https://prezi.com/ssfllao79vp_/sapui5-overview/

SAP SE. “SAPUI5 Developer Guide for SAP HANA”, (2014).

Schubert, H., “How to Write Testable SAPUI5 Applications.” (2013). Obtido em 04 de Janeiro de 2015, de SAP Community Network: http://scn.sap.com/community/developer-center/front-end/blog/2013/03/19/how-to-build-testable-sapui5-applications

Su, Y., Yang, C., “A structural equation model for analyzing the impact of ERP on SCM.”, *Expert Systems with Applications*, 37, 1 (2010), 456–469.

Sutherland, J., Heuvel, W. V., “Enterprise Application Integration and Complex Adaptive Systems”, *Communications of the ACM*, 54, 10 (2002), 59-64 .

Wylie, L., “A vision of the next-generation MRP II.”, *Scenario*, (1990), 300-339.

Modelação de Processos de Gestão de Recursos Humanos

Sebastião Rodrigues Pedro, Vítor Santos
NOVA IMS - Universidade Nova de Lisboa, Portugal
m2013144@novaims.unl.pt
vsantos@novaims.unl.pt

Resumo

Este artigo tem como objetivo propor modelos BPMN que possam representar as boas práticas em gestão recursos humanos, a fim de facilitar a identificação dos principais requisitos dos pacotes de *software* para a Gestão de Recursos Humanos.

Assim, foram estudadas as subáreas: recrutamento e seleção, avaliação de desempenho, formação e desenvolvimento, cálculo e pagamento de salário e recompensas, e a gestão de carreira.

Para o desenho dos modelos BPMN foi adotada a metodologia *design science*, adequada aos projetos de investigação em tecnologias e sistemas de informação que têm como objetivo o design de artefactos a partir de bases teóricas.

Os modelos desenvolvidos foram disponibilizados para avaliação por gestores e executivos da área de gestão de recursos humanos o que permitiu refinar e melhorar os modelos.

Palavras-chave: Modelação de Processos; Gestão de Recursos Humanos; Sistemas de Informação

1. Introdução

Nos últimos anos tem-se notado um crescimento significativo no uso das tecnologias de informação visando automatizar os processos de gestão das organizações, pois é assumido que este investimento pode ajudar a atingir objetivos estratégicos como a excelência operacional, melhor tomada de decisão, vantagens competitivas e sobrevivência. Por esta razão são cada vez maiores os investimentos feitos em sistemas e tecnologias de informação nesta área.

Por um lado a indústria de *software* deparou-se desde sempre com a questão de conseguir entender as necessidades dos utilizadores, por outro, os utilizadores sempre tiveram dificuldades em expor aquilo que pretendiam obter com o *software*.

As empresas que desenvolvem produtos de *software* visando um amplo mercado consumidor, geralmente com clientes distribuídos em várias localidades, sofrem dificuldades para atender satisfatoriamente clientes com diferentes realidades organizacionais e requisitos específicos. Ademais, diante da crescente procura por produtos de qualidade e da crescente exigência dos mercados consumidores, empresas produtoras de *software* têm reconhecido a importância de programas de melhoria do processo de desenvolvimento de *software* (Alves, 2012).

O presente artigo tem como finalidades a identificação e modelação em notação BPMN os principais processos de negócio das principais áreas associadas à Gestão de Recursos Humanos com recurso à plataforma de modelação Bizagi, com o intuito de auxiliar, por comparação, a critica e melhoria de processos de Recursos Humanos existentes em diferentes organizações e auxiliar os produtores de *software* da área de Gestão de Recursos Humanos, apresentando modelos que traduzam as boas

práticas da área, e dessa forma facilitar a identificação dos principais requisitos dos pacotes de *software* afetos à Gestão de Recursos Humanos.

2. *Software* de Gestão de Recursos Humanos

Segundo Bell (1996) a maioria dos levantamentos de requisitos feitos são incompletos, inconsistentes e ambíguos, e consequentemente, é inevitável o impacto negativo na qualidade do *software*. Este impacto torna-se significante visto que todas as atividades do processo de desenvolvimento de *software*, principalmente as etapas iniciais, são dependentes da participação e entendimento das reais necessidades dos clientes e/ou utilizadores finais, e que o mau entendimento dos requisitos é frequentemente um fator de insucesso, atrasos e custos adicionais durante o desenvolvimento do *software*.

De acordo com Davis (1993), citado por Kujala (2002), foram analisados os resultados de três estudos empíricos, indicando que pode ser até 200 vezes mais caro detetar e reparar erros na fase de manutenção em comparação com detetar e reparar durante a fase de análise de requisitos. O autor argumenta ainda que o impacto de erros nos requisitos é substancial:

- O *software* resultante pode não satisfazer as reais necessidades dos utilizadores;
- Múltiplas interpretações de requisitos podem causar desentendimentos entre clientes e desenvolvedores e, talvez, resultando em ações judiciais;
- Tanto tempo e dinheiro podem ser desperdiçados na construção de um sistema ineficaz.

Entretanto, os produtores de pacotes de *software* para Gestão de Recursos Humanos também têm enfrentado dificuldades em entender por parte dos utilizadores, as reais necessidades dos sistemas que eles necessitam.

Por outro lado, os utilizadores de tais *software*, por sua vez, também não têm conseguido explicar com sucesso aos criadores dos programas, o que eles realmente esperam que o *software* faça.

3. Modelação de Processos em BPMN

Para melhor compreensão das atividades de Gestão de Recursos Humanos o recurso à modelação de processos em BPM (Business Process Management) pode constituir uma grande ajuda, uma vez que permite executar a modelação de processos, através de uma notação gráfica bem definida. A notação mais utilizada é o BPMN (Business Process Management Notation), que segundo Gallo (2012), define um conjunto de objetos gráficos que representam as principais características de um processo de negócio, como atividades e fluxos de controlo que definem como o processo organizacional funciona. Segundo Weske, (2007), o BPMN 2.0 tem como principal objetivo disponibilizar uma notação que seja facilmente compreensível por todos os utilizadores de negócio, desde os analistas de negócio que criam os desenhos iniciais dos processos, até aos técnicos responsáveis pela implementação da tecnologia que suportará esses processos e, finalmente, pelas pessoas que irão gerir e monitorizar esses processos.

Estes processos são geralmente criados por analistas de negócio e/ou gestores.

4. Desenho de Processos de Gestão de Recursos Humanos

A figura 1 ilustra o processo de recrutamento e seleção que na visão de Chiavenato (1992), é um conjunto de procedimentos que visa atrair candidatos potencialmente qualificados e capazes de ocupar cargos dentro da organização, ao passo que a seleção, visa classificar e escolher os candidatos que melhor se adaptam às necessidades da empresa, bem como do cargo. Neste processo existe o envolvimento do departamento/área de recrutamento e seleção, do gestor de contratação, da contabilidade, bem como dos candidatos, que são os responsáveis pela realização das diversas atividades/tarefas que constituem o processo.

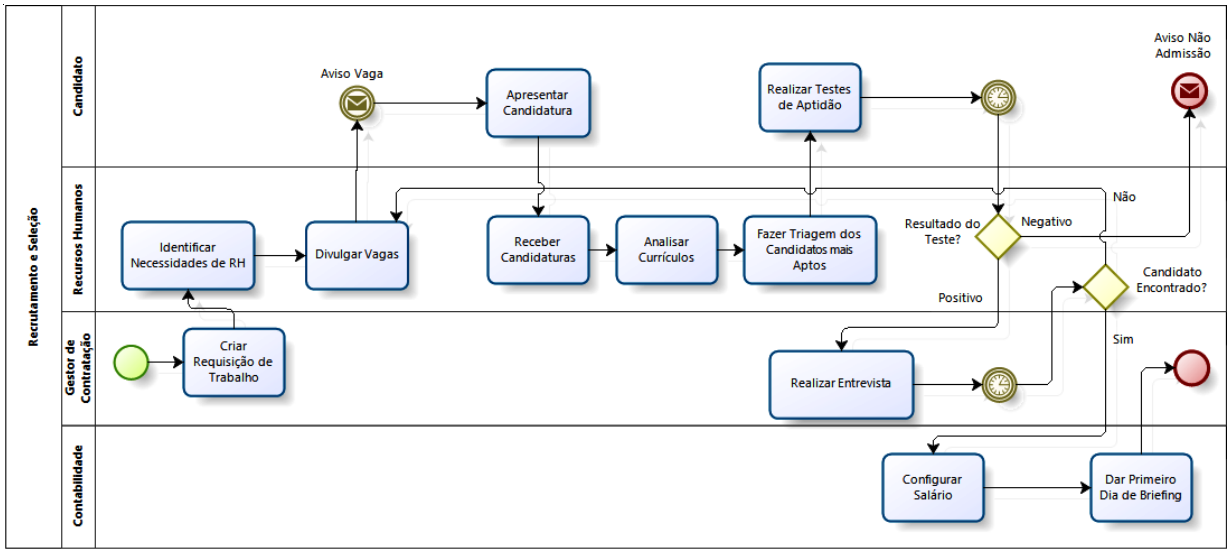


Figura 1 – Recrutamento e Seleção

A figura 2 apresenta o processo de avaliação de desempenho que segundo Fourgous & Lamber (1991), cit. por Guerreiro (2010), pode ser definida como o processo sistemático, através do qual as organizações avaliam a performance dos trabalhadores, relativamente ao cargo que ocupa. A partir da figura podemos verificar que neste processo há o envolvimento do departamento de recursos humanos, do responsável pelo funcionário, bem como do próprio funcionário que está sendo avaliado.

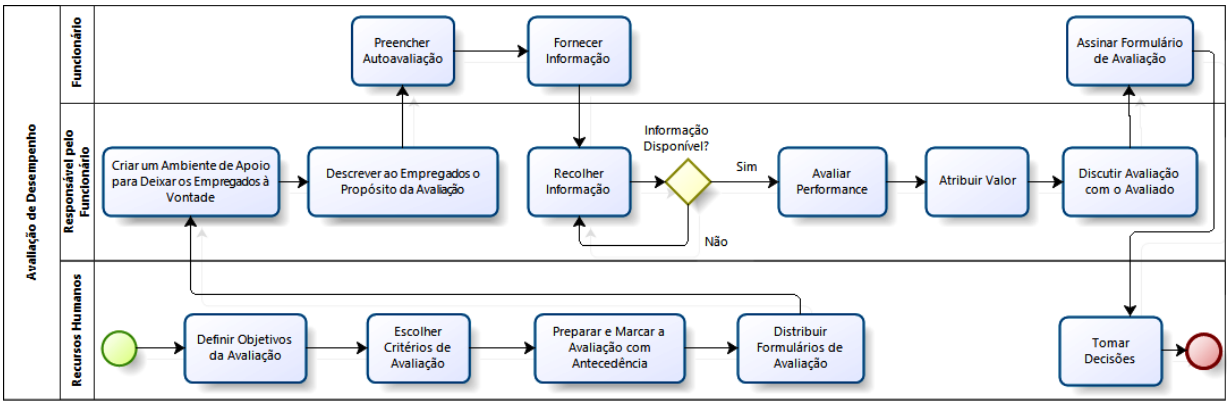


Figura 2 – Avaliação de Desempenho

A figura 3 ilustra as atividades e os intervenientes (atores) inerentes ao processo de Formação e Desenvolvimento que de acordo com a DGAEP (2015), é o nome que se dá ao processo global e permanente pelo qual os trabalhadores, através da aquisição e desenvolvimento de competências, se preparam para o exercício de uma atividade profissional ou para a melhoria do seu desempenho.

Ademais, ainda segundo o autor, deve haver um documento devidamente aprovado que deverá conter: a identificação das necessidades de formação; a caracterização das ações de formação; a identificação do conteúdo programático; o calendário das ações de formação; a identificação dos formadores, bem como as entidades formadoras; a indicação dos custos das ações de formação; a definição dos objetivos a serem alcançados em cada ação de formação; a definição dos critérios e grelha de indicadores de avaliação e validação da formação.

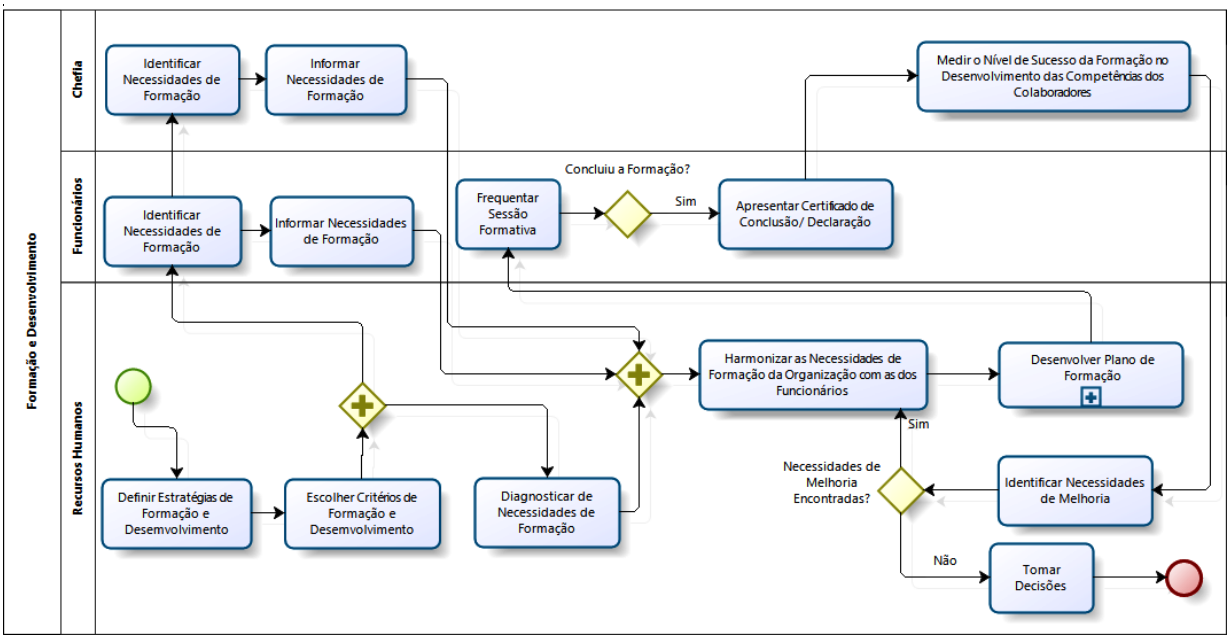


Figura 3 – Formação e Desenvolvimento

As fases da elaboração do plano de formação (subprocesso) são apresentadas na figura 3.1.

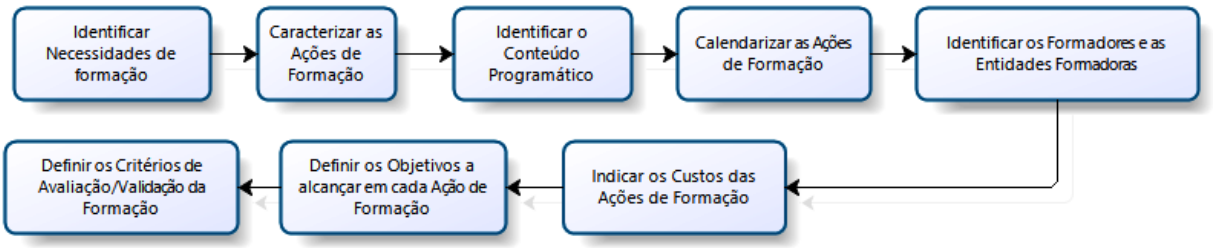


Figura 3.1 – Plano de Formação

A figura 4 representa o processo de Gestão de Carreiras, que segundo Ferreira (2007), pode ser visto como sendo uma sequência de experiências obtidas ao longo da vida, dentro ou fora do contexto organizacional e de trabalho, valorizando-se as mudanças nos desejos, conceções e posturas dos indivíduos perante a sua vida profissional. Na figura 4 são apresentados as tarefas necessárias e suficientes a fim de se levar a cabo uma gestão de carreira eficaz, à luz das boas práticas encontradas nas diversas bibliografias disponíveis, bem como os pareceres de gestores afetos a área gestão de Recursos Humanos.

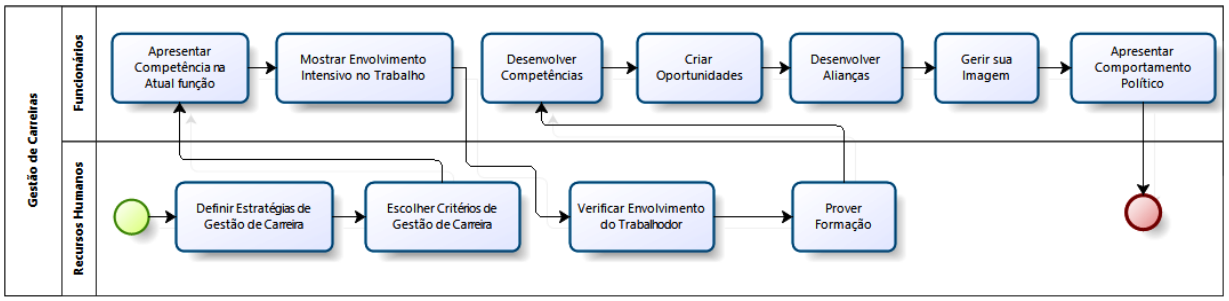


Figura 4 – Gestão de Carreiras

A figura 5 esquematiza as fases que envolvem o processo de cálculo e pagamento de salários, à luz das boas práticas identificadas na revisão da literatura. Nunes (2015), definiu o processo supracitado como sendo o procedimento que visa compensar um trabalho prestado. Outrossim, define que o salário é o rendimento que os trabalhadores recebem em troca do trabalho que despendem no processo produtivo, ou seja, a compensação periódica como forma de pagamento do trabalho.

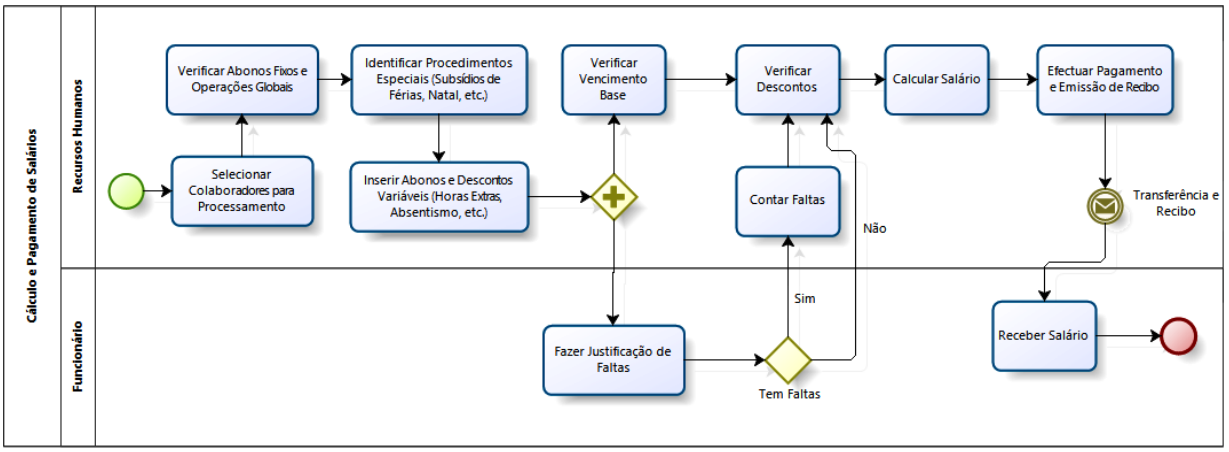
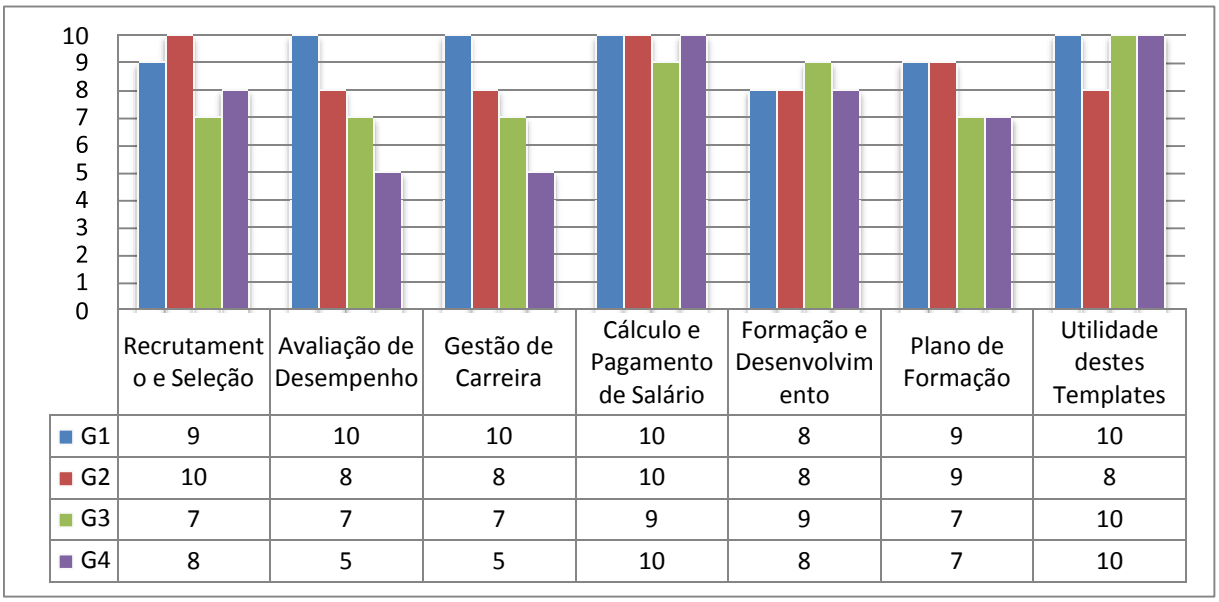


Figura 5 – Cálculo e Pagamento de Salário

5. Conclusões e Trabalho Futuro

Neste trabalho abordamos a modelação de processos das principais áreas de gestão de recursos humanos. Após a revisão da literatura, foi possível fazer uma análise dos vários processos existentes na área gestão de recursos humanos, identificando os principais processos e as melhores práticas que serviram de base para a criação dos modelos propostos.

A tabela 1 apresenta os pareceres dos gestores de recursos humanos, obtidos através do recurso a questionários relativos à utilidade dos modelos desenhados.



	Recrutamento e Seleção	Avaliação de Desempenho	Gestão de Carreira	Cálculo e Pagamento de Salário	Formação e Desenvolvimento	Plano de Formação	Utilidade destes Templates
G1	9	10	10	10	8	9	10
G2	10	8	8	10	8	9	8
G3	7	7	7	9	9	7	10
G4	8	5	5	10	8	7	10

Tabela 1 – Parecer dos gestores de recursos humanos

A tabela reflete os pareceres dos gestores de recursos humanos relativamente ao ajuste dos modelos BPMN à situação particular das suas organizações no que respeita às atividades de: Recrutamento e Seleção; Avaliação de Desempenho; Gestão de Carreira; Cálculo e Pagamento de Salário; Formação e Desenvolvimento; e Plano Anual de Formação. Após a análise das respostas dos gestores de recursos humanos de diferentes organizações, concluímos que o modelo proposto poderá funcionar como facilitador para os produtores de pacotes *software*, no que toca a interpretação das reais necessidades dos sistemas de gestão de recursos humanos.

Todavia, durante o desenvolvimento deste trabalho, foram aparecendo áreas sujeitas a melhoramento, algumas em virtude à limitação temporal do projeto de investigação, acabaram por ser descuradas, outras porque se achou que já não faziam parte do trabalho realizado. Assim, como trabalho futuro almejamos que:

- Os modelos propostos sejam apresentados a um número maior de gestores e funcionários afetos a gestão de recursos humanos, a fim de se detetar falhas não identificadas até a presente data;
- Se execute séries de simulações/testes sobre os processos apresentados, com o auxílio de um *software*, a fim de se e analisar os respetivos resultados, visando determinar os pontos de falha e apresentar soluções para os mesmos;

Por outro lado, acreditamos que se houver maior apoio por parte das entidades responsáveis pela gestão de recursos humanos, provavelmente se refletirá numa maior efetividade dos modelos.

6. Referências

Archer, L.B. (1984). Systematic method for designers. Developments in design methodology. Association of Business Process Management Professionals. ABPMP. (2009). Guia para o Gerenciamento de Processos de Negócio Corpo Comum de Conhecimento (BPM CBOK). (Versão 2.0). Chicago.

Bell, E. & Thayer, A. (1996). Software Requirements: Are They Really a Problem Second International Conference on Software Engineering. San Francisco.

Chiavenato, I. (1992). Recursos Humanos. Edição Compacta. São Paulo: Atlas.

Chiavenato, I. (1999). Gestão de Pessoas: O novo papel dos recursos humanos nas organizações (6ª ed.) Rio de Janeiro: Campus S.A.

DGAEP, Ministério das Finanças. (2015), Formação profissional e desenvolvimento de competências conceitos, obtido de //www.dgaep.gov.pt/stap/infoPage.cfm?objid=afb25c9d- ac5a-42a4-8edd-ae45e2dd0f51&KeepThis=true&TB_iframe=true&height=580&width=520 (07 de Janeiro de 2015).

Ferreira, A. P. V.G. (2007). Conteúdo e quebra do contracto psicológico e comportamentos individuais de gestão de carreira (Tese de Doutoramento em Ciências Empresariais), Universidade do Minho, obtido de https://repositorium.sdum.uminho.pt/bitstream/1822/7818/1/TESE%20VERS%C3%83O%20FINAL_Ana%20Paula%20Ferreira.pdf (08 de Janeiro de 2014).

Gallo, J. (2012). Comparativo entre As Versões 1.2 E 2.0 da Notação BPMN a sua Aplicação em Diagramas de Processos de Negócios, (Tese de Pós Graduação), Universidade Tecnológica Federal do Paraná – UTFPR, obtido de http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/1112/3/MD_ENGESS_I_2012_11.pdf (20 de Abril de 2014).

Guerreiro, S., R., O., L. (2010). Analise de um Processo de Avaliação de Desempenho "O caso da Sumolis", obtido de http://bdigital.ufp.pt/bitstream/10284/1013/1/Monografia.pdf (07 de Janeiro de 2015).

Kujala, S. (2002). User Involvement: A Review of the Benefits and Challenges. In Soininen, T. (Ed.), Preprints, Software Business and Engineering Institute, Helsinki University of Technology, Report no.: HUT-SoberIT-B1. Espoo, Finland

Nunes,P. (2015). Conceito de Salário, obtido de http://www.knoow.net/cienceconempr/economia/salario.htm (22 de Janeiro de 2015).

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2008). A design science research methodology for information system research, Journal of management information system.

Weske, M. (2007). Business Process Management: Concepts, Languages. Berlin: Springer.

Modelação de Processos de Negócio: Análise Comparativa de Linguagens

Diogo Silva [1], José Luís Pereira [2]

1) Universidade do Minho, Portugal
a58899@alunos.uminho.pt

2) Universidade do Minho, Portugal
jlmp@dsi.uminho.pt

Resumo

Subjacente a qualquer projeto de *Business Process Management* (BPM) está a necessidade de representar processos de negócio, recorrendo a linguagens adequadas para o efeito. Neste artigo, partindo de uma revisão aprofundada da literatura mais relevante, faz-se uma análise comparativa de cinco das linguagens de modelação de processos de negócio atualmente mais utilizadas no âmbito dos projetos BPM. O objetivo principal é conhecer os pontos fortes e as maiores limitações de cada uma, de modo a poder traçar uma perspetiva comparativa entre elas. Para esse efeito, descrevem-se sucintamente as respetivas notações fazendo-se uma breve análise aos elementos utilizados por cada técnica de modelação. Depois de devidamente analisadas estabelece-se um *framework* comparativo de cada uma das notações relativamente a um conjunto de critérios relevantes para a sua comparação, assinalando os trabalhos mais relevantes.

Palavras-chave: Modelação de processos de negócio, linguagens de modelação de processos.

1. Introdução

Com este trabalho pretende-se analisar e comparar algumas das principais linguagens para modelação de processos de negócio utilizadas na implementação de projetos BPM (*Business Process Management*) nas organizações. A oferta ao nível destas notações é já bastante extensa dificultando aos modeladores dos projetos uma escolha sustentada. Estes projetos têm sido cada vez mais valorizados pelas organizações, na perspetiva destas melhorarem os seus processos de negócio, garantindo assim a operacionalização das suas estratégias de negócio, um maior alinhamento entre essas estratégias e os seus SI/TI e, de uma forma geral, melhorando as suas capacidades de gestão do seu negócio [van der Aalst 2013].

Entre as várias linguagens de modelação de processos de negócio, destacam-se cinco casos como os mais influentes nos dias de hoje: a linguagem BPMN *(Business Process Model and Notation*), atualmente na sua versão 2.0, é a mais utilizada e mundialmente considerada como um *standard* na prática da atividade; a EPC (*Event-driven Process Chain*), desenvolvida no âmbito da conhecida ferramenta ARIS, com o propósito de modelar processos de negócio que, rapidamente

e de uma forma simples, fossem percebidos e executados por qualquer indivíduo numa organização; a UML (*Unified Modelling Language*), criada pela OMG (*Object Management Group*), inicialmente com o principal objetivo de dar suporte ao desenvolvimento de *software* mas mais tarde com um âmbito de utilização mais generalizado; a IDEF (*Integration DEFinition*) conhecida também por ser uma família de notações com propósitos muito diversos, inicialmente desenvolvida para um programa específico, levado a cabo pela Força Aérea dos Estados Unidos; finalmente a linguagem RAD (*Role Activity Diagram*), com características muito particulares, cujas principais preocupações de modelação estão associadas aos intervenientes de um processo de negócio, e às suas interações no decorrer do mesmo.

De entre a panóplia existente de linguagens para modelação de processos de negócio interessa clarificar os aspetos em que estas se distinguem, de modo a permitir uma escolha mais informada por parte dos agentes que as utilizam nas organizações, quer os analistas/modeladores que as usam para documentar e definir os processos, quer pelos utilizadores que terão de validar os modelos construídos pelos primeiros.

2. A Gestão de Processos de Negócio (Business Process Management)

É possível definir a gestão de processos de negócio como um conjunto de competências e métodos ou, simplesmente, como uma abordagem que através de uma simbiose da gestão com as TI procura identificar e definir os processos de negócio mais adequados a uma determinada organização, assim como a sua otimização.

Esta é uma área em franca expansão, já que as organizações procuram melhorar o seu desempenho através de uma gestão mais eficaz e eficiente dos seus processos de negócio. Espera-se que com os processos organizacionais bem arrumados, fiquem perfeitamente identificadas as atividades que devem, ou não, ser realizadas e assim garantir uma rentabilização de recursos, tempo e custos [ABPMP 2013].

Devillers [2011] defende que a gestão de processos de negócio tem um ciclo de vida suportada em 5 fases (ver Figura 1). A fase de **concepção**, onde se identificam os processos já presentes na organização (*as is*) e se projetam os futuros (*to be*). Trata-se de uma fase em que são levantadas, identificadas e registadas as tarefas, as responsabilidades, os recursos e os resultados inerentes a cada processo. É uma fase que se centra no desempenho do processo [Pereira 2011]. A segunda fase é a fase de **modelação**, que procura transformar as informações obtidas na fase anterior em modelos de processo de negócio, geralmente através de ferramentas tecnológicas e com base na utilização de uma linguagem de modelação [Devillers 2011]. Aparece em destacado na figura, dado ser a área em que este trabalho se enquadra.

A próxima fase é a de **execução**, onde são implementadas as soluções informáticas de forma a suportar a automatização dos processos de negócio. Ou seja, tenta-se garantir uma integração e sincronização entre as várias aplicações utilizadas pela organização que até aqui tinham funcionado individualmente. Com o desenvolvimento das ferramentas tecnológicas de suporte ao BPM, é possível aos sistemas computacionais interpretarem e executarem os próprios processos de negócio reais, baseadas na sua anterior modelação. Executam-se os processos provenientes das fases anteriores [Pereira 2011].

A quarta fase é a fase de **controlo**, onde será medido o desempenho dos processos. Esta análise tem como base a tecnologia implementada anteriormente. Dependendo da qualidade e características das soluções desenvolvidas o nível de análise pode ser global, mas também individual a cada processo e até a cada atividade. Estas análises podem depois ser trabalhadas e visualizadas por todos os interessados dentro da organização, permitindo aos gestores tomarem as decisões mais adequadas ao nível organizacional [Devillers 2011].

Por fim a última fase destacada por *Devillers* [2011] é a fase de **otimização**, nesta fase os processos de negócio devem ser otimizados tendo por base os resultados e as análises realizadas na fase de controlo. Estas otimizações podem ir desde a identificação e modelação de novos processos como a remodelação de processos que não estejam a ser tão eficazes como o esperado. O resultado desta fase pode até ser mesmo o reinício de uma nova fase de conceção, completando assim um ciclo de vida, num ciclo que se espera ser de melhoria continua.

Neste estudo, como foi referido atrás, será focada a fase dois descrita por *Devillers* [2011]. Ou seja, incidirá especificamente sobre a **modelação de processos de negócio.**

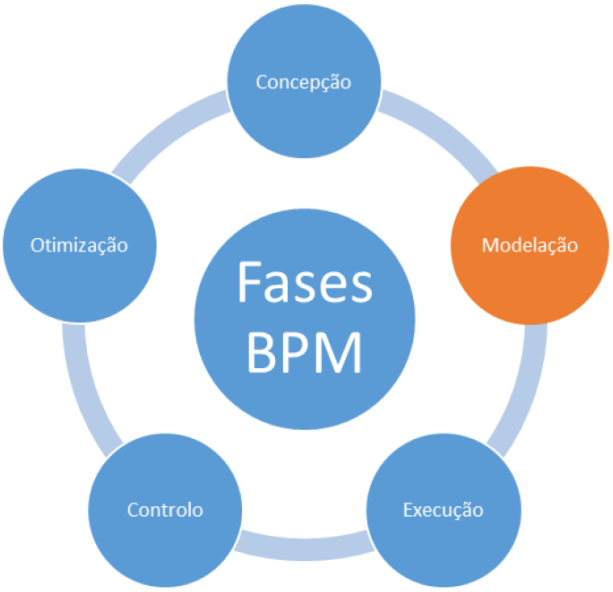


Figura 1 - Fase BPM (Adaptado de [Devillers 2011])

Convém ainda reafirmar que a área BPM tem sido nos últimos anos alvo de grande interesse, em virtude da sua capacidade e potencial para, comprovadamente, garantir melhorias na produtividade das organizações e diminuição de custos, melhorando assim significativamente a sua sustentabilidade [van der Aalst 2013].

3. Modelação de Processos de Negócio

De acordo com Chinosi & Trombetta [2012], o termo *modelação* associado a *processos de negócio* surge pela primeira vez em 1960, num estudo denominado "*Business Process Modeling Improves Administrative Control*". Daí para a frente o termo foi alvo de muito estudo e desenvolvimento (sobretudo a partir de 1990, quando se criou nas organizações uma grande consciência dos conceitos de processo e fluxo de trabalho), tendo surgido desde então inúmeras linguagens para a modelação de processos.

Linda Padilla [2014] descreve a modelação de processos de negócio como uma atividade que representa, numa dada notação gráfica, como um negócio se deve comportar, ou seja, como os processos vão cumprir as suas tarefas e atingir os seus objetivos.

Esta autora é da opinião que a modelação de processos de negócio é das principais etapas em qualquer projeto BPM, sendo a principal responsável pelo sucesso dos mesmos. Para *Padilla* [2014], se nesta fase o trabalho não for bem executado irão ocorrer erros nas fases posteriores, sendo que alguns deles podem comprometer o projeto global. Por isso aconselha-se a escolha e utilização de ferramentas de modelação e notações simples e intuitivas, de modo a que todos os interessados (analistas e utilizadores) compreendam facilmente o comportamento dos processos de negócio.

Outra definição de modelação de processos de negócio é dada por *Pourshahid* [2009], que diz que esta tarefa tem como principal foco documentar e visualizar graficamente processos de negócio, sendo assim uma tarefa estrutural que dá suporte aos interessados na análise dos processos e na procura constante das suas otimizações. Na modelação destes processos devem estar representados aspetos como os objetivos e resultados de cada processo mas também da organização, para além de respostas às 5 famosas questões (*O quê? Onde? Quando? Quem? Porquê?*), pois através delas conseguimos ter a definição completa dos processos e saber quais os seus requisitos e objetivos na organização e para a organização.

A modelação de processos baseia-se na utilização de técnicas de representação designadas por notações ou linguagens de modelação de processos. Dada a importância desta atividade, foram propostas múltiplas notações ou linguagens, embora muitas tenham caído em desuso com o passar do tempo. Cada linguagem de modelação de processos tem associadas a si diferentes perspetivas e características. Nas próximas secções serão apresentadas as linguagens consideradas mais

relevantes nos dias de hoje. Em termos de exemplificação, até para facilitar posteriores comparações entre linguagens, recorre-se sempre ao mesmo exemplo de processo de negócio.

3.1 Business Process Model and Notation

A *Business Process Model and Notation* (BPMN) é, atualmente, a mais representativa linguagem de modelação de processos de negócio. Foi desenvolvida com esse mesmo objetivo e com a preocupação principal de ser amplamente compreendida por todos os utilizadores e analistas (tanto de negócio como de TI), desde os profissionais de modelação de processos, implementadores até aos gestores organizacionais, independentemente dos seus níveis de especialização, para que todos realizem as suas tarefas de desenvolvimento, implementação e monitorização de uma forma simplificada.

Esta notação foi publicada pela primeira vez em 2004 pela *Business Process Management Initiative* (BPMI) que a descreveu como uma notação gráfica capaz de representar processos de negócio, e que tinha de certa forma sido inspirada pelos diagramas de atividade da *Unified Modeling Language* (*UML*), apresentada mais à frente.

De 2004 em diante foram inúmeras as demonstrações de interesse e adoções desta nova notação por parte das organizações, facto que não passou despercebido ao *Object Management Group* (OMG) - organização responsável pela aprovação de padrões para sistemas de informação, que não tardou a adotar e a manter a notação como um dos seus padrões [Chinosi & Trombetta 2012]. A partir de 2011 a notação BPMN passa, de facto, a ser considerada um *standard* entre as várias linguagens de modelação de processos.

Pode-se afirmar que esta notação também é muito baseada nos moldes principais de fluxogramas, mas apresenta uma maior variedade de representações, distanciando-se assim destes pelo seu maior poder de expressividade [Mili et al. 2010].

Por seu lado, a ABPMP [2013] reforça a ideia que esta notação é especialmente recomendada quando se pretende realizar a modelação de processos de negócio para a análise e execução por parte de diferentes públicos-alvo e com diferentes especializações. No entanto, para além desta sua mais-valia, também é aconselhada para a simulação de processos de negócio executáveis e para gerar aplicações a executar em *Business Process Management Systems* (BPMS) através de modelos de processo de negócio. Basicamente, é uma notação que é capaz de responder aos vários tipos de modelação de processos (estáticos ou executáveis) [White 2006].

Esta linguagem de modelação tem evoluído ao longo dos anos e apresenta um leque considerável de elementos para representar os diversos detalhes associados aos processos de negócio, que lhe permitem definir relacionamentos como os fluxos de atividades e ordens de precedência ou a

caracterização dos diversos atores intervenientes nos processos através de raias paralelas onde estará definido o que certo ator deve fazer naquele processo.

Algumas das características desta notação mais rapidamente identificáveis são a agregação dos seus componentes simbólicos em grandes grupos. Ao contrário de outras notações é capaz de representar eventos iniciais, intermédios e finais, diferentes fluxos como os de atividades e mensagens, e representação de relacionamentos externos à organização [ABPMP 2013].

A notação BPMN está baseada em quatro grandes agregados de elementos simbólicos para a representação dos processos: os objetos de conexão; os objetos de fluxo; as *swimlanes* ou raias; e os artefactos. Através destas categorias é possível encontrar os elementos representativos de eventos, atividades, papéis, fluxos de trabalho, etc. [Chiarello, Emer & Neto 2014].

Relativamente à primeira grande categoria, os objetos de conexão são responsáveis por garantir três tipos de conexão entre elementos simbólicos: fluxo de sequência tradicional; fluxo de mensagens; e associação de elementos.

Os objetos de fluxo são compostos por elementos capazes de representar todas as ações que estão associadas ao longo de um processo de negócio e aos seus estados comportamentais, ou seja, elementos responsáveis por representar, por exemplo eventos, atividades, decisões (*gateways*).

As *swimlanes* são um agregado de simbologias muito característico desta notação e que na modelação é responsável por agrupar todos os elementos de um *Business Process Modeling Diagram* (BPD), por papéis ou secções de trabalho. Os dois tipos simbólicos aqui presentes são as *lanes* e as *pool's* [Chinosi & Trombetta 2012].

O último conjunto de elementos básicos da BPMN são os artefactos. Este conjunto de símbolos serve para fornecer informações adicionais aos processos que não interferem diretamente com o mesmo, são exemplos as anotações, grupos e *data objects* [Chinosi & Trombetta 2012].

Ao longo das várias versões da notação (sendo a mais recente a BPMN 2.0) o número de elementos simbólicos tem vindo a crescer e a possibilitar um maior número de representações facilitando também a modelação, embora a grande maioria dos utilizadores usem apenas um subconjunto destes com destaque, naturalmente, para os elementos principais da notação, que foram referidos.

Além de um maior número de elementos representativos a BPMN 2.0 veio trazer novas características e mudar algumas propriedades de certos elementos já existentes. A versão 2.0 da notação, comparada à sua antecessora (versão 1.2) possibilita uma maior representação de interações humanas nos processos, refina a composição e correlação dos eventos, formaliza a execução semântica para todos os elementos BPMN, define modelos de conversação e coreografia e resolve ambiguidades e inconsistências da versão anterior [Chinosi & Trombetta 2012].

Na Figura 2, apresenta-se um pequeno processo modelado em BPMN, onde é possível observar alguns dos elementos principais desta linguagem.

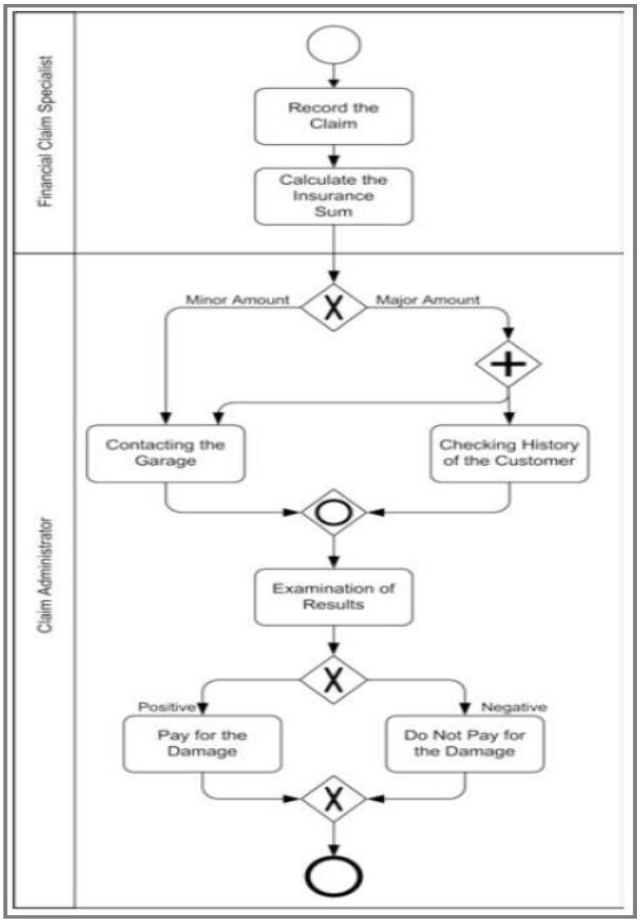


Figura 2 - Processo modelado em BPMN
(Retirado de [Korherr 2008])

3.2 Event-driven Process Chain

A linguagem *Event-driven Process Chain* (EPC) é também uma das linguagens de modelação de processos de negócio mais utilizadas mundialmente. É uma notação que foi desenvolvida em 1992, por investigadores da *Universidade de Saarland* (entre eles, como grande impulsionador, o Professor *August Wilhelm Scheer*) em pareceria com a *SAP*, uma das maiores potencias mundiais na produção de software integrado de gestão [Mili et al. 2010].

Dos anos 90 em diante esta linguagem cresceu e expandiu-se, tornando-se mesmo uma das mais reconhecidas notações de modelação de processos [Devillers 2011]. Trata-se de uma linguagem de modelação que se baseia em conceitos base das *Petri nets* (clássica notação de modelação para sistemas distribuídos) [Korherr 2008]. Tal como a maioria das notações para a modelação de processos, a EPC também se assemelha na sua estrutura a fluxogramas para a representação de

dependências lógicas e temporais entre atividades na construção dos processos de negócio [Keletso, Chioasca & Zaho 2014].

O principal foco da notação EPC é permitir aos seus utilizadores uma representação gráfica dos processos organizacionais, de uma forma intuitiva e que seja de rápida e simples compreensão tanto para analistas de processos como para o pessoal do negócio [Van Wel 2013].

Para além disso, a EPC é a principal linguagem para a representação de processos de negócio da metodologia ARIS (*Architecture of Integrated Information Systems*), que agrega os recursos referentes ao negócio (sistemas, dados, etc.) e organiza-os de forma a garantir o desenvolvimento de sequências de atividades/tarefas que produzem valor [Davis & Brabander 2007]

Esta notação é constituída por um leque de elementos básicos e complexos. O primeiro conjunto é constituído por funções, eventos, conectores lógicos e fluxos de controlo, enquanto o conjunto de elementos complexos é representado por unidades organizacionais/papéis, objetos informacionais e objetos de entrega.

Os eventos são mecanismos responsáveis por realizarem o início de cada processo, definirem o estado do processo ou terminar o mesmo, representando assim pré-condições e pós-condições de uma função. Nunca poderão ocorrer situações com dois eventos sucessivos e cada evento no máximo só terá uma conexão de entrada e outra de saída.

As funções são elementos que têm como objetivo representar as atividades ou tarefas presentes no processo de negócio. Geralmente estas atividades são executadas por pessoas ou por sistemas computacionais. Uma função deve ser ativada por um evento antecessor e pode dar origem a um ou mais eventos sucessores. Tal como os eventos, não poderão ocorrer duas funções sucessivas nem podem ter mais que uma conexão de entrada nem mais que uma de saída.

Os conectores lógicos servem essencialmente para representar as tomadas de decisão ao longo do processo, sendo que existem três tipos de conectores lógicos: o OU (representado por V) o E (representado por Λ) e o OU EXCLUSIVO (representado por XOR) [Thomas 2004]. Já os fluxos de controlo são utilizados para definir os fluxos sequenciais de trabalho o dito *workflow*. As unidades organizacionais definem quem é responsável pela realização e desempenho de determinada atividade ou tarefa.

Por sua vez, os objetos informacionais, são utilizados para apoiar as funções através da entrada de informações adicionais que servem de base para o desempenho das mesmas ou, por outro lado, podem corresponder a um conjunto de informações produzidas aquando da execução de determinada função a si associada.

Por fim os objetos de entrega, representam os resultados (serviço ou produtos) das funções ou o que elas precisam para serem executadas [Korherr 2008].

De seguida, na Figura 3, apresenta-se o mesmo exemplo anterior de um processo de negócio, agora modelado com a linguagem EPC.

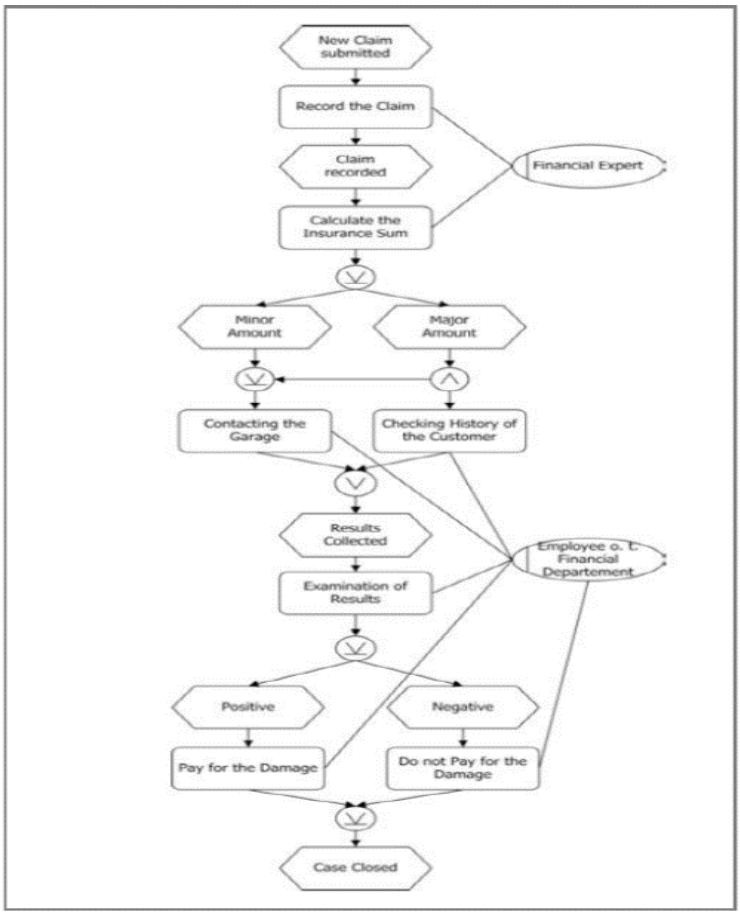


Figura 3 - Processo modelado em EPC
(Retirado de [Korherr 2008])

3.3 Unified Modeling Language: Activity Diagrams

A *Unified Modeling Language* (UML) é uma notação de utilização geral nos dias de hoje, desenvolvida originalmente pela OMG (a primeira versão em 1995) com a intenção de ser utilizada no âmbito da Engenharia de Software. É uma linguagem de modelação padrão da OMG (desde 1997) e em 2000 foi aceite como um *standard* pela *International Standards Organization*, como um *standard*.

Esta técnica como foi referido anteriormente, foi desenvolvida com o propósito de ser uma notação gráfica para a análise, especificação, construção e documentação de suporte ao desenvolvimento de software. O principal desafio aquando do desenvolvimento desta notação era o de criar uma linguagem padronizada e com uma semântica única para ser capaz de representar sistemas de diferentes níveis de complexidade [Geambaşu 2012].

A versão atual é a UML 2.4.1 lançada em 2011. As atualizações UML 2 melhoraram significativamente as anteriores UML 1, de tal forma que fez a notação distanciar-se um pouco do seu propósito inicial de suporte ao desenvolvimento de software para ir de encontro a uma visão mais global dos sistemas "gerais", diversificando assim os seus domínios de aplicação. Esta linguagem é composta por uma panóplia de diferentes diagramas, alguns dos quais podem ser utilizados para a modelação de processos de negócio. Essa estrutura é composta segundo *Szilagyi* [2010] por:

- Diagrama de *Use-Case*, onde são representados os processos organizacionais de alto nível e a interação destes com as entidades externas;
- Diagrama de Classes, que permitem a identificação das entidades presentes no negócio e como elas se relacionam entre si;
- Diagramas de Sequência, que demonstram as várias interações entre os intervenientes num processo e as tarefas que estão associadas aos mesmos;
- **Diagramas de Atividades**, estes são os diagramas mais importantes na modelação de processos de negócio pois são eles os responsáveis por demonstrar graficamente o fluxo de trabalho de cada processo de negócio.

Sendo o diagrama de atividades o mais frequentemente utilizado para a modelação de processos de negócio, é importante sublinhar que estes foram significativamente redesenhados a partir da versão 2.0 da UML, tanto em termos de sintaxe como também de semântica. Alterações essas que melhoraram significativamente a capacidade destes diagramas para representar processos de negócio [Geambaşu 2012].

Geralmente os diagramas de atividade são constituídos por cinco conjuntos de elementos, são eles as ações, as sub-atividades, os objetos de dados, os controlos e as partições. As ações são os elementos responsáveis por definir o comportamento do processo de negócio, geralmente recebem conectores de entradas e saídas e são os principais responsáveis pelas alterações no estado do sistema. As sub-atividades têm apenas como objetivo esconder a complexidade do modelo de processo. Os objetos de dados são utilizados para indicar uma instância de um estado particular associado a um determinado ponto da atividade. Os controlos servem para coordenar todos os fluxos entre os elementos. Por fim as partições têm como função dividir partes do fluxo de trabalho do processo geralmente por unidades organizacionais (ou seja, o que as *swimlanes* fazem no BPMN) [Eloranta, Kallio & Terho 2006].

Do mesmo modo que nas secções anteriores na próxima figura (Figura 4) apresenta-se como exemplo o processo anterior modelado com um Diagrama de Atividades UML.

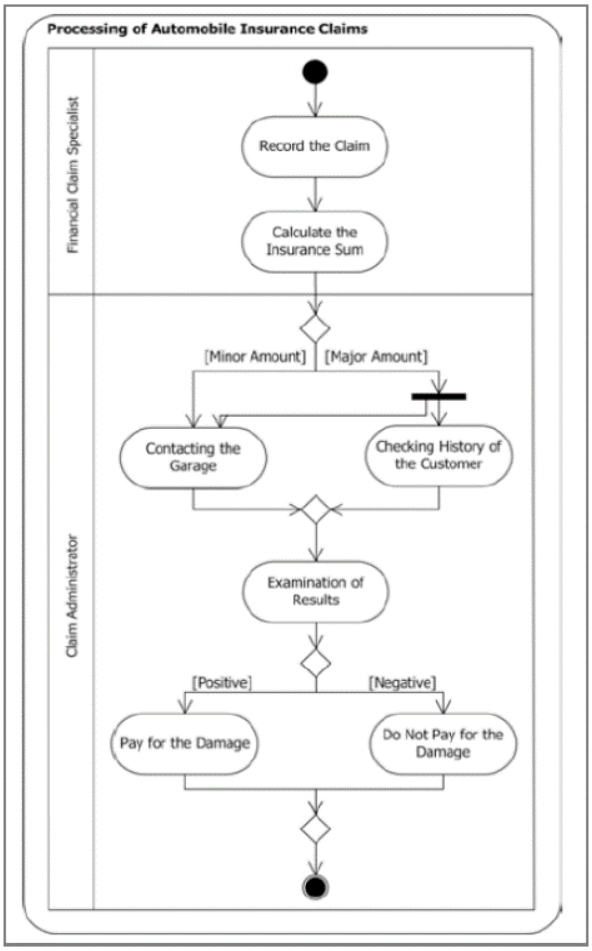


Figura 4 - Processo modelado em UML-AD
(Retirado de [Korherr 2008])

3.4 Role Activity Diagrams

A linguagem *Role Activity Diagram* (RAD) foi desenvolvida inicialmente para a modelação da coordenação de trabalho em ambientes de programação, mas nos dias de hoje é de maior utilização na modelação de processos de negócio, tanto os existentes (*as is*) como os que se pretendem para o futuro (*to be*) [Van Wel 2013].

É uma linguagem de modelação criada em 1986 por *Ould* e *Roberts*, que desde então integra a STRIM (*Systematic Technique for Role and Interaction Modeling*), metodologia desenvolvida pela Praxis (uma *software house* a que *Ould* se juntou e foi diretor técnico de qualidade antes desta ser adquirida pela *Deloitte Consulting* para se tornar na sua secção de engenharia de software), que descreve a RAD como "Técnica de levantamento, modelação e análise de processos de negócio". Já a metodologia STRIM foi desenvolvida através do trabalho de pesquisa de *Ould* e *Roberts* sobre a modelação de processos de *software*.

Ould no desenvolvimento desta metodologia STRIM, identificou cinco elementos necessários para a modelação de processos: os papéis; os atores; as interações; as atividades; e as entidades.

Os papéis são responsabilidades atribuídas a um determinado indivíduo representado no processo. Atores representam no processo os indivíduos ou sistemas que desempenham funções específicas (papel) num determinado momento no tempo. Interações são os elementos responsáveis pela sincronização, comunicação e partilha de dados entre os atores do processo. Atividades são os elementos que representam o que um determinado ator faz na execução do seu papel, o que ele deve realizar e quando deve realizar, sendo que na realização destas atividades os atores podem precisar de interagir entre si. Por fim as entidades representam os objetos que os autores trocam entre si ao longo das suas interações.

De certa forma os RAD são semelhantes aos Diagramas de Atividade da UML, só que enquanto na RAD foca-se mais e dá-se enfase à atribuição de responsabilidades ao longo de todo o processo, os Diagramas de Atividade da UML centram-se mais no alinhamento orquestrado das tarefas do processo.

Para além disso a RAD é uma notação mais centrada na modelação estática dos processos não permitindo a sua direta automatização e execução. Outra característica da RAD é que esta notação não possui uma formalização da sua semântica, devido ao facto de, ao contrário das linguagens de modelação de processos referidas anteriormente, a RAD apenas ter como objetivo exclusivo, permitir a compreensão de uma forma simples, por parte de todos interessados, do funcionamento dos processos de negócio e não de fazer dos seus modelos uma base para a execução ou simulação dos processos. No entanto existem vários estudos com o objetivo de formalizar a RAD para depois a traduzir noutras notações de forma a realizar a execução dos processos [Devillers 2011].

Na Figura 5 apresenta-se o processo anterior, agora modelado em RAD.

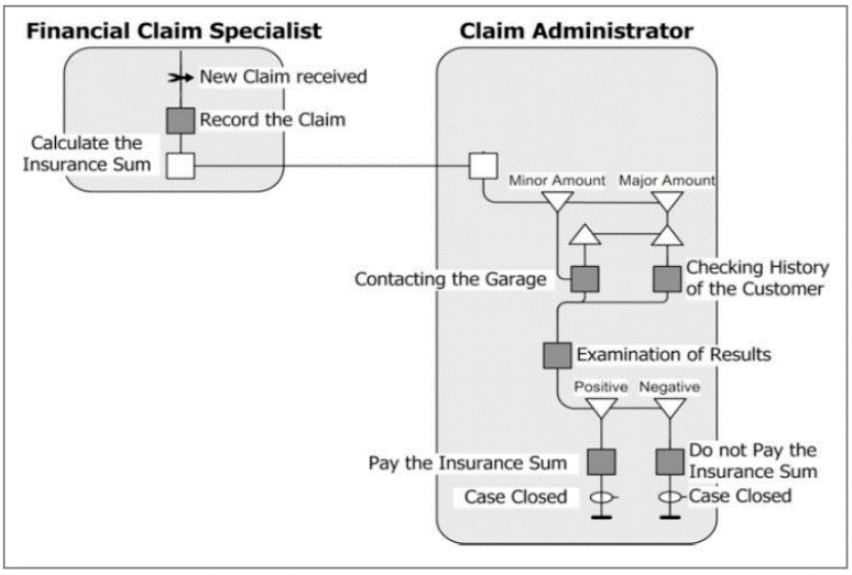


Figura 5 - Processo modelado em RAD
(Retirado de [Korherr 2008])

3.5 Integration DEFinition

A *Integration DEFinition* (*IDEF*) é uma "família" de 16 técnicas para a modelação e análise de sistemas, inicialmente projetada para ser utilizada no âmbito da Engenharia de Software. Cada uma destas técnicas é utilizada em diferentes âmbitos de aplicação, no entanto a IDEF0, mais direcionada para modelar funções de negócio, e IDEF3, para modelar concretamente os processos de negócio, são de facto as duas notações mais úteis e que se podem complementar na modelação de processos, embora cada uma destas também o possam realizar individualmente, em particular a IDEF3 [Costin & Fox 2004].

3.5.1. IDEF0

A técnica IDEF0 é a primeira da "família" IDEF e é capaz de representar um conjunto de ações suportando-se nos ICOMs (*Input Control Output Mechanism*). O ICOM é uma representação gráfica de uma ou mais atividades, que para além de dados e informações é capaz de descrever os vários elementos que estão associados a um processo.

Cada ICOM é composto por conectores ou "terminais" de entrada e saída, através dos quais se gera o fluxo de trabalho do processo. Ao nível de conectores de entrada cada ICOM recebe o elemento central do processo a ser tratado (por exemplo um produto), recebe também os controlos que representam as responsabilidades associadas à execução da tarefa, recebe ainda mecanismos que definem quem é o responsável pela execução da atividade (pessoa ou máquina) e deve ter como elemento de saída o resultado daquela atividade que foi realizada sobre o elemento de entrada.

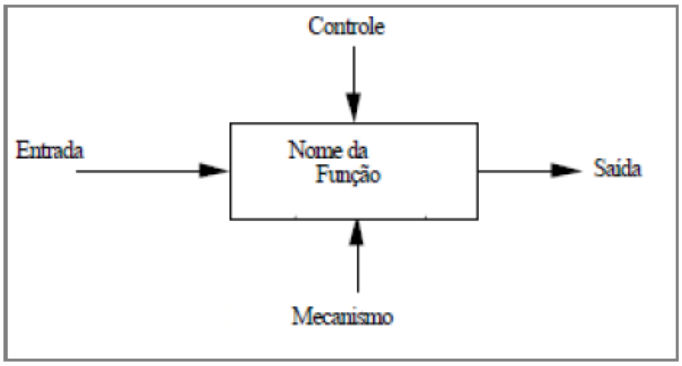


Figura 6 - Exemplo de ICOM (Retirado de [Mykolayczky & Júnior 2001])

Posto isto pode-se afirmar que o modelo funcional IDEF0 é representado por um agregado de ICOMs, em que cada um deles representa uma determinada atividade pertencente ao processo. Cada uma das atividades pode depois ser detalhada, ou seja, baixa-se um nível de detalhe e cria-se um sub-processo. Este processo pode ser repetido, criando-se assim uma hierarquia de processos no modelo IDEF0 [Mykolayczky & Júnior 2001].

Ao nível da semântica esta técnica, para alguns autores, é considerada limitada, pois tudo se baseia em atividades que são representadas por "caixas", linhas e setas [Devillers 2011].

Para além de elementos gráficos responsáveis pela representação gráfica das atividades um diagrama IDEF0 pode ainda conter elementos como textos e glossários. Os textos são muitas vezes usados para resumir um determinado diagrama e geralmente destaca algumas características do mesmo com o objetivo de clarificar o intuito e funcionalidade do diagrama em questão. Já os glossários servem como o próprio nome indica para definir significados de expressões utilizadas nos diagramas e têm também por objetivo dar apoio à interpretação do diagrama.

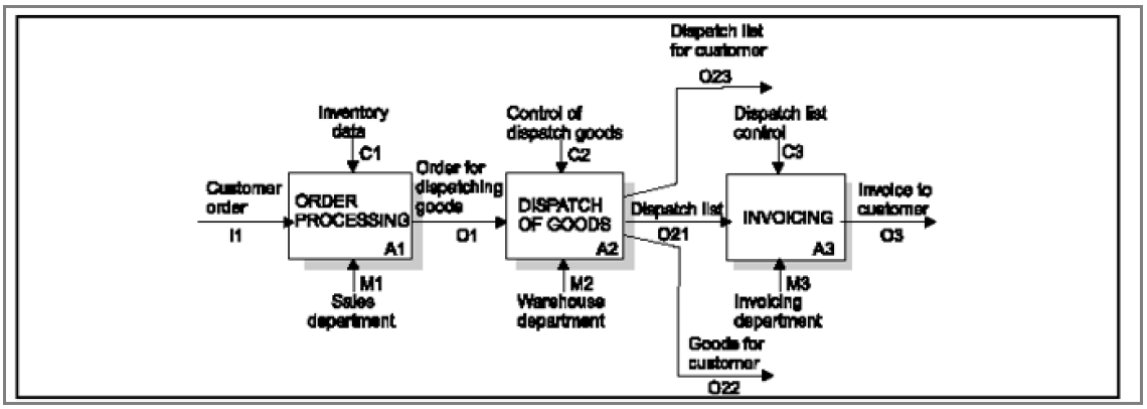


Figura 7 - Processo modelado em IDEF0 (Retirado de [Demirag, Johnson, Nazzal, & Wan 2000])

3.5.2. IDEF3

A técnica IDEF3 foi desenvolvida, especificamente, para descrever o aspeto dinâmico dos processos de negócio e com o objetivo de facilitar o levantamento e descrição de sistemas de informação [Mili et al. 2010].

É uma técnica que se centra no aspeto temporal dos processos e que veio então responder à necessidade anteriormente identificada na IDEF0. Esta técnica descreve dois tipos de linguagens de modelação, uma com o objetivo de descrever os fluxos de trabalho dos processos de negócio e outra para definir as transições de estado dos objetos. O nosso foco de pesquisa centra-se mais no primeiro tipo e relativamente a isso a IDEF3 descreve um processo como uma sequência de atividades. As atividades em IDEF3 são reconhecidas como unidades de comportamento, e as relações entre atividades são denominadas de ligação de precedência. Além destes elementos também existem os fluxos de controlo que servem para realizar divisões ou junções ao longo do fluxo de trabalho. Esta é uma técnica muito semelhante na sua composição de elementos aos Diagramas de Atividade da UML ou aos EPC, com a exclusão dos eventos que neste caso não

têm uma representação própria, ou seja, não existem estruturas capazes de representarem o início e fim do processo explicitamente.

Tanto esta técnica como a IDEF0 têm características únicas relativamente às outras linguagens de modelação, pois têm uma sintaxe bastante elaborada, sendo assim técnicas muito formais comparando com algumas das outras aqui apresentadas. Além disso estas notações destacam a importância da decomposição funcional dos processos e incentivam à sua utilização. O que nem sempre é uma boa prática pois se o processo for longo e complexo ele tem de ser dividido em bastantes diagramas o que por vezes dificulta muito a compreensão e impossibilita ter uma visão global do processo (pois embora exista um diagrama de raiz ele não tem grande profundidade para dar uma noção muito estruturada de todo o processo). Por tudo isto muitas vezes os diagramas IDEFs são acompanhados de um diagrama em árvore que descreve as relações entre os vários diagramas [Devillers 2011].

A figura seguinte (Figura 8) apresenta o processo que temos vindo a utilizar como exemplo agora modelado na notação IDEF3.

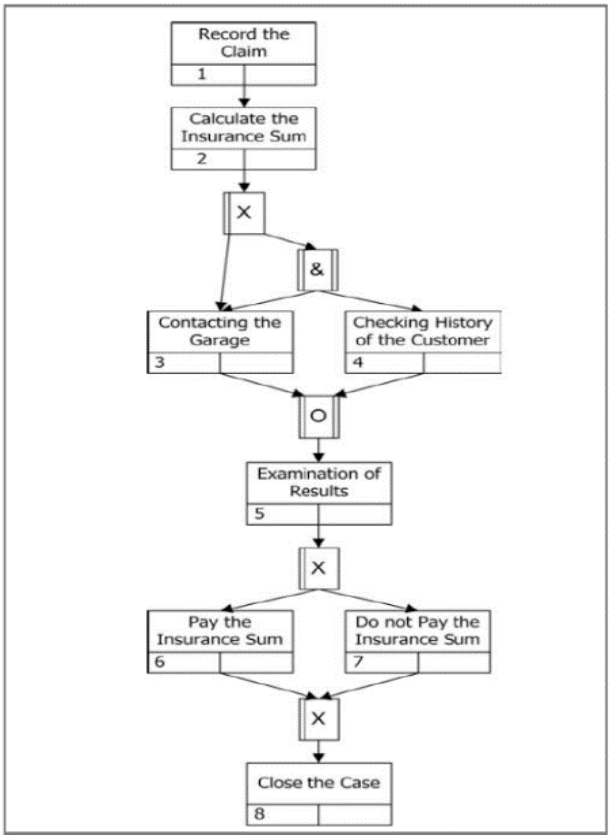


Figura 8 - Processo modelado em IDEF3 (Retirado de [Korherr 2008])

4. Análise Comparativa das linguagens

Como foi referido anteriormente, a modelação de processos de negócio tem grande impacto no sucesso de qualquer projeto BPM. Por essa razão a escolha da linguagem de modelação a utilizar não é indiferente, devendo depender dos objetivos específicos do projeto em causa (simples documentação de processos de negócio com vista à sua comunicação e divulgação entre os intervenientes? Modelação dos processos para posterior otimização do seu funcionamento? Levantamento dos processos com vista ao seu suporte usando TI? Etc.) e das características do próprio processo de negócio (Envolvendo sobretudo recursos humanos? Ou integrando principalmente sistemas de processamento automatizados? Constituídos por fluxos de trabalho (workflows) bem estruturados? Ou envolvendo predominantemente interações entre pessoas? Etc.). Com o estudo realizado até ao momento foi possível verificar que a oferta em termos de linguagens/notações para modelação de processos de negócio já é de certa forma alargada e que embora todas partilhem o mesmo objetivo, cada uma tem as suas características próprias.

As cinco linguagens para modelação de processos de negócio sucintamente apresentadas nas secções anteriores, sendo bem representativas do que é possível encontrar neste domínio, possuem diferentes características, pontos fortes e limitações. Interessa agora encontrar um meio de as comparar de modo a sistematizar as suas diferenças e semelhanças. Para esse efeito foi já realizada uma revisão alargada da literatura relevante que se debruça sobre a caracterização das linguagens de modelação de processos.

Diferentes autores recorrem a diferentes critérios para avaliar as linguagens de modelação de processos, embora alguns critérios tendam a ser mais ou menos universais:

- **Expressividade** – a capacidade da linguagem permitir representar as mais diversas situações organizacionais, quer em termos comportamentais, funcionais, estruturais ou informacionais;
- **Legibilidade** – a facilidade com que os modelos de processos definidos na linguagem são compreendidos e interpretados pelas pessoas;
- **Usabilidade** – embora relacionado com a legibilidade, este critério foca-se sobretudo na complexidade da utilização da linguagem;
- **Amigabilidade** – também um tanto relacionado com a legibilidade, este critério refere-se à maior ou menor atratividade da linguagem, no sentido em que pode ser de utilização mais ou menos agradável ou intuitiva;
- **Formalidade** – o rigor com que a semântica da linguagem está definida, reduzindo/removendo as ambiguidades na interpretação dos modelos;
- **Universalidade** – a maior ou menor popularidade da linguagem em termos de divulgação no mercado, com implicações ao nível do suporte à sua utilização.

Na tabela seguinte (Tabela 1) sistematizam-se os critérios comparativos referidos assim como os autores que, na literatura da especialidade, os valorizam para caracterizar as linguagens de modelação de processos. Como se pode verificar, a expressividade das linguagens é, de longe, o critério que mais atenção tem recebido.

Critérios/Notações	BPMN	EPC	UML:AD	RAD	IDEF0/IDEF3
Expressividade	[Heidari, Loucopoulos, Brazier & Barjis 2012] [Korherr 2008] [Mohammadi & Mukhtar 2012] [Mili et al. 2010] [Kelemen, Kusters, Trienekens & Balla 2013] [Aldin & Cesare 2009]	[Heidari et al. 2012] [Korherr 2008] [Mohammadi & Mukhtar 2012] [Mili et al. 2010] [Kelemen et al. 2013]	[Heidari et al. 2012] [Korherr 2008] [Mohammadi & Mukhtar 2012] [Mili et al. 2010] [Kelemen et al. 2013]	[Heidari et al. 2012] [Korherr 2008] [Mili et al. 2010] [Aldin & Cesare 2009]	[Heidari et al. 2012] [Mohammadi & Mukhtar 2012] [Mili et al. 2010] [Korherr 2008]
Legibilidade	[Aldin & Cesare 2009] [Van Wel 2013] [Kelemen et al. 2013]	[Van Wel 2013] [Kelemen et al. 2013]	[Van Wel 2013] [Kelemen et al. 2013]	[Aldin & Cesare 2009] [Van Wel 2013]	[Van Wel 2013]
Usabilidade	[Aldin & Cesare 2009] [Van Wel 2013]	[Van Wel 2013]	[Van Wel 2013]	[Aldin & Cesare 2009] [Van Wel 2013]	[Van Wel 2013]
Formalidade	[Mili et al. 2010]	[Mili et al. 2010]	[Mili et al. 2010]	[Mili et al. 2010]	[Mili et al. 2010]
Amigabilidade	[Aldin & Cesare 2009]			[Aldin & Cesare 2009]	
Universalidade	[Kelemen et al. 2013]	[Kelemen et al. 2013]	[Kelemen et al. 2013]		

Tabela 1 - Critérios de Avaliação vs Linguagens - Cobertura na literatura existente

Esta tabela será ainda completada com outros critérios, por nós julgados relevantes para fazer uma comparação mais completa das linguagens de modelação de processos. Por exemplo, critérios como a *concisão* (isto é, a maior ou menor capacidade de representar as várias facetas de um processo recorrendo a um menor conjunto de elementos), a *simplicidade / curva de aprendizagem* (isto é, o maior ou menor esforço necessário para dominar e ser produtivo na utilização da linguagem), o *suporte à inovação* (isto é, a maior ou menor facilidade com que induz os modeladores a descobrir novas soluções), a *amplitude de utilização* (isto é, se a linguagem se adequa apenas à documentação e análise dos processos, ou permite a execução e/ou simulação dos mesmos), até outros critérios mais pragmáticos como a existência de ferramentas de suporte à linguagem, a portabilidade dos modelos desenvolvidos, ou até as garantias de evolução futura da linguagem, poderão ser também critérios a ter em consideração na escolha de uma linguagem de modelação de processos.

Depois de encontrados os critérios necessários à análise comparativa entre as linguagens de modelação de processos pretende-se valorizar, critério a critério, qual ou quais as notações que melhor os verificam, para de uma forma estruturada apresentar as soluções mais adequadas. Naturalmente, distintas organizações / projetos BPM poderão valorizar de modo diferente cada um dos critérios comparativos. O que significa que a seleção será feita atribuindo pesos diferentes a cada um dos critérios, de acordo com os objetivos da organização / projeto BPM . Desse modo, consegue-se fazer a seleção da linguagem mais adequada a cada contexto de utilização, tomando em consideração as suas especificidades.

O grande propósito deste trabalho é construir um artefacto consistente que, perante um determinado contexto organizacional ou projeto BPM, permita a análise comparativa das linguagens de modelação de processos, com vista a selecionar aquela que mais se adequa às suas especificidades. Esta solução permitirá uma seleção mais informada da notação para servir determinados desígnios por parte da equipa de projeto sem a necessidade desta efetuar um grande estudo sobre as diversas linguagens existentes. Bastando apenas indicar, entre os vários critérios identificados, aqueles que para o seu projeto ou contexto organizacional lhe parecem mais importantes, atribuindo-lhes a valorização correspondente.

5. Conclusão

Até ao momento este trabalho permitiu alargar o leque de conhecimentos relativamente à área BPM, especificamente numa fase crítica desta abordagem como é o caso da modelação de processos de negócio. Reconhecidamente, esta fase assume um papel fundamental na melhoria da gestão dos processos de negócio, e consequentemente, na melhoria do desempenho das organizações em resposta às novas exigências da sua envolvente.

Para uma melhor atuação nesta área de modelação de processos ficou claro que é fulcral a utilização de uma notação específica para o efeito. Através de uma inventariação das linguagens de modelação mais importantes nesta área foi possível verificar que a oferta é alargada e que, embora partilhem o mesmo objetivo, cada uma tem as suas características específicas.

Daqui em diante pretende-se através da continuação deste estudo encontrar uma solução viável para um problema/necessidade encontrado, isto é, a realização de uma base de diferenciação de notações de modelação de processos. A resolução deste problema passa pelo desenvolvimento de uma ferramenta ou artefacto de suporte à comparação de várias notações de modelação de processos de negócio, que facilitará a seleção da linguagem mais adequada aos propósitos e características de cada projeto BPM particular. Neste momento, a ferramenta encontra-se em fase de protótipo.

6. Referências

ABPMP. (2013). *Association Business Process Management Professionals - Common Body of Knowledge 3.0*.

Aldin, L., & Cesare, S. De. (2009). A comparative analysis of business process modelling techniques. *UKAIS 2009l, Oxford, UK*, (Ukais), 1–17. Retrieved from http://bura.brunel.ac.uk/handle/2438/4078

Chiarello, M. a., Emer, M. C. F. P., & Neto, A. G. S. S. (2014). An Approach of Software Requirements Elicitation Based on the Model and Notation Business Process (BPMN). *Lecture Notes on Software Engineering*, *2*(1), 65–70. doi:10.7763/LNSE.2014.V2.96

Chinosi, M., & Trombetta, A. (2012). BPMN: An introduction to the standard. *Computer Standards & Interfaces*, *34*(1), 124–134. doi:10.1016/j.csi.2011.06.002

Costin, B., & Fox, C. (2004). Hybrid IDEF0 / IDEF3 Modelling of Business Processes : Syntax , Semantics and Expressiveness. *Concurrent Engineering*, 3–5.

Davis, R., & Brabander, E. (2007). *Platform, Aris Design : Getting Started with BPM*. (Springer, Ed.).

Demirag, O., Johnson, A., Nazzal, D., & Wan, Y.-T. (2000). Integrated Definition (IDEF) Modeling Techniques.

Devillers, M. (2011). *Business Process Modeling The Business-IT Divide*. Radboud University Nijmegen, Master's Thesis.

Eloranta, L., Kallio, E., & Terho, I. (2006). A Notation Evaluation of BPMN and UML Activity Diagrams Lauri Eloranta Eero Kallio Ilkka Terho. *Information Systems Journal*. Retrieved from http://www.soberit.hut.fi/T-86/T-86.5161/2006/UML_PBMN_Final_Presentation.pdf

Geambaşu, C. V. (2012). BPMN vs UML Activity Diagram for Business Process Modeling. *Accounting and Management Information Systems*, *11*(4), 637–651.

Heidari, F., Loucopoulos, P., Brazier, F., & Barjis, J. (2012). A Unified View of Business Process Modelling Languages 1, *004*.

Keletso, L., Chioasca, E., & Zaho, L. (2014). An Integrative Approach To Support Multi-Perspective Business Process Modeling.

Korherr, B. (2008). *Business Process Modelling - Languages, Goals and Variabilities*. Vienna University of Technology, PhD Thesis.

Mili, H., Tremblay, G., Jaoude, G. B., Lefebvre, É., Elabed, L., & Boussaidi, G. El. (2010). Business process modeling languages: Sorting Through the Alphabet Soup. *ACM Computing Surveys*, *43*(1), 1–56. doi:10.1145/1824795.1824799

Mykolayczky, J. L., & Júnior, J. T. (2001). IDEF0-Método de Representação de Processos em Forma de Fluxo. *Núcleo de Pesquisa Em Engenharia Simultânea (NuPES) Centro Federal de Educação Tecnológica Do Paraná (CEFET-PR)*. Retrieved from http://graco.unb.br/alvares/pub/idef0/idef0_cefet.pdf

Padilla, L. (2014). *Transformation of Business Process Models : A Case Study*. Univerity of Porto, Master's Thesis.

Pereira, J. (2011). *Business process management : proposta de framework comparativa das tecnologias XPDL e BPEL*. Universidade do Minho, Tese de Mestrado.

Pourshahid, A., Amyot, D., Peyton, L., Ghanavati, S., Chen, P., Weiss, M., & Forster, A. J. (2009). Business process management with the user requirements notation. *Electronic Commerce Research*, *9*(4), 269–316. doi:10.1007/s10660-009-9039-z

Szilagyi, D. (2010). *Modelagem de Processos de Negócio - um Comparativo entre BPMN e UML. Business*. Universidade Católica São Paulo, Tese de Mestrado.

Thomas, O. (2004). Process Modeling using Event-driven Process Chains, 1–21.

Van der Aalst, W. M. P. (2013). Business Process Management: A Comprehensive Survey. *ISRN Software Engineering*, *2013*, 1–37. doi:10.1155/2013/507984

Van Wel, R. (2013). *Business Best practices in Agile software development*. Leiden University, Master's Thesis.

White, S. A. (2006). Introduction to BPMN. *IBM Software Group*, *Websphere* .

Modelo de Avaliação do Desempenho das Concessões Automóveis

Jorge Pires [1], Bráulio Alturas [2]

1) AXT Capital
jpires.pt@gmail.com

2) Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Portugal
braulio.alturas@iscte.pt

Resumo

O setor automóvel a nível mundial está reduzido a poucos fabricantes que por sua vez agrupam algumas marcas de veículos, apresentando assim evidentes caraterísticas de oligopólio. Por outro lado, o setor sempre se destacou pelos elevados investimentos de marketing, criando fortes imagens de marca e um conjunto de ideias pré-concebidas que dificultam a mudança de marca automóvel no momento em que o consumidor decide trocar de viatura.

Deste modo a discussão passou a ser a avaliação baseada em critérios, centrados nas questões da qualidade de serviços, performance de vendas, satisfação do cliente e outros de índole operacional.

Contudo estes critérios vão variando de fabricante para fabricante, são tratados de forma departamentalizada, nem sempre são consensuais, não têm uma metodologia racional nem global o que conduz à criação de ideias pré-concebidas, que nem sempre são razoáveis.

O presente artigo pretende apresentar um modelo de avaliação do desempenho de gestão das concessões automóveis, que seja uniforme e transversal a todas as marcas automóveis.

Palavras-chave: *concessão automóvel, desempenho, modelo de avaliação, framework.*

1. Introdução

O setor do retalho automóvel apresenta caraterísticas singulares em termos de funcionamento da sua atividade, comparativamente com outros setores de atividade, que assentam no princípio da livre concorrência.

Na europa este setor é regulado pelo *Block Exemption Regulation* [1] (Regulamento de Isenção por Categoria) o que cria algumas barreiras à entrada de novos concorrentes e dá aos fabricantes grande poder sobre a sua rede.

Apesar deste normativo ter sido revisto na última década, beneficiando essencialmente os consumidores, os fabricantes fizeram várias reestruturações e alteraram as políticas comerciais

[1] http://ec.europa.eu/competition/state_aid/legislation/block.html

passando da atribuição de margens fixas para margens variáveis, de acordo com um conjunto de critérios supostamente baseados na boa gestão das concessões automóveis.

Outra das alterações importantes ao Regulamento de Isenção por Categoria inicial foi a seleção dos concessionários pelos fabricantes. Durante anos, os fabricantes nomeavam os seus representantes exclusivos, sendo-lhes atribuída uma área geográfica que exploravam tanto a nível de vendas de automóveis e peças, como de assistência após-venda. Atualmente já não existe a exclusividade territorial nem a exclusividade de todo o negócio (vendas e assistência). A marca define pois um conjunto de requisitos mínimos de investimento, de imagem de marca, de procedimentos de gestão, etc., que ao serem cumpridos tornam a empresa candidata, num novo distribuidor automóvel de viaturas, peças, ou reparador autorizado.

Não obstante estas mudanças, a forma organizativa do setor e o enfoque numa imagem de marca permaneceu igual ao longo dos tempos.

As decisões estratégicas são tomadas no topo e ações vão "deslizando" em sintonia pela organização estendendo-se por todas as regiões do globo de forma organizada, controlada e uniforme.

Geralmente existe uma estrutura organizacional geográfica organizada por continentes, que englobam os seus países e estas as suas regiões.

Neste contexto, os fabricantes têm formas eficazes de assegurar o "controlo" da rede de concessionários. Importa pois garantir um desempenho qualitativo de gestão de uma rede de retalho para tornar as estratégias definidas mais eficazes.

É neste campo, que existem divergências conceituais, dispersão e visões departamentalizadas sobre o que significa o desempenho de uma concessão automóvel, pelo que se mostrou muito relevante o desenvolvimento de um modelo de avaliação do desempenho de gestão das concessões automóveis, que seja uniforme e transversal a todas as marcas automóveis.

2. Modelos de Avaliação do Desempenho

A literatura científica sobre os sistemas de medição de desempenho fornece vários exemplos de sistemas integrados, considerando medidas financeiras e não financeiras, de longo prazo e de curto prazo, destacando-se, entre outros autores, os trabalhos de [Keegan et al. 1989], [Kaplan e Norton 1992], [Kaplan e Norton 2004], [Olve et al. 1997] e [Neely et al. 2002].

O Balanced Scorecard (BSC) [Kaplan e Norton 1992], [Kaplan e Norton 1996], [Kaplan e Norton 2004] é uma solução para as empresas implementarem e assegurarem um sistema com várias dimensões de medida do desempenho organizacional, de forma alinhada com a estratégia

definida, que possibilita identificar o que realmente cria valor e sustenta o nível de competitividade, bem como possibilita uma atuação mais rápida e eficaz ao nível de gestão [Piedade e Alturas 2011].

O BSC reflete o equilíbrio entre objetivos de curto e longo prazo, entre medidas financeiras e não financeiras, entre indicadores de tendência e de ocorrência e entre as perspetivas internas e externas [Kaplan e Norton 1996]. A divisão dos objetivos em quatro perspetivas: Aprendizagem e crescimento, Processos internos, Clientes e Económico-Financeira numa relação de causa e efeito são essenciais no que concerne ao sucesso das métricas [Piedade e Alturas 2011].

No entanto, apesar de considerar uma perspetiva de Clientes, a aplicação do BSC ao serviço pós-venda é quase ausente. Por outro lado, a literatura que lida com o serviço pós-venda, apresenta um quadro altamente fragmentado, sem foco específico em sistemas de medição de desempenho. As diversas contribuições, como por exemplo [Cohen e Lee 1990], [Armistead e Clark 1991], [Cohen et al. 1997], [Cohen e Whang 1997] evidenciam a falta de uma abordagem sistémica ligando a estratégia, a configuração da cadeia de abastecimento e sistemas de medição de desempenho. Um quadro de referência integrado (*framework*) para medição de desempenho pós-venda foi proposta por [Gaiardelli et al. 2005]. O *framework* integra alguns modelos relevantes na literatura [Lynch e Cross, 1991], [Kaplan e Norton 1992], [Supply Chain Council 2003] aplicando-os às atividades de pós-venda. O *framework* é articulado em quatro níveis: a área de negócios, o nível de processo, o nível de atividade, cujas medidas de desempenho estão focados no curto prazo, e o nível de desenvolvimento e inovação, que considera uma perspetiva de longo prazo.

Num primeiro nível, o nível estratégico da área de negócios é normalmente considerado o desempenho global financeiro pós-venda. Este pode ser medido por indicadores como os Resultados Operacionais, ROA, ROI, etc. É importante ressaltar que os resultados financeiros são gerados, tanto por resultados de mercado (quota de mercado, penetração no mercado, etc.) que têm impacto sobre as receitas, como pelo consumo eficiente de recursos (gastos). O segundo nível do *framework* é o processo. De acordo com [Lynch e Cross 1991], o desempenho do processo pode ser medido no que diz respeito à satisfação do cliente, flexibilidade e produtividade. O terceiro nível considera o desempenho da unidade organizacional pós-venda em lidar com suas atividades específicas. É útil distinguir entre atividades de *front office*, que têm impacto direto na satisfação do cliente, e as atividades de *back office*, que são responsáveis pela eficiência e prazos de entrega. Cinco dimensões de desempenho no nível de atividade são considerados: confiabilidade, capacidade de resposta, prazos de entrega internos (que são medidos principalmente por meio de indicadores não-financeiros), resíduos, custos e utilização

de ativos (que são principalmente avaliado por meio de indicadores financeiros). Finalmente, o quarto nível do *framework* proposto por [Gaiardelli et al 2005], avalia o desenvolvimento e a dimensão inovação [Kaplan e Norton 1992], com o objetivo de apurar os determinantes de futuros resultados competitivos e financeiros.

A partir da análise da literatura, torna-se claro que, para avaliar o desempenho pós-venda, se devem ter em consideração diferentes níveis de análise (unidade de negócios, processos e atividades, unidades organizacionais), bem como dimensões diferentes (desempenho financeiro e não financeiro, eficácia e eficiência, tanto a curto prazo como a longo prazo). Assim o modelo proposto por [Gaiardelli et al 2005] foi adaptado para analisar o papel do serviço pós-venda de produtos eletrónicos de consumo e tecnologias da informação [Saccani et al. 2006], tendo-se verificado uma orientação para melhorar a imagem da empresa, satisfação do consumidor e retenção de clientes.

Mais recentemente surgiu o Business Model Canvas, que é uma ferramenta de gestão estratégica, que permite desenvolver e esboçar modelos de negócio novos ou existentes. É um mapa visual pré-formatado contendo nove blocos do modelo de negócios [Osterwalder e Pigneur 2010]. Partindo de vários outros modelos, os autores propõem um modelo único de referência baseado nas semelhanças de um vasto número de conceitualizações de modelo de negócios.

Outro aspeto importante é a implementação do *framework* através de um sistema de informação, pois os sistemas de informação são uma importante ajuda no controlo, para medição e observação das atividades, permitindo fazer correções quando necessário [Alturas 2013]. Mostra-se fundamental informatizar os dados para ajudar na recolha, transformação e disponibilização da informação de desempenho, comunicar a informação de performance à gestão e utilização da informação para construir conhecimento.

3. Avaliação do Desempenho de Concessionários Automóvel

São poucas as abordagens à avaliação do desempenho de concessionários automóveis. O mais próximo foi o trabalho de [Zhang et al. 2013] sobre o desempenho do serviço no setor do aluguer de automóveis na China, tendo concluído que quanto mais personalizado for o serviço, maior a satisfação do consumidor e respetiva fidelidade.

A metodologia aqui proposta para um modelo de avaliação de desempenho de concessões automóveis foi definida a partir de um conjunto de entrevistas com profissionais do setor, e assenta numa lógica de dimensões de análise, contribuindo no seu conjunto para a obtenção de um valor final numa escala numérica.

A primeira assunção é a de que todos os aspetos da gestão interna de uma concessão estão englobados nos domínios dos rácios financeiros, indicadores de gestão operacional, elemento humano e organizacional, satisfação do cliente, infraestruturas da concessão e notoriedade da concessão. A cada um destes domínios é atribuído um peso relativo.

Por sua vez, cada um destes domínios é composto por um conjunto de indicadores relevantes, com pesos relativos. O total do peso dos indicadores relevantes de um determinado domínio corresponde ao peso que esse domínio tem comparativamente com os restantes.

A área financeira foi considerada uma dimensão crítica do negócio, suportada na dimensão dos rácios financeiros. Estes rácios devem ser entendidos como uma ferramenta que permite racionalizar a enorme quantidade de informação económica e contabilística.

É certo que uma análise exaustiva da situação financeira de uma empresa não se limita a um conjunto de rácios, mas o que se pretende é identificar o desempenho da gestão da concessão através dos indicadores mais importantes e atribuir-lhes individualmente um peso relativo.

Esta análise estática, apoiada principalmente nas principais demonstrações financeiras, é a mais adequada ao objetivo proposto, porque permite estabelecer valores de referência credíveis por sector de atividade.

Em resumo, o método dos rácios financeiros cumpre os objetivos propostos porque consegue identificar a situação financeira da empresa e estabelecer valores de referência credíveis.

Estes valores de referência [Cohen, 1996] podem corresponder aos valores obtidos quando calculados em bases estatísticas, determinados a partir das contas de empresas-tipo ou que correspondem a objetivos que os dirigentes desejam atingir.

Devido às caraterísticas do modelo organizativo do setor, que permanece idêntico ao longo dos tempos, não existe margem para alterar a estratégia que a marca define à priori, restando ao concessionário focar-se na gestão operacional do negócio, que se assume como uma dimensão de análise crítica do negócio de distribuição automóvel.

Para tal, é necessário definir indicadores para analisar o estado da concessão nas suas áreas de atividade principais identificadas no [Mitac, 1997], que são a área de Venda de Viaturas Novas e Usadas, de Assistência e de Peças.

Assim sendo, o conhecimento profundo de todos os aspetos operacionais das atividades principais do negócio é fundamental para aumentar a eficácia e a produtividade, que se traduzirá no aumento de rentabilidade da empresa.

As pessoas e a forma como se relacionam nas organizações são um fator crítico de sucesso para qualquer empresa, devendo ser enquadradas numa dimensão de análise própria.

A gestão dos recursos humanos numa empresa é hoje uma tarefa complexa de carácter instável que envolve atividades relacionadas com o recrutamento, formação, desenvolvimento, compensação e despedimento quando necessário.

Esta gestão de recursos humanos exige que as empresas desenvolvam modelos de organização e níveis de comunicação e liderança, que têm que se adaptar às constantes mudanças, sob pena de desaparecerem dos mercados onde estão inseridas.

Segundo [Camara, Guerra e Rodrigues, 2001] a nova relação de trabalho implica uma grande flexibilização do conteúdo funcional dos indivíduos nas organizações, tornando-se pois evidente, que a rentabilidade das concessões tem uma forte correlação com o fator humano.

Para a obtenção de bons resultados, é necessário criar estruturas organizacionais muito flexíveis e com funções adequadas.

É necessário também criar bons esquemas de remuneração e ambientes de trabalho, que motivem os trabalhadores a atingir os objetivos que são pedidos a uma concessão.

Por outro lado, os gestores confrontam-se com a necessidade de ajustar o número de pessoas, as respetivas funções e o próprio modelo organizacional, sempre que tal seja necessário.

Assim, uma empresa terá uma vantagem competitiva em relação a outra, quanto melhor for a competência dos seus trabalhadores, a liderança, a motivação, a formação, a eficácia do quadro organizacional e a riqueza de funções, o esquema de remuneração, entre outros aspetos adiante desenvolvidos.

O grau de satisfação e confiança que os clientes depositam numa concessão, influência de forma determinante a sua rentabilidade global no médio e longo prazo, pelo que deve ser considerada como uma componente critica do negócio.

O crescente nível de informação e conhecimento faz com que o consumidor de hoje seja muito mais exigente no seu relacionamento com o vendedor, alterando-se assim a relação fornecedor/cliente.

Como o automóvel um bem duradouro, nesse período de tempo de utilização, é posta à prova a relação com o produto e com a concessão, que resultará num exercício de avaliação da satisfação, determinante para continuar ou não a ser cliente.

Deste modo, se uma concessão conseguir atingir um elevado grau de satisfação, é provável que o cliente se mantenha fiel á marca e à concessão, comprando uma nova viatura e continuando a fazer as manutenções de oficina no mesmo local.

Este cliente, uma vez satisfeito será um veículo de publicidade, contando a sua experiência, canalizando direta e indiretamente mais pessoas a comprarem produtos e serviços na concessão.

Caso contrário, segundo [Chougule et al, 2012] os clientes insatisfeitos neste setor não só divulgam a sua má experiência, contribuindo para uma má imagem do produto e de quem o vendeu ou prestou um serviço relacionado com o automóvel.

O nível de infraestruturas numa concessão deve ser considerado uma dimensão crítica de performance de uma concessão, porque influência os colaboradores, bem como a satisfação dos clientes.

Entende-se por infraestruturas, todos os meios materiais existentes no local da concessão e sua área envolvente, que são utilizados por colaboradores e/ou clientes.

Para dar a conhecer os produtos que a organização vende e os serviços que presta é necessário dispor de um conjunto de meios que vão desde as instalações da concessão ao nível de equipamento necessário para se poder desenvolver a atividade.

Para além da quantidade, é também necessário que esses meios postos à disposição tenham a qualidade necessária para que a atividade de desenvolva da melhor forma.

Assim, para uma concessão automóvel é imprescindível que estejam disponíveis, um conjunto de meios físicos, para receber o cliente desde o início do processo de venda até à visita deste para reparação da viatura.

Estes meios devem também permitir aos colaboradores desenvolver as suas tarefas da melhor forma, aumentando a produtividade do trabalho e consequentemente a rentabilidade da empresa.

A notoriedade da concessão é a imagem ou conjunto de perceções que os clientes têm da concessão. Vários fatores contribuem para que os clientes estabeleçam uma determinada imagem acerca da concessão.

A compra de um automóvel é muitas vezes o resultado da tomada de decisões com uma carga emotiva e subjetiva muito forte, em vez de decisões baseadas em argumentos objetivos e racionais.

Esta carga emotiva está associada a uma estratégia de marketing que os fabricantes desenvolvem há muito tempo com poderosos meios financeiros e que podemos considerar de grande sucesso.

Lógico será dizer que todo o marketing desenvolvido pela marca, afeta a imagem da concessão, mas esta também pode contribuir para a retenção da marca. Segundo [Verhoef et al, 2007] tanto a marca como o concessionário podem contribuir para reter o cliente, apesar desta relação variar em função da imagem da marca (volume, prestigio, económica).

O número de veículos que a concessão vendeu também contribui para a sua notoriedade que aumentará à medida do número de veículos vendidos pela concessão.

A notoriedade das pessoas que estão ligadas à concessão faz com que o público em geral associe essa notoriedade à da própria concessão.

A satisfação e a competência dos colaboradores contribuem de forma decisiva para a imagem da concessão porque são percebidas pelos clientes através do seu relacionamento. Estes clientes por sua vez transmitiram as experiências que tiveram, para o exterior.

Por último, a responsabilidade social das empresas ajuda-as a divulgar a sua própria imagem no meio onde estão inseridas, através dos seus colaboradores, clientes e beneficiários diretos das ações sociais implementadas pela concessão.

Em suma, quanto maior for a notoriedade que uma concessão conseguir criar, maior será o número de potenciais clientes, nível de faturação da empresa e grau de popularidade.

4. Metodologia

Com base na revisão da literatura e num conjunto de entrevistas realizadas a responsáveis por concessões automóveis, foi desenvolvido um *framework* para avaliação do desempenho de uma concessão automóvel.

Esta abordagem de análise do desempenho teve também como base a experiência profissional adquirida por um dos autores do presente artigo, em várias marcas automóveis e numa empresa de consultoria automóvel de âmbito internacional. Ao longo do seu percurso profissional foi cruzando as várias perspetivas dos fabricantes acerca do desempenho das suas redes de retalho, que conjuntamente com um trabalho empírico resultou numa metodologia de análise da performance das concessões automóveis. Para criar a respetiva metodologia, foram realizadas várias reuniões com profissionais do setor em diversas atividades do setor de retalho automóvel. Como resultado destas reuniões foi possível definir os seis domínios de análise presentes no *framework* proposto, bem como os respetivos indicadores (indicadores críticos de análise do desempenho).

O passo seguinte consistiu na elaboração um questionário dirigido a quadros do setor automóvel e outros profissionais com experiência relevante de gestão e conhecedores deste setor.

Este questionário teve como objetivo determinar o peso dos domínios e indicadores críticos de análise do desempenho de gestão de uma concessão, sendo entrevistadas presencialmente 50 pessoas.

O questionário foi elaborado da seguinte forma:

1. Inicialmente apresentaram-se os domínios de análise da atividade de uma concessão automóvel, solicitando-se a sua classificação na seguinte escala:

– Sem importância

– Pouco importante

– Importante

– Grande importância

– Importantíssimo

2. Após a determinação do peso relativo dos domínios de análise, apresentaram-se todos os fatores críticos do negócio (indicadores) que foram classificados com a mesma escala.

O peso de um domínio de análise corresponde ao somatório dos pesos dos indicadores críticos que lhe estão associados.

3. O resultado de cada indicador crítico foi multiplicado pelo peso do respetivo domínio de atividade.

4. O somatório de todos os indicadores críticos, obtidos de acordo com o respetivo domínio de atividade, corresponde a 100%.

5. Proposta de Modelo

Os indicadores críticos do negócio representam os fatores de atuação da concessão para melhorar a sua gestão e desempenho e fazem parte de um domínio da atividade da concessão.

Estes indicadores são variáveis internas que permitam classificar as tomadas de decisão dos gestores relativamente à esfera de ação da concessão.

Faz parte do trabalho dos gestores maximizar os indicadores relevantes de forma a conseguir uma excelente performance de gestão da concessão.

Os resultados obtidos, ou seja, os pesos relativos obtidos, estão identificados na Figura 1.

Domínios	Peso
Rácios Financeiros	14,5%
Gestão Operacional	17,4%
Elemento Humano e Organizacional	19,1%
Satisfação do Cliente	19,8%
Infraestruturas da Concessão	15,3%
Notoriedade da Concessão	13,9%
TOTAL	100%

Figura 1 – Domínios de Análise de desempenho de uma concessão automóvel

De seguida descrevem-se quais os indicadores que compõem os respetivos domínios da atividade da concessão:

Rácios Financeiros – Indicadores	Peso
Resultados Líquidos s/ Vendas	1,702%
Nível de endividamento	1,739%
Rotação das Vendas (Ativo)	1,588%
Rentabilidade dos Capitais Investidos	1,617%
Rentabilidade dos Capitais próprios	1,617%
Liquidez Geral	1,503%
Autonomia financeira	1,560%
Solvabilidade	1,617%
Enc. Financeiros s/ Vendas	1,532%
SUBTOTAL	14,5%

Figura 2 – Indicadores do domínio Rácios Financeiros

Gestão Operacional - Indicadores	**Peso**
Viaturas Novas	4,703%
Viaturas Usadas	4,035%
Assistência	4,552%
Peças	4,078%
SUBTOTAL	17,4%

Figura 3 – Indicadores do domínio Gestão Operacional

Desta forma foi necessário separar o domínio Gestão Operacional em quatro subdomínios: viaturas novas, viaturas usadas, assistência e peças.

Gestão Operacional - Viaturas Novas	
Realização do Potencial de vendas da AC	0,260%
Realização dos bónus comerciais da marca	0,243%
Política de descontos e ofertas	0,236%
Política de Financiamento e Seguros	0,242%
Política de viaturas de demonstração	0,220%
Animação dos Pontos de Venda	0,201%
Exposição de viaturas nos Pontos de Venda	0,223%
Envolvência do espaço reservado à entrega de VN	0,211%
Horários de funcionamento das áreas de vendas	0,227%
Ações de Marketing/Publicidade desenvolvidas	0,242%
Organização do CRM	0,247%
Prospeção comercial	0,240%
Reuniões periódicas c/a equipa comer.	0,245%
Organização da atividade dos vendedores	0,238%
Política de comissionamento da equipa comercial	0,233%

Especialização dos vendedores	0,228%
Sectorização da equipa comercial	0,228%
Mapas de controlo da atividade comercial	0,240%
Performance dos Chefes de vendas	0,250%
Performance dos vendedores	0,248%

Figura 4 – Indicadores do subdomínio Viaturas Novas

Gestão Operacional – Viaturas Usadas	
Política de Retomas (lucro ou lucro 0)	0,339%
Qualidade da avaliação de Retomas	0,351%
Canais de escoamento das retomas	0,337%
Autonomia da equipa da atividade de usados	0,283%
Existência de uma equipa de vendas separada	0,282%
Nível de stock de usados	0,299%
Idade do stock de usados	0,307%
Taxa de retomas	0,306%
Existência de política de custo recondicionamento	0,306%
Prospeção comercial	0,318%
Existência de um espaço próprio p/ usados	0,295%
Ficha de avaliação das viaturas usadas	0,306%
Existência de procedimentos operativos nesta atividade	0,307%

Figura 5 – Indicadores do subdomínio Viaturas Usadas

Gestão Operacional - Assistência	
Taxa Fidelização	0,290%
Taxa de Absentismo	0,268%
Taxa de formação dos técnicos	0,269%
% horas internas não produtivas	0,271%
Taxa de Produtividade	0,280%
Taxa de Eficiência	0,282%
Taxa de Eficiência Global	0,269%
Número de reclamações de oficina	0,265%
Execução correta das reparações	0,273%
Tempo de Imobilização Viatura	0,262%
Tempo de espera para reparação viatura	0,271%
Preço da mão-de-obra	0,255%
Execução de todos serviços de oficina de mecânica	0,255%
Existência serviço de pintura na concessão	0,255%
Existência serviços rápidos	0,237%
Existência equipamento recomendado p/marca/org.	0,262%
Taxa de Absorção Após-venda	0,289%

Figura 6 – Indicadores do subdomínio Assistência

Gestão Operacional - Peças	
Cobertura regional do mercado peças na AC	0,389%
Política de prospeção de mercado	0,358%
Taxa de encomendas planeadas	0,368%
Política de crédito	0,383%
Eficácia recebimentos	0,406%
Rotação do stock	0,391%
Idade do stock	0,379%
Valor do stock	0,377%
Existência de vendedores itinerantes	0,340%
Existência de política de merchandising	0,329%
Taxa de Atendimento	0,358%

Figura 7 – Indicadores do subdomínio Peças

Elemento Humano e Organizacional - Indicadores	
Número de pessoas	1,453%
Relação trabalhadores Diretos/Indiretos	1,445%
Formação académica dos empregados	1,422%
Formação específica de certas funções	1,491%
Formação específica da marca representada	1,407%
Existência de uma política de objetivos	1,506%
Esquema de remuneração da empresa	1,460%
Esquema de avaliação dos colaboradores	1,468%
Rotatividade dos empregados	1,514%
Nível de comunicação	1,415%
Estrutura das funções organizacionais	1,369%

Motivação dos colaboradores	1,514%
Liderança das chefias	1,658%
SUBTOTAL	19,1%

Figura 8 – Indicadores do domínio Elemento Humano e Organizacional

Satisfação do Cliente - Indicadores	
Existência de Inquéritos SSI/CSI	1,806%
Nº Reclamações dos Clientes	1,940%
Taxa de retenção de Clientes de Oficina	1,788%
Aspeto visual dos colaboradores	1,628%
Aspeto Geral e conforto do espaço circ. Clientes	1,566%
Preço justo pago pelo produto ou serviço prestado	1,815%
Cumprimento dos prazos estabelecidos	1,842%
Tempo de espera para ser atendido	1,833%
Acompanhamento do Cliente	1,806%
Satisfação das necessidades do cliente	1,869%
Qualidade do atendimento	1,931%
SUBTOTAL	19,8%

Figura 9 – Indicadores do domínio Satisfação do Cliente

Infraestruturas da Concessão - Indicadores	
Nível de investimento	1,873%
Layouts interiores e exteriores preconizados p/ Marca	1,829%
Cobertura do território	1,916%
Área das instalações	1,884%
Funcionalidade do "espaço" produtivo	1,948%
Funcionalidade do "espaço" clientes	1,884%
Acessibilidade da concessão	1,959%

Nível de equipamento	1,970%
SUBTOTAL	15,3%

Figura 10 – Indicadores do domínio Infraestruturas da Concessão

Notoriedade da Concessão - Indicadores	
Quantidade de Ações de publicidade da concessão	1,532%
Dimensão da concessão	1,640%
Notoriedade das pessoas afetas a concessão	1,715%
Notoriedade da marca que a concessão representa	1,898%
Parque circulante da concessão	1,769%
Perceção do nível de satisfação dos empregados	1,758%
Perceção do nível de competência dos empregados	1,877%
Responsabilidade social da concessão	1,758%
SUBTOTAL	13,9%

Figura 11 – Indicadores do domínio Notoriedade da Concessão

6. Análise dos Resultados

- Média

A análise aos 50 questionários preenchidos diz-nos que a Satisfação do Cliente é o domínio que os inquiridos consideram de maior importância na gestão de uma concessão, atingindo 4,52 na escala de 0 a 5.

A seguir, encontra-se o Elemento Humano e Organizacional com uma importância média de 4,36.

Em terceiro lugar aparecem os Indicadores de Gestão operacional com uma importância média de 3,96, onde a atividade de viaturas novas assume o maior destaque.

Continuando a análise da importância dos domínios de gestão por ordem decrescente, as Infraestruturas da Concessão e os Rácios Financeiros apresentaram valores médios de 3,48 e 3,30 respetivamente.

Por último, os aspetos relacionados com a Notoriedade da Concessão foram considerados os menos importantes para os inquiridos, com um valor médio de 3,18.

Quanto aos indicadores críticos convém salientar que a Liderança das Chefias e a Reclamação de Clientes registaram a importância média mais elevada (4,36) seguida da Qualidade do atendimento (4,34).

Por outro lado, os inquiridos atribuíram a menor importância à Dimensão da Concessão (3,04) e à Quantidade de Ações de Marketing realizadas (2,84).

- Mediana

Se analisarmos os resultados da mediana de cada um dos domínios de análise da performance de gestão de uma concessão, concluímos que na pergunta relativa à Satisfação do Cliente, a atribuição de Importância Elevadíssima foi a "ocorrência central".

Em termos práticos, quer dizer que a Importância Elevadíssima foi a resposta mais escolhida pelos inquiridos (58%).

O Nº de Reclamações de Clientes, que faz parte do domínio mencionado no parágrafo anterior, teve um comportamento idêntico, em que 52% dos inquiridos atribuíram uma Importância Elevadíssima, sendo esta a resposta central à pergunta em causa.

Em termos gerais, o valor da mediana da maior parte dos indicadores críticos é 4, na escala de 0 a 5, o que significa que os inquiridos atribuem com muita frequência Grande Importância às perguntas efetuadas no questionário.

- Desvio-Padrão

De uma forma geral não existe grande dispersão de valores relativamente à importância dos domínios e indicadores analisados.

A Satisfação do Cliente é o domínio que apresenta valores com menor dispersão, com um desvio-padrão de 0,614.

Quer isto dizer que em média, a importância atribuída à Satisfação do Cliente afasta-se da sua média 0,614 (na escala de 1 a 5).

Por outro lado a Notoriedade da Concessão é o domínio com maior dispersão, ou seja, as respostas relativas à sua importância, são as que se afastam mais da média da importância atribuída no questionário realizado.

Após a definição de todos os pesos dos indicadores, foi elaborado o *framework* que será alimentado a partir de um questionário final em que o avaliador atribuirá uma classificação de 1 a 5 a cada indicador, de cuja ponderação com os respetivos pesos resultará num número. O somatório de cada número de cada indicador dará uma classificação final.

Cada indicador está devidamente parametrizado, ou seja, a atribuição da classificação (1 a 5) estará escrita de forma explícita.

O passo seguinte será o desenvolvimento do sistema de informação, para implementar o modelo proposto. Nesse sentido foi realizada uma análise inicial utilizando a linguagem UML [Booch et al. 2005], da qual se apresentam o Diagrama de Classes (Figura 12) e o Diagrama de *Use Cases* (Figura 13).

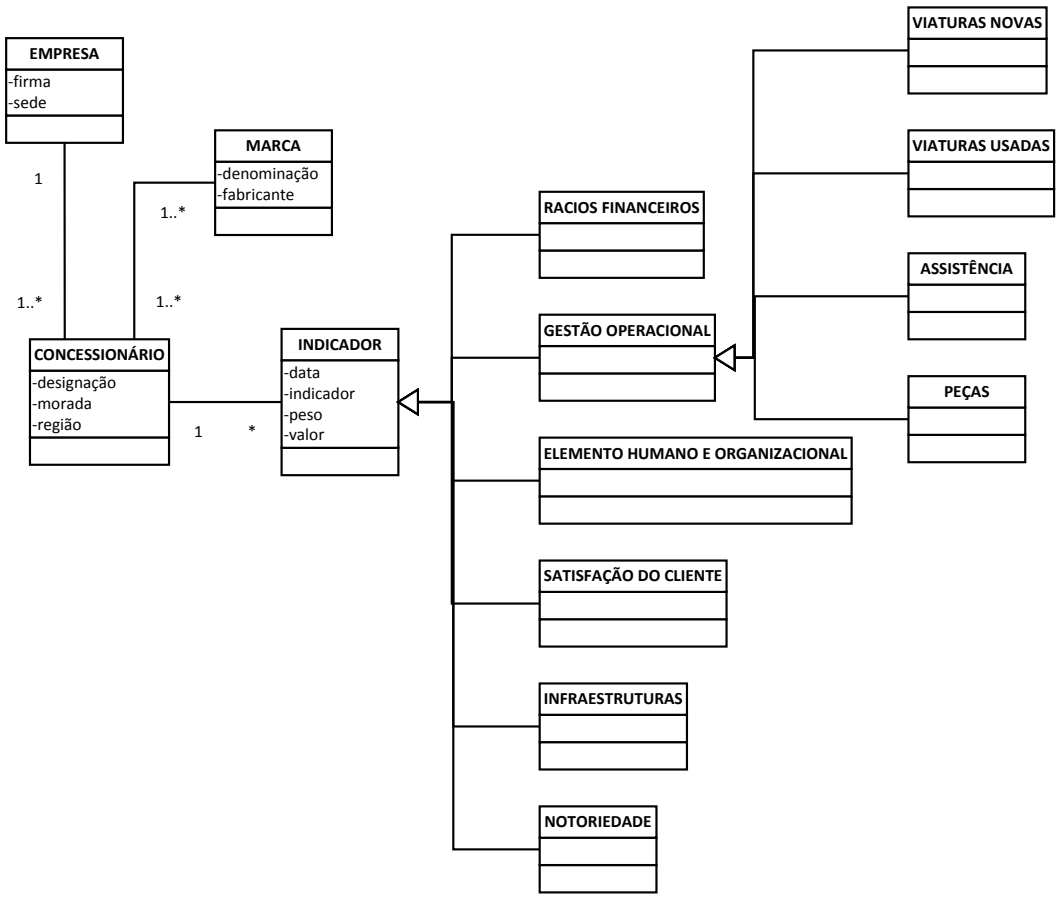


Figura 12 – Diagrama de Classes do Modelo de análise de desempenho de uma concessão automóvel

Segundo [Ramos 2007], o diagrama de classes da UML está essencialmente vocacionado para representar a componente estrutural de um sistema, nomeadamente a estrutura de informação que o suporta. Essa estrutura é habitualmente estática durante um considerável período de tempo.

Por seu lado os diagramas de *use cases* são ferramentas simples para descrever o comportamento de um sistema. Um *use case* deve conter uma descrição textual de todas as funcionalidades que o utilizador futuramente poderá vir a utilizar no sistema através da sua interface [Alturas 2013].

O passo seguinte passará por criar um sistema de informação, baseado em *dashboards*, para implementar o *framework*.

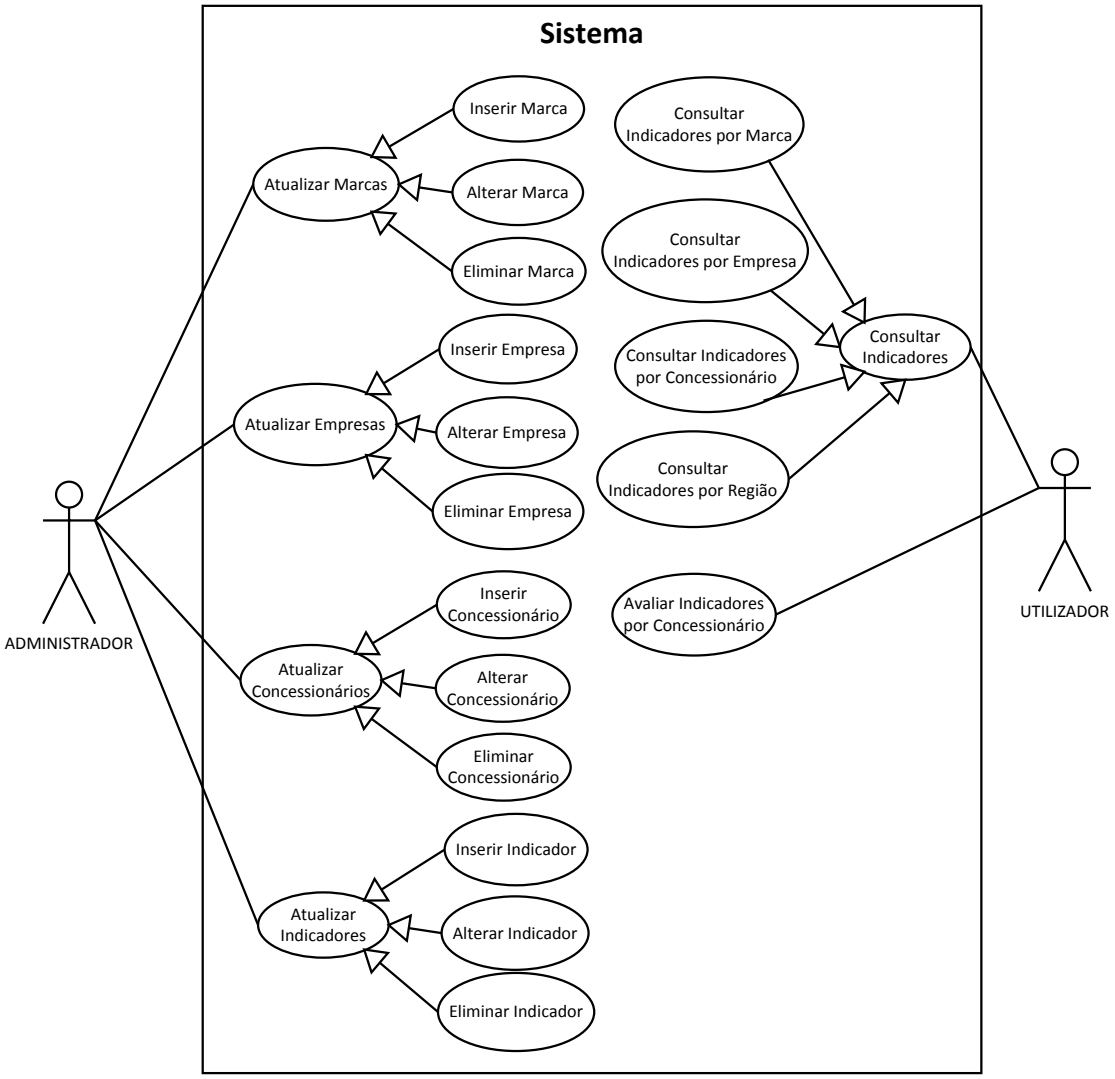


Figura 13 – Diagrama de *Use Cases* do Modelo de análise de desempenho de uma concessão automóvel

7. Limitações

A virtude desta metodologia está na classificação da gestão de uma empresa de forma rápida, abrangente e integradora.

Contudo, esta avaliação reflete supostamente uma situação presente e pode não espelhar o passado. Deste modo, a utilização inicial deste modelo numa empresa não permite a análise temporal que é fundamental para uma interpretação global da gestão das empresas.

Os resultados presentes podem também não ser suficientes para projetar o futuro, uma vez que a informação obtida neles não tem em conta circunstâncias futuras que alteram conceitos subjacentes à performance de gestão. Um exemplo desta problemática verificou-se no tempo de espera das peças automóveis. A introdução de sistemas integrados de gestão, processos logísticos aperfeiçoados e o maior interesse dos fabricantes na satisfação do cliente, fizeram alterar o conceito de tempo de espera nas reparações automóveis.

Outra problemática prende-se com a eventual dificuldade em obter dados no mesmo período de tempo. Por norma as informações financeiras e de satisfação de clientela, entre muitas outras, têm períodos de recolha específicos que podem não ser coincidentes, pondo em causa a veracidade do desempenho de gestão num determinado período.

Relativamente à dimensão financeira já se referiu a importante função dos rácios financeiros, contudo nem sempre são suficientes para avaliar a verdadeira situação financeira de uma empresa, que obriga à análise de informação complementar.

Um outro ponto fraco do modelo, decorre da sua própria estrutura que não assimila o dinamismo inerente a qualquer setor. Novas tendências de negócios podem não ser enquadráveis assim como a pressão da indústria, dos clientes, dos colaboradores, etc, podem fazer alterar os parâmetros que servem de base à avaliação.

Por último, este modelo sofre ainda da limitação que existe em transformar aspetos que não são exatos em expressões numéricas.

8. Conclusão

Da literatura analisada conclui-se que existem muitas metodologias de análise de desempenho das empresas que consideram diversos aspetos que são tratados na metodologia proposta.

O retalho automóvel é constantemente avaliado pelas marcas que representam, contudo essa análise é departamentalizada nas atividades principais do negócio e orientadas para atingir objetivos comerciais.

O questionário permitiu validar a metodologia de análise da performance de uma concessão de acordo com as seguintes dimensões: rácios financeiros, indicadores de gestão operacional, elemento humano e organizacional, satisfação do cliente, infraestruturas da concessão e notoriedade da concessão. As dimensões de análise apresentam pesos relativos relativamente idênticos, destacando-se a satisfação do cliente como a mais importante e a notoriedade da concessão como menos importante.

9. Referências

Alturas, B., *Introdução aos Sistemas de Informação Organizacionais*, Edições Sílabo, 2013.

Armistead, C. e Clark, G., *Customer Service and Support: Implementing Effective Strategies*, FT Books, London, 1992.

Booch, G.; Rumbaugh, J. e Jacobson, I., *The Unified Modeling Language User Guide (2nd Edition)*, Addison-Wesley, 2005.

Camara, P. B.; Guerra, P. B. e Vicente Rodrigues, J., Recursos Humanos & Sucesso Empresarial. Humanator, 4ª edição, Lisboa, (2001).

Chougule, R.; Khare, V. R.; e Pattada, K., "A fuzzy logic based approach for modeling quality and reliability related customer satisfaction in the automotive domain", *Expert Systems with Applications*, 40, 2013, 800–810.

Cohen, E., *Análise Financeira*, Editorial Presença, 1996.

Cohen, M. A. e Lee, H. L., "Out of touch with customer needs? Spare parts and after sales service", *Sloan Management Review*, 31, 2, 1990, 55-66.

Cohen, M. A. e Whang, S., "Competing in product and service: a product life-cycle model", *Management Science*, 43, 4, 1997, 535-545.

Cohen, M. A., Zeng, Y. e Agrawal, V., "Service parts logistics: a benchmark analysis", *IEEE Transactions*, 29, 1997, 627-39.

Gaiardelli, P.; Saccani, N. e Songini, L., "Performance measurement systems in the after-sales service: an integrated framework", *International Journal of Business Performance Measurement*, 2005.

Kaplan, R. S. e Norton, D. P., "The balanced scorecard. measures that drive performance ", *Harvard Business Review*, 70, 1, 1992, 71-79.

Kaplan, R. S. e Norton, D. P., *The Balanced Scorecard - Translating Strategy in Action*, Harvard Business School Press, Boston, MA, 1996.

Kaplan, R. S. e Norton, D. P., *Strategy Maps*, Harvard Business School Press, Boston, MA, 2004.

Keegan, O. P.; Eiler, R. G. e Jones, C. R., "Are your performance measures obsolete?", *Management Accounting*, June, 1989, 45-50.

Lynch, R. L. e Cross, K. F., Measure Up: *The Essential Guide to Measuring Business Performance*, Mandarin, London, 1991.

Mitac Ltd, "The Guide To Better Performance", MPG Books Limited, 1997.

Neely, A. D., Adams, C. e Kennerley, M., *The Performance Prism: The Scorecard for Measuring and Managing Business Success*, FT Prentice-Hall, London, 2002.

Olve, N. G.; Roy, J. e Wetter, M., *Performance Drivers: A Practical Guide to Using the Balanced Scorecard*, John Wiley & Sons, Chichester, 1997.

Osterwalder, A. e Pigneur, Y., *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, John Wiley & Sons, 2010.

Piedade, O. e Alturas, B., "Implementação do Balanced Scorecard numa Companhia de Seguros de Assistência", *CISTI 2011 - 6ª Conferência Ibérica de Sistemas e Tecnologias de Informação*, Chaves, Portugal, Vol. I, 2011, 409-413.

Ramos, P. N., *Desenhar Bases de Dados com UML*, 2ª Edição, Sílabo, 2007.

Saccani, N.; Songini, L.e Gaiardelli, P, "The role and performance measurement of after-sales in the durable consumer goods industries: an empirical study ", *International Journal of Productivity and Performance Measurement*, 55, 3/4, 2006, 259-283.

Supply Chain Council, *Supply-Chain Operations Reference Model (SCOR)*, Ver. 6.0, Supply Chain Council, Pittsburgh, PA, www.supply-chain.org, 2003.

Verhoef, P. C.; Langerak, F.; Donkers B., "Understanding brand and dealer retention in the new car market: The moderating role of brand tier", *Journal of Retailing*, 8, 2007, 97–11.

Zhang, M.; Xie, Y.; Huang, L. e He, Z., "Service quality evaluation of car rental industry in China", *International Journal of Quality & Reliability Management*, 31, 1, 2014, 82-102.

Modelos de Simulação de Processos por Junção de Componentes

Manuel Almeida [1], José Luís Pereira [2]

1) Universidade do Minho, Portugal
a4757@alunos.uminho.pt

2) Universidade do Minho, Portugal
jlmp@dsi.uminho.pt

Resumo

Com este artigo pretende-se abordar o papel que um repositório de componentes de simulação previamente desenvolvidos pode desempenhar em projetos de simulação de processos de negócio. Recorrendo aos conhecidos *padrões workflow* (*workflow Patterns*), resultantes da identificação de situações recorrentemente encontradas em modelos de processos organizacionais, e utilizando mecanismos específicos de componentização disponíveis em algumas ferramentas de simulação, constroem-se componentes a (re)utilizar no desenvolvimento de modelos de simulação. Como exemplo, usou-se um processo de negócio - "*Reclamação de Encomendas de Clientes*" - em vigor numa empresa do sector têxtil, em que um dos autores colabora.

Palavras-chave: Processos de Negócio, Modelos de Processos, Desenvolvimento por Junção de Componentes, Simulação

1. Introdução

As organizações, perante um ambiente complexo e cada vez mais competitivo à escala global, necessitam de se adaptar rapidamente às mudanças que ocorrem à sua volta. É, pois, importante que estas sejam dotadas de mecanismos que permitam uma adaptação contínua às exigências a que estão sujeitas.

Contudo, as mudanças organizacionais apresentam sempre riscos, devido ao impacto que estas têm nos processos organizacionais estabelecidos e, de uma forma geral, nos vários componentes da organização, sendo que muitos esforços de reengenharia ou reformulação de processos acabam por falhar quando são levados à prática [Gregoriades et al. 2008].

A simulação computorizada, ao permitir a avaliação de cenários futuros, ou a otimização dos processos existentes, sem incorrer nos custos e riscos da sua implementação prática, pode dar

um contribuir inestimável para suportar as decisões de mudança nas organizações. De facto, a simulação de processos ajuda a prever potenciais impactos das modificações nos processos de negócio atuais e comparar alternativas de implementação. Além disso, a simulação fornece uma base quantitativa que possibilita apoiar ou, por outro lado, descartar uma iniciativa de mudança.

Estas ferramentas permitem reproduzir em computador um sistema real, facilitando a alteração de variáveis caso os resultados não sejam os esperados, evitando assim, os riscos de uma experimentação no terreno. A simulação computorizada tem portanto um vasto campo de aplicação nas organizações, como defende Oliveira [2008]: "A simulação tem vindo ao longo dos tempos a contribuir para a análise e compreensão de sistemas complexos. Habitualmente utilizada e divulgada na indústria e com resultados credíveis, a simulação, com o advento da gestão de processos de negócio, passou a ter também aqui, um campo de aplicação".

A simulação de processos de negócio, enquadrada numa abordagem BPM, contribui para a avaliação de cenários futuros e de novas opções sem incorrer nos custos e riscos da sua implementação. Além disso, apoia a gestão da mudança, visto que torna compreensíveis as razões de tais mudanças. A simulação funciona como uma ferramenta de diagnóstico para a identificação de problemas e ineficiências na organização. Por meio da simulação é possível testar e analisar diferentes cenários, para verificar o impacto que têm nos sistemas e avaliar o retorno antes de realizar as mudanças ou planear a implementação da transformação do negócio.

Neste contexto, a simulação de processos de negócio, como forma expedita de analisar e explorar diferentes cenários de organização do trabalho nas organizações pode trazer grandes benefícios, em particular:

- Encontrar a "melhor" configuração para a utilização dos seus recursos;
- Desenvolver a estrutura lógica mais adequada para os processos organizacionais.

No entanto, pode ser uma atividade bastante morosa e sujeita a falhas, já que, por norma, implica a criação dos modelos de simulação de raiz. Mas da qualidade dos modelos de processo, da exatidão dos dados de entrada e dos componentes a utilizar, dependerão os resultados da simulação. Para obter resultados fiáveis e em tempo útil da simulação, deve ser assegurada a correta aplicação dos mesmos.

Neste âmbito, um trabalho de particular importância, liderado pelo Prof. Van der Aalst, vulgarmente designado por *Workflow Patterns Initiative*, resultou na identificação e documentação de um conjunto de "situações" recorrentemente observadas em processos de negócio – os chamados *padrões workflow* [Aalst et al.. 2003a].

Neste contexto, emerge a seguinte questão de investigação, que está subjacente ao trabalho referido neste artigo:

> *Como tornar mais eficientes e eficazes os esforços de simulação em iniciativas de melhoria de processos de negócios?*

Em termos simples, pretende-se analisar o contributo que um repositório de componentes de simulação (correspondentes aos referidos padrões de *workflow*) pode dar no desenvolvimento de modelos de simulação. A expectativa é que, qualquer modelo de simulação de processos, ao recorrer a estes componentes pré-fabricados e previamente testados, possa ser rapidamente desenvolvido, sendo que fica também parcialmente validado.

Em termos de estrutura, este artigo começa por apresentar, resumidamente, os conceitos em torno da abordagem BPM, em particular o conceito fundamental que lhe está subjacente – *processo de negócio*. Seguidamente, aborda-se o papel da simulação de processos em que, recorrendo a um processo concreto, se procede à identificação de alguns dos padrões *workflow* que o integram e que serão traduzidos por componentes do respetivo modelo de simulação, previamente desenvolvidos e disponibilizados num repositório.

2. Gestão por Processos Organizacionais (BPM)

Até muito recentemente, as organizações estruturavam sua forma de operar com base numa abordagem hierárquica e funcional do trabalho, inspirada no modelo Taylorista. Ora, a excessiva segmentação hierárquica e funcional, perante a maior agressividade e turbulência dos mercados, tornaram os modelos de gestão tradicionais obsoletos. As organizações procuraram novas formas de se estruturar pois necessitam de ser ágeis e eficazes para responderem mais facilmente às pressões e oportunidades do mercado. De facto, "*uma organização estruturada em unidades funcionais semiautónomas, mais preocupadas com o seu próprio desempenho do que com o desempenho da organização como um todo, não está preparada para compreender as necessidades do mercado e responder de forma efetiva às suas solicitações*" (Pereira 2004). É precisamente neste contexto que surge o conceito de processo organizacional, como forma de integrar todos os recursos da organização, independentemente da divisão funcional existente.

Face a esta necessidade de mudança, surge um novo paradigma organizacional em que a Gestão por Processos de Negócio (BPM - *Business Process Management*) aparece como um dos pilares fundamentais e que preconiza uma constante mudança e melhoria dos processos de negócio na prossecução dos objetivos das organizações.

A popularização do conceito BPM entre as organizações, estimulou o mundo das TI no sentido de desenvolver produtos e ferramentas para suportar o conceito de processo organizacional nas

suas várias vertentes. A gestão por processos requer um compromisso da organização como um todo, desde o nível operacional até ao planeamento estratégico a longo prazo. De acordo com Harmon, citado em Jeston e Nelis [2006], o BPM pode ser definido como "*uma disciplina de gestão focada em melhorar a performance de uma organização pela gestão dos seus processos organizacionais*".

Com a segmentação do mercado e com a redução dos ciclos de vida da produção, entre outros fatores, os gestores e investigadores procuram estruturas organizacionais mais adequadas às condições de mercado e infraestruturas das empresas. As organizações passam a ser retratadas pelos seus processos de negócio e não pelo seu organograma.

Segundo Cichocki et al. [1998], "*um processo representa uma descrição e ordenação de atividades ao longo do tempo e espaço com o objetivo de produzir produtos ou serviços específicos, de forma a cumprir os objetivos da organização*". O processo fornece uma base conceptual para a integração e coordenação dos recursos distribuídos, tarefas e indivíduos. Na Figura 1, Aalst [2003] apresenta quatro perspetivas (processos, tarefas, recursos e informação), em que na base se encontra a perspetiva de processo, suportando esta as outras três perspetivas.

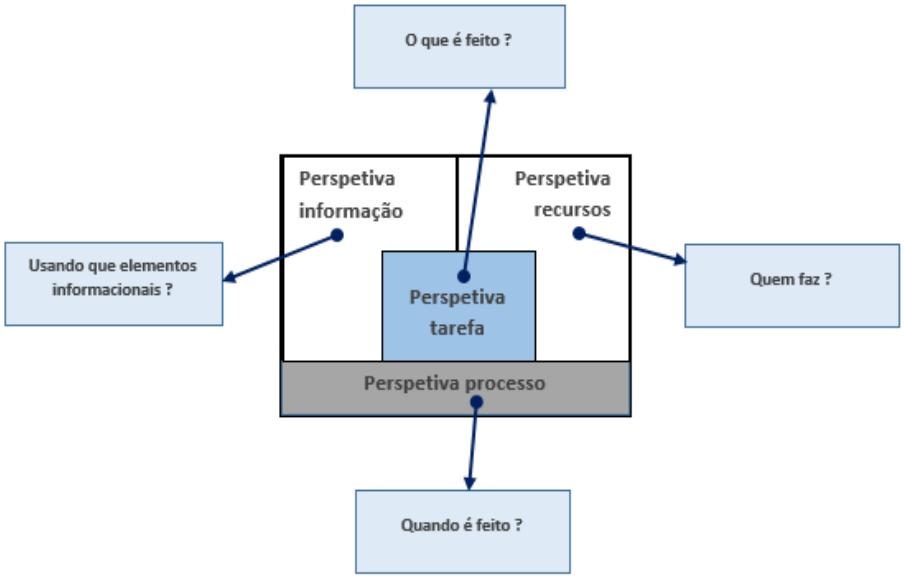


Figura 1 - Perspetivas de processo (adaptado de Aalst, 2003b)

3. Modelos de Processos

Com a abordagem BPM tornou-se importante a modelação dos processos como forma de representar o funcionamento da organização. O modelo de um processo traduz a forma como este é realizado, identifica os recursos necessários ao seu funcionamento, permite a

eliminação de etapas ambíguas, a automatização do fluxo e a indicação dos seus responsáveis.

Pode-se dividir a modelação de processos de negócio em duas fases distintas: a fase *AS-IS* e a fase *TO-BE*. Garratt et al. [2007, p. 24] definem a fase *AS-IS* como a "*a captura, documentação, e entendimento dos processos atuais*", correspondendo a fase *TO-BE* à "*fase na qual os processos podem ser alterados, permitindo melhorias baseadas nas informações e análise dos processos mapeados na fase AS-IS*".

Os modelos de processos constituem para os gestores de processo uma importante ferramenta para a tomada de decisão. DE facto, o uso de modelos é essencial para que o gestor teste alternativas para assim tomar uma decisão mais informada. A modelação de processos pode ser também uma oportunidade para a organização clarificar e formalizar os seus processos de negócio a fim de promover a representação e entendimento das suas atividades de trabalho, para minimizar problemas e maximizar o seu desempenho.

Na modelação de processos surge a necessidade de disponibilizar uma notação universal, compreensível para os utilizadores e analistas do negócio, bem como para os técnicos que irão implementar e monitorizar as soluções tecnológicas que irão suportar os processos. As notações com maior relevância atualmente são a *Business Process Model and Notation* (BPMN), Diagramas de Atividades UML e os *Event-driven Process Chain* (EPC).

A BPMN, uma das mais utilizadas, cria uma ponte entre a modelação do processo de negócio e a sua implementação. A BPMN fornece uma notação para expressar os processos de negócio num único diagrama de processo de negócio (BPD - *Business Process Diagram*). Um dos objetivos da BPMN é criar um mecanismo simples para o desenvolvimento dos modelos processos de negócio, e ao mesmo tempo poder garantir a complexidade inerente aos processos. Procura diminuir a distância entre a modelação e a implementação dos processos de negócio. Trata-se de uma notação que é compreensível por todos os utilizadores, analistas e técnicos, além de garantir que linguagens projetadas para a execução de processos de negócio, tais como a BPEL4WS (*Business Process Execution Language for Web Services*) e a BPEL (*Business Process Execution Language*) sejam visualmente representadas com uma notação comum.

O progressivo grau de importância da modelação de processos, associada aos projetos BPM, levou alguns investigadores a direcionarem os seus esforços para a modelação de processos [Aalst, 2003a], [Russel 2006], [Thom 2006]. Nas suas pesquisas foram identificados padrões de atividades, os quais representam aspetos do negócio recorrentemente encontrados em processos de negócio – os chamados *padrões workflow* (*workflow pattern*). Estes focam, sobretudo, as dependências entre as atividades de um processo, numa determinada etapa do desenvolvimento de aplicações orientadas a processos.

Um padrão é a abstração de uma forma concreta a qual se mantem recorrente em contextos específicos não arbitrários [Riehle 1996]. Mais recentemente Thom [2006] define padrão de *workflow* como a descrição de uma função recorrente encontrada com frequência em processos de negócio (notificação, aprovação, decisão, solicitação de execução de tarefa).

Neste âmbito, um trabalho de particular importância, liderado pelo Prof. Van der Aalst, vulgarmente designado por *Workflow Patterns Initiative*, resultou na identificação e documentação de um conjunto de situações recorrentemente observadas em processos de negócio – os padrões *workflow*. Este investigador defende a ideia de que a padronização de estruturas de processo e a reutilização dos padrões resultantes podem aumentar a produtividade na fase de modelação de processos, assim como na sua documentação e manutenção. Para além disso, todos reconhecem que os padrões *workflow* ajudam a promover boas práticas de desenvolvimento na modelação de processos.

4. Padrões Workflows

A *Workflow Patterns Initiative* foi estabelecida com o objetivo de identificar e descrever as situações recorrentes que surgem nos modelos de processos organizacionais. [Aalst et al. 2003b]. Com a identificação dos padrões *workflow* obtêm-se múltiplas vantagens, entre as quais:

- Uma maior eficiência no esforço de modelação pois, com poucos padrões *workflow* é possível modelar uma grande variedade de processos;
- A reutilização dos padrões *workflow* garante maior probabilidade de correção na automatização dos processos;
- A aplicabilidade em diferentes níveis de abstração, sendo que alguns padrões *workflow* podem ser úteis tanto para a modelação dos processos de negócio, como ao nível da sua implementação.

Padrões *Workflow* simplificam o processo de modelação e possuem semântica simples e bem definida. Aalst [2003b], juntamente com outros investigadores, propôs uma notação, conhecida como YAWL (*Yet Another Workflow Language*) com a qual descreveram um conjunto de padrões complexos de controlo de fluxo encontrados em processos organizacionais e que não eram diretamente representáveis por outras linguagens. A linguagem YAWL foi proposta para preencher esta lacuna, permitindo analisar qualitativamente estes padrões.

Através da análise estrutural e do particionamento de cada processo, evidenciou-se que os padrões classificados representam o conjunto necessário e suficiente de construtores para a

modelação de todos os processos. Tais padrões podem, eventualmente, simplificar o processo de modelação, visto que possuem semântica simples e bem definida. Os padrões de *workflow* identificados podem, ainda, ser utilizados para teste de completude e na comparação da expressividade das linguagens de modelação de processos existentes e futuras.

Existem diferentes perspetivas a considerar na modelação de um processo, das quais as mais relevantes são: fluxo de controlo, fluxo de dados, organizacional e tratamento de exceções. Grande parte dos trabalhos teóricos e ferramentas desenvolvidas para a modelação de processos de negócio focam a perspetiva de fluxo de controlo. A enfase dada a essa perspetiva é compreensível, dado que fornece uma visão sobre a especificação de um processo de negócio que é essencial para a avaliação de sua efetividade. As outras perspetivas assumem um papel mais secundário, uma vez que elas oferecem uma visão complementar da estrutura do processo.

Inicialmente foram identificados cerca de vinte padrões na perspetiva de controlo de fluxo, obtidos a partir de uma pesquisa minuciosa sobre os sistemas de *workflow* e modelação de processos de negócio, com o intuito de identificar as referidas situações recorrentes. Mais tarde verificou-se que existiam um certo número de padrões que necessitavam de uma descrição mais minuciosa, de forma a eliminar ambiguidades no que se refere aos conceitos que se propunham representar. Surgiu assim a identificação de vinte e três novos padrões, sendo alguns destes novos padrões baseados em especificações de padrões existentes [Aalst et al. 2003ª].

5. Simulação de processos

A eficácia e eficiência dos processos de negócio de uma organização são importantes para a sua sobrevivência num mundo cada vez mais competitivo. Se eficácia é a capacidade de atingir o objetivo com a qualidade e efeito desejado, já eficiência é a capacidade de realizar um objetivo sem desperdício de recursos e no menor tempo possível. Nesse sentido, a qualidade de um processo é definida tanto pela sua eficácia quanto pela sua eficiência.

A simulação computorizada tem vindo ao longo dos tempos a contribuir para a análise e compreensão de sistemas complexos. A simulação pode ser entendida como o processo de construção de um modelo representativo de um sistema real, bem como da realização de experiências com esse modelo com o objetivo de conhecer melhor o seu comportamento e avaliar o impacto de estratégias alternativas de operação [Shanon, 1998]. Desta forma, a simulação fornece a capacidade de rapidamente prever o sucesso ou insucesso que, determinadas decisões irão ter no contexto do processo de negócio. A simulação pode ser usada para explorar novas regras de escalonamento dos recursos, procedimentos operativos, fluxos de informação, regras de decisão, estruturas organizacionais sem ser necessário interromper o normal funcionamento do sistema [Shannon 1998].

Numa organização, os processos de negócio estão interligados e sujeitos a uma dependência, variabilidade e complexidade o que torna muitas vezes impossível prever o seu comportamento e desempenho. Urge a necessidade de construir um modelo que permita estudar o desempenho do sistema em determinadas condições, permitindo a análise do comportamento do modelo ao longo do tempo [Ball 1996], a comparação de diferentes cenários e a determinação de modelos alternativos. Ou seja, a simulação associada à gestão de processos de negócio passou a ter, também aqui, um importante campo de aplicação.

A simulação de processos pode ser considerada, um meio de avaliar o impacto das alterações efetuadas num processo, seja ele novo ou já existente, num ambiente controlado pela criação de cenários *what-if* (o que acontece se…), sendo utilizada para testar e analisar decisões antes da sua implementação no ambiente real. A simulação permite a inclusão de incerteza e variabilidade nas previsões de desempenho do processo sem afetar o ambiente real. Por outro lado, a simulação de processos pode também ser considerada uma ferramenta de gestão da mudança, uma vez que permite tornar visíveis as razões da referida mudança (evidenciando a causa-efeito) e permitindo obter explicações para o processo de decisão [Barnett 2003], sempre com o intuito da melhoria continua.

Os sistemas de simulação podem ser enquadrados por diferentes paradigmas, consoante o âmbito do sistema a ser simulado. Na simulação computorizada, é possível identificar três paradigmas distintos: Simulação Baseada em Eventos; Dinâmica de Sistemas; e Simulação Baseadas em Agentes. Neste trabalho recorre-se, unicamente ao paradigma de simulação baseada em eventos uma vez que este paradigma tem subjacente a si o conceito de processo, recorrendo à gestão de eventos para fazer avançar o fluxo de trabalho nos processos. Uma sequência de eventos, pré-definida ou aleatória, pode ser utilizada para imitar o comportamento dinâmico dos sistemas.

Para a realização de um projeto de simulação de processos, é necessário que um conjunto de passos seja seguido, de modo a que, o estudo da simulação seja eficaz. Estes passos, fases ou processos são conhecidos na literatura como “metodologias de simulação” ou “ciclos de vida de um modelo de simulação” [Law e McComas 1991].

O processo de criação de uma simulação baseada em eventos, começa com uma análise detalhada do processo que se pretende criar ou melhorar, permitindo definir os requisitos, indicadores de desempenho e exceções que vão possibilitar a construção do modelo de simulação.

Depois da análise detalhada do processo a simular, seguem-se as restantes fases do processo de simulação, desde a criação e melhoramento de modelos de simulação até à geração e

processamento de simulações. No final de cada iteração, é avaliado o resultado da execução da simulação: Se foram detetados erros ou caso se verifique que os resultados não refletem de forma correta os dados de entrada, então o modelo de simulação pode ser alterado. A simulação prossegue até que seja encontrado um comportamento satisfatório para o processo e por esta razão, o número de iterações não é fixo e depende em grande medida da complexidade do processo.

A simulação de processos é constituída por cinco fases [Schiefer et al. 2007]:

1. Fase de aquisição de conhecimentos e análise
2. Fase de modelação
3. Fase de execução
4. Fase de análise do sistema
5. Fase de desenvolvimento e reformulação

A Figura 2 fornece uma visão geral sobre o processo de simulação, as fases e as tarefas envolvidas.

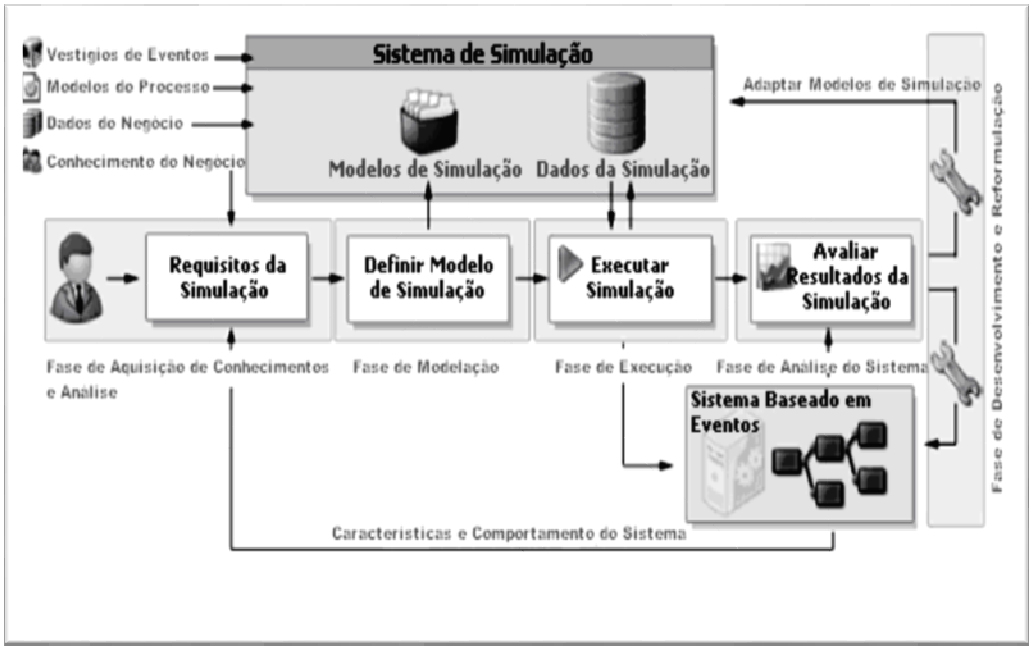


Figura 2 – Fases da Simulação por Eventos Discretos [Schiefer et al. 2007]

6. Ferramentas de simulação

Depois de analisadas algumas ferramentas de simulação comercialmente disponíveis (Arena, Simio, Simul8, etc.), a decisão sobre a ferramenta a utilizar neste trabalho recaiu sobre a ferramenta Simio. A Simio (SIMulation Intelligent Objects) é uma ferramenta baseada na simulação por eventos discretos e que possui uma estrutura de modelação fácil, rápida e flexível, permitindo avaliar rapidamente alternativas para reduzir custos e maximizar o impacto das decisões.

Ao invés da abordagem mais tradicional das ferramentas de simulação por eventos discretos, o Simio é baseado num paradigma orientado a objetos, mais genérico, na qual objetos executam em processos, mudando radicalmente a forma como os modelos são contruídos e utilizados. Os sistemas são modelados pela combinação de objetos que representam os componentes físicos destes sistemas. Possui uma biblioteca de objetos padrão que permite rapidamente modelar uma diversidade de sistemas, tais como processos de negócio. Outro aspeto a salientar, é a sua arquitetura extensível que permite aos utilizadores adicionar novos objetos (criados a partir de objetos existentes na biblioteca padrão para aplicação em problemas recorrentes) e mostrar o comportamento dos seus processos para validação e visualização através de um interface gráfico sofisticado, do tipo 3D. Alguns conceitos base desta ferramenta, a ter em atenção:

- **Objectos**

  No mundo atual, a programação orientada a objetos é aceite pela maioria dos profissionais TI e a simulação de eventos discretos move-se no sentido da modelação orientada a objetos. O uso de objetos permite reduzir a complexidade dos problemas ajudando a melhorar a confiabilidade do modelo, robustez, extensibilidade, reutilização e facilidade de manutenção. Na ferramenta Simio um objeto define-se como uma construção de um modelo autónomo que possui propriedades, estados, eventos, visão externa e lógica interna. As propriedades do objeto são valores internos que podem ser definidos pelo utilizador. Os utilizadores do Simio podem construir os seus próprios objetos e bibliotecas de objetos.

- **Processos**

  Cada modelo construído em Simio utiliza processos porque a lógica de todos os objetos são especificados usando processos. Um processo é um conjunto de atividades que ocorrem ao longo do tempo, o que pode alterar o estado do modelo.

- **Simbits**

Um dos mecanismos mais inovadores desta ferramenta, e que vai servir de base para a criação dos componentes de simulação que estão na base deste trabalho, é a sua biblioteca de Simbits. Simbits são pequenos modelos, bem documentados e em que cada um ilustra um conceito de modelação ou explica como resolver um problema comum. Cada modelo de Simbit é acompanhado de informação detalhada que descreve como construir um modelo.

Com base no trabalho desenvolvido por Aalst et al. [2003a], em que foram identificados os diversos padrões *workflow*, surge a necessidade da criação de um repositório de componentes de simulação correspondentes aos referidos padrões de *workflow*. Deste modo, espera-se que qualquer modelo de simulação de processos, ao recorrer a estes componentes previamente desenvolvidos e testados, possam ver a sua implementação acelerada e parcialmente pré-validada.

De seguida apresenta-se o processo de negócio que serve de exemplo para demostração do que poderá ser o desenvolvimento de um modelo de simulação recorrendo à junção de componentes pré-fabricados.

7. Processo "Reclamação de Encomendas de Clientes"

Para demonstrar o contributo proveniente da construção de um repositório de componentes de simulação usou-se como caso de estudo o processo "Reclamação de Encomendas de Clientes", atualmente em vigor numa empresa têxtil de que um dos autores é colaborador. Este encontra-se modelado na linguagem BPMN na sua versão de mais alto nível (ver Figura 3). Devido ao espaço disponível, decidiu-se omitir os detalhes dos subprocessos "Análise de Reclamação" e "Criar Movimentos Contabilísticos", assinalados no modelo.

Neste processo, aqui utilizado como mero exemplo, é possível reconhecer alguns padrões *workflow* no respetivo modelo. Em termos simples, este inicia-se com a receção da reclamação (normalmente, a sua origem é por via email ou carta), procedendo-se ao registo da mesma na base de dados do sistema de informação. De seguida é efetuada uma pré-análise da reclamação, por forma a solicitar ao Departamento Comercial, ao Departamento de Qualidade e ao Departamento Contabilístico informação importante para a tomada de decisão. Depois de toda a informação recolhida, esta será reencaminhada para o Departamento de Análise de Reclamações para atribuição de responsabilidades, verificando-se se o cliente tem ou não razão. Caso o cliente tenha razão é criado um documento com a informação recolhida e os elementos adicionais de acordo com a análise efetuada. Este documento é enviado para a chefia ou

administração para tomada de conhecimento, seguindo para a contabilidade para se efetuarem os movimentos necessários. Quer a reclamação seja ou não aceite, será sempre enviada uma resposta ao cliente.

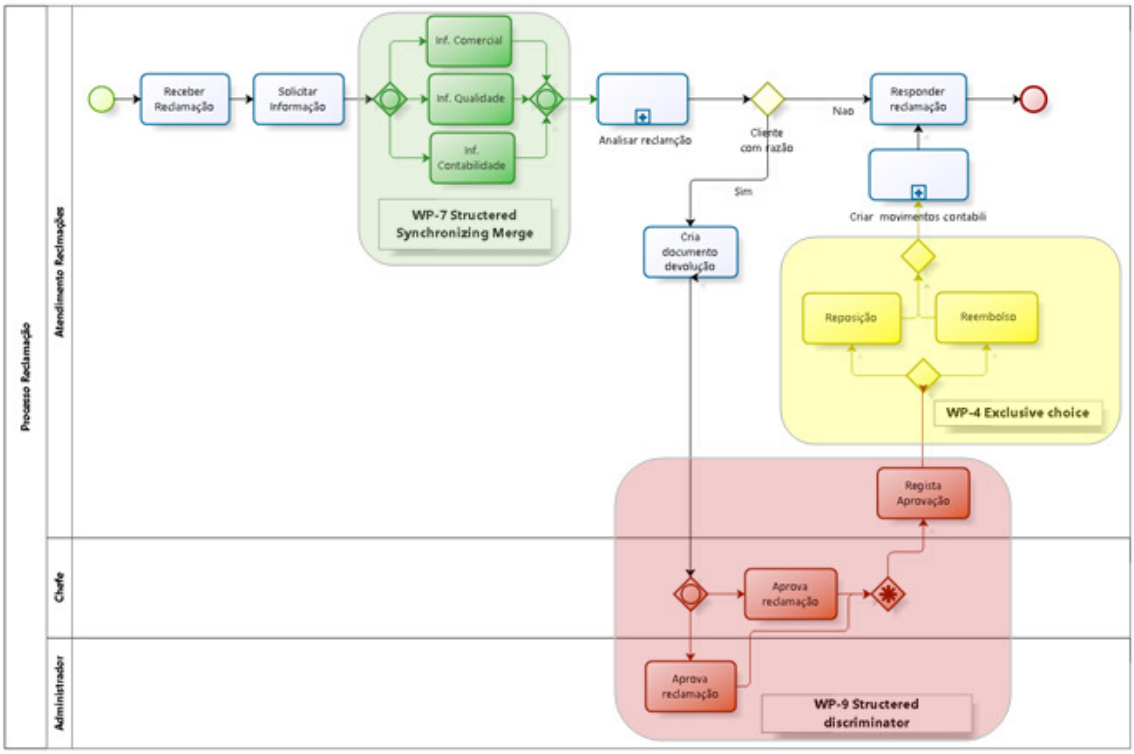


Figura 3 - Processo "Reclamação de Encomendas de Clientes"

Confirmando a pesquisa efetuada por Aalst et al. [2003a], que defendem que na modelação de qualquer processo de negócio surgem recorrentemente padrões *workflow*, o processo "Reclamação de Encomendas de Clientes" não é exceção. Conforme se pode verificar na figura anterior (Figura 3), foi possível identificar imediatamente três padrões:

- **Padrão *Workflow* 4: (*WP-4 – Exclusive choice*)** – Assinalado a amarelo.

  Descrição: Consiste na seleção de uma linha de controlo num conjunto de múltiplas linhas, através de uma decisão ou informação do controlo de *workflow* em determinado ponto do processo.

  Caso/exemplo: Após o registo da aprovação da reclamação segue-se a decisão do cliente: se deseja a reposição do produto ou o reembolso do mesmo. Apenas um dos caminhos é selecionado.

- **Padrão *Workflow* 7: (*WP-7 – Structured Synchronizing Merge*)** – Assinalado a verde.

  Descrição: Um ponto no processo de negócio onde dois ou mais subprocessos ou atividades convergem para uma única linha de controlo. Nesta fusão pressupõe-se que se mais que uma atividade foi executada em paralelo então é necessária sincronização na sua fusão, caso contrário, não necessita de sincronização.

  Caso/exemplo: Para analisar a responsabilidade da reclamação, é necessário recolher várias informações em departamentos diferentes, podendo ocorrer paralelamente. A atividade seguinte apenas se iniciará quando toda a informação solicitada, dos vários departamentos, se encontrar disponível.

- **Padrão *Workflow* 9: (*WP-9 – Structured Discriminator*)** – Assinalado a vermelho.

  Descrição: Ponto do processo que aguarda a conclusão de uma das atividades convergentes antes de iniciar a atividade subsequente. A partir desse momento, as outras atividades serão concluídas, mas a atividade subsequente não será realizada novamente. Isto pode ocorrer quando temos atividades que ocorrem em paralelo e permitem despoletar a atividade subsequente, mas que porém podem demorar tempos diferentes. Dessa forma, a primeira atividade que concluir continuará o fluxo de controlo ativando a atividade subsequente. Quando as outras atividades concluírem, os seus resultados serão descartados.

  Caso/exemplo: A aprovação está a cargo de duas entidades, sendo apenas uma necessária para aprovar. Quando uma das entidades aprova a atividade seguinte fica imediatamente em condições de ser iniciada.

Com base nos padrões *workflow*, são desenvolvidos os atrás referidos Simbits e efetuada a sua documentação para utilização futura de acordo com as seguintes perspetivas:

- Problema;
- Categorias;
- Conceitos Chave;
- Hipótese;
- Abordagem técnica:
- Detalhes para a construção do modelo:
- Configuração sistema:

Uma vez criados Simbits para os múltiplos padrões *workflow* identificados por Aalst et al. [2003a], qualquer processo de negócio pode ser transformado num modelo de simulação na ferramenta Simio, por simples reconhecimento dos padrões *workflow* que o constituem, a que

correspondem outros tantos Simbits. Por exemplo, relativamente ao processo "Reclamação de Encomendas de Clientes" poderemos reutilizar três dos Simbits correspondentes aos padrões de *workflow* WP-4, WP-7 e WP-9.

Com base no trabalho desenvolvido por Aalst et al. [2003a] e com a criação dos Simbits será criado um repositório de componentes de simulação correspondentes aos referidos padrões de *workflow*. Deste modo, espera-se que qualquer modelo de simulação de processos, ao recorrer a estes componentes previamente desenvolvidos e testados, possam ver a sua implementação acelerada e parcialmente pré-validada.

8. Conclusão

Atualmente, as organizações estão sujeitas à grande turbulência dos mercados e pressões constantes de mudança, perante este cenário, têm necessidade de melhorar e reformular constantemente seus negócios. O esforço para a melhoria contínua é uma das principais prioridades para as organizações, de forma a garantir a sua sobrevivência e sustentabilidade nos mercados reais.

BPM ajuda as organizações na resposta rápida e adequada às pressões em torno delas, defendendo uma estrutura organizacional centrada no conceito de processo. Contudo, os processos de negócio são sistemas complexos que envolvem pessoas, atividades e tecnologia, e carregam um alto grau de incerteza sobre os resultados obtidos em iniciativas de melhoria contínua. Neste contexto a simulação computorizada pode desempenhar um papel importante para o sucesso das iniciativas de mudança organizacional.

A simulação de processos pode ser considerada uma ferramenta de gestão da mudança, uma vez que permite tornar visíveis as razões da referida mudança e a obtenção de explicações para o processo de decisão, sempre com o intuito da melhoria continua. Mas a simulação de processos de negócio apresenta algumas limitações que comprometem a sua relevância como ferramenta para a tomada de decisões de gestão. Normalmente, é uma atividade bastante morosa e sujeita a falhas, pois implica a criação dos modelos de simulação de raiz. Por esta razão, é importante a criação de um repositório de componentes de simulação baseados em pequenos modelos - Simbits (construídos com a ferramenta de simulação Simio), que retratem um número considerado de situações padrão (padrões de *workflow*). Desse modo consegue-se rapidamente montar modelos de simulação de processos, parcialmente validados e prontos a servir os propósitos da simulação.

9. Referências

Aalst W.M.P. van der, Hofstede, A.H.M. ter, Kiepuszewski, B., e Barros, A.P., 14, 01, (2003a) "Workflow Patterns". *Distributed and Parallel Databases*, 5-51.

Aalst W.M.P. van der, Hofstede, A.H.M. ter, & Weske, M. Business Process Management: A Survey. Lecture Notes in Computer Science, pp. i – xiii. 2003b.

Ball, P., *Introduction to Discrete Event Simulation*, 2nd DYCOMANS Workshop on Management and Control: Tools in Action, in Algarve, Portugal. 15th-17th May 1996.

Barnett, M. W. *Modeling & Simulation in Business Process Management*. Gensym Corporation, 2003.

Cichocki, A. *Workflow and Process Automation – Concepts and Technology*. Kluwer Academic Publishers, 1998.

Garrat, Andy et. al. *Human-Centric Business Process Management with WebSphere Process Server V6. 2007*. Disponível em: https://obahy.files.wordpress.com/2013/04/human-centric-business-process-management-with-websphere-process-server_sg247477.pdf.

Gregoriades, A., Sutcliffe, A. *A socio-technical approach to business process simulation*. Decision Support Systems, v. 45, n. 4, p. 1017-1030. 2008.

Harmon, P. *Business Process Change: a manager´s guide to improving, redesigning, and automating processes*. EUA: Morgan Kaufmann Publishers, 2003.

Jeston, J. and Nelis, J. Business Process Management : Practical Guidelines to Successful Implementations. Elsevier/ButterworthHeinemann, Amsterdam. 2006.

Law, A., McComas, M., Secrets of Successful Simulation Studies, Preceedings of 1991 Winter Simulation Conference, 1991.

Pereira, J. L., *Sistemas de Informação para o Novo Paradigma Organizacional: O Contributo dos Sistemas de Informação Cooperativos*, Tese de Doutoramento, Universidade do Minho. 2004.

Riehle, D. and Zullighoven H. *Understanding and Using Patterns in Software Development*. Theory and Practice of Object Systems. 1996.

Russell, N.C., Hofstede, A.H.M. ter, Aalst W.M.P. van der & Mulyar N.A. *Workflow control-flow patterns*: a revised view. 2006.

Schiefer, J., Roth, H., Suntinger, M., Schatten, A., Simulating Business Process Scenarios for Event-Based Systems, Preceedings of 15th European Conference on Information Systems (ECIS'07), St. Gallen, Suiça, 2007.

Shannon, R. E., “Introduction to the Art and Science of Simulation”, in Proceedings of 1998 Winter Simulation Conference, D. J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, eds. 1998.

SIMIO. Case Study: Simio Models Movement of Armed Forces Between Germany and Afghanistan. Http://www.simio.com/case-studies/Konekta/. 2012.

Thom, L. H., Iochpe, C., Amaral, V. L. do, Viero, D. M. Toward block activity patterns for reuse in workflow design. 2006.

O impacto da utilização de RFID: Um estudo de caso numa empresa de manutenção e engenharia

David Ribeiro [1], Winnie Picoto [2] e Paulo Almeida Gonçalves[3]

1) ISEG Lisbon School of Economics and Management, Universidade de Lisboa, Portugal
davidtribeiro@gmail.com

2) ISEG Lisbon School of Economics and Management, Universidade de Lisboa, Portugal
w.picoto@iseg.ulisboa.pt

3) ISEG Lisbon School of Economics and Management, Universidade de Lisboa, Portugal
pagoncalves@iseg.ulisboa.pt

Resumo

A manutenção de aeronaves representa um custo elevado, pelos processos envolvidos e também pelas perdas associadas ao tempo fora de serviço. A identificação por rádio frequência (RFID) tem despertado o interesse deste setor pela sua capacidade de identificação de objetos de forma inequívoca, podendo assim contribuir para o aumento da eficiência destes processos complexos.

O presente estudo de caso pretende analisar o impacto da utilização da RFID numa empresa de manutenção de aeronaves, como a perceção da qualidade do sistema influencia a utilização do sistema e como a utilização do mesmo afeta os colaboradores da organização. Foram realizadas diversas entrevistas com utilizadores do sistema e com o gestor do projeto para obter informação das diferentes perspetivas.

A análise dos dados e as conclusões pretendem promover e melhorar a utilização da RFID de forma generalizada, bem como incentivar o desenvolvimento de frameworks para desenvolvimento e implementação de sistemas de RFID.

Palavras chave: Radio Frequency Identification (RFID); Setor aeronáutico; Manutenção, Processos complexos.

1. Introdução

Na indústria aeronáutica, a gestão dos equipamentos e produtos com vidas úteis longas e configurações complexas é um desafio. As empresas de *Maintenance Repair and Overhaul* (MRO - Manutenção, Reparação e Revisão Geral em português) neste setor precisam de garantir a performance contínua dos seus produtos e por isso, o conceito de gestão do ciclo de vida do produto é essencial no setor, e os objetivos da sua aplicação passam por reduzir o tempo de manutenção de uma aeronave, geralmente referido como TAT (*Turnaround Time*), reduzir os custos com manutenção e aumentar o tempo entre operações de manutenção [Lee et al. 2008].

Os custos de MRO correspondem a 12% do total dos custos operacionais de um avião [Lampe et al. 2004] e os custos por hora de ter um avião parado para manutenção não planeada rondam os 23.000 dólares americanos [Brown 2003]. É por isso visível a importância de melhorar a gestão e a qualidade destas atividades.

Na produção e manutenção de aeronaves, a RFID tem ganho reconhecimento pela sua capacidade de rastreio de objetos de forma inequívoca, em tempo real e com um elevado nível de precisão [Harun et al. 2008; Ferrer et al. 2011].

O presente estudo pretende responder a duas questões de investigação, nomeadamente: Quais os impactos organizacionais do uso de RFID?", "Como é que a utilização da RFID afeta os colaboradores da empresa?"; e "De que forma a qualidade do sistema de RFID influencia o uso do mesmo?". Para responder a estas questões, foi realizado um estudo de caso. Para tal, foi escolhida uma empresa aeronáutica de Manutenção e Engenharia (M&E) como unidade de análise, cujo nome não será referido por razões de confidencialidade. Esta empresa incorporou o uso da RFID na sua oficina para identificar componentes de motores de aviões. A principal motivação para este tema prende-se com a necessidade de aprofundar o conhecimento em relação a esta tecnologia em termos da qualidade dos sistemas RFID e do impacto destes nos utilizadores. Este estudo pretende contribuir para que as organizações compreendam melhor os benefícios da implementação de sistemas RFID na sua atividade.

Este artigo está organizado em cinco seções, sendo esta a primeira. Nas seções seguintes apresentam-se respetivamente a revisão da literatura relevante para o estudo, a metodologia, a caracterização do caso, a apresentação e discussão dos resultados obtidos e por último são apresentadas as conclusões e limitações do estudo.

2. Revisão da literatura

A identificação por rádio frequência (*Radio Frequency Identification* ou RFID) é uma tecnologia de identificação automática que utiliza ondas de rádio para identificar um animal, uma pessoa ou um objeto de forma inequívoca [Wyld 2006].

A utilização da RFID pode ter três tipos de impactos numa organização – automatização, informatização e transformação [Tellkamp 2006]. A automatização do processo de recolha e transmissão de dados pode eliminar a necessidade de ter um trabalhador numa zona de entrada de mercadoria para identificar os produtos rececionados [Gille & Strüker 2008]. Sendo a poupança associada à utilização da RFID, dependente da frequência da atividade de recolha de dados [Subirana et al. 2003]. A automatização é mais evidente em empresas que apresentem muitas atividades de recolha de dados, como é o caso do retalho [Gille & Strüker 2008]. Por

exemplo, a Gillette eliminou os processos de leituras e contagens manuais através do recurso a RFID e conseguiu um aumento da eficiência operacional de 400% e a redução do número de erros humanos associados [Roh et al. 2009].

A informatização refere-se à capacidade de recolher mais informação e com maior qualidade em termos de dados precisos, objetivos, atempados e completos [Tellkamp 2006]. As decisões de coordenação de operações entre parceiros podem assim ser melhoradas através de, por exemplo, comparações automáticas em tempo real de quantidades encomendadas e recebidas. Relativamente à transformação ou reengenharia e novos processos de negócio, os novos processos de *picking* num centro de distribuição podem resultar na redução de inventário, libertando fundos de capital [Gille & Strüker, 2008].

Além dos tipos de impacto da RFID, também é possível identificar diferentes níveis de integração. Wamba & Chatfield [2009] identificaram três níveis de integração da RFID nas organizações. No nível mais baixo, *Slap and Ship*, verifica-se uma baixa integração tecnológica e uma baixa transformação organizacional. No segundo nível – projeto intraorganizacional, já existe integração com os SI (Sistemas de Informação) da empresa e é considerada alguma reestruturação organizacional. No último nível – projeto interorganizacionais considera-se a integração da tecnologia com os SI dos parceiros e também a uniformização de processos de negócio e interfaces entre as organizações. O Quadro I apresenta em resumo, as características de cada nível de integração.

Âmbito Projeto	Integração tecnológica	Integração/Transformação Organizacional	Benefício	Custo
Nível 1: *Slap and ship*	Limitada a etiquetas	Ligeira transformação organizacional	Controlo do inventário interno	Compra de etiquetas, impressoras e *staff*
Nível 2: Intra-organizacional	Integração com os SI da organização (ERP)	Restruturação do armazém	Inventário em tempo real; Redução custos operacionais; Aumento da integridade de dados	Infraestrutura de RFID e integração com os SI; Formação
Nível 3: Inter-organizacional	Integração externa com os SI dos parceiros	Uniformização dos processos de negócio e interfaces entre empresas e fornecedores	Dados em tempo real entre a cadeia de fornecimento; Melhorias na previsão e planeamento da procura	Uniformização de processos e interfaces; Confiança entre empresas

Quadro I - Níveis de integração da RFID nas organizações

Fonte: Adaptado de Wamba & Chatfield [2009]

À medida que a RFID continua a amadurecer, o conhecimento técnico e teórico são aprofundados e o custo do *hardware* e dos sistemas diminui, o que leva ao aumento do interesse em aplicações de RFID e à respetiva difusão pelas mais variadas indústrias [Castro & Wamba 2007].

2.1 Benefícios da RFID

A RFID alarga o alcance dos SI ao incluir objetos sem capacidade inerente de processamento de informação, como produtos num hipermercado ou peças numa oficina, o que pode representar uma oportunidade estratégica para as empresas [Asif & Mandviwalla 2005]. Roh et al. [2009] decompõem os benefícios da RFID em diretos e indiretos, sendo os diretos operacionais por natureza (reduções de custos e de tempos de execução) e os indiretos intangíveis como, por exemplo, a criação de novos processos de negócio.

Para Zelbst et al. [2010], a utilização da RFID melhora a capacidade de uma organização responder às necessidades dos clientes. Os autores referem o exemplo de organizações produtivas que adotaram RFID e que se tornaram mais ágeis, concluindo que a utilização de RFID influencia diretamente a agilidade e performance operacional e logística de uma organização. Chappell et al. [2003] consideram que a utilização da RFID na produção permite a realização de benefícios como o aumento das vendas até um por cento derivado do aumento da qualidade e serviço ao cliente, a diminuição de capital fixo de um a cinco por cento, a redução do capital operacional de dois a oito por cento, a redução dos *stocks* através de menores ciclos e maior visibilidade e por fim, a redução do custo dos produtos vendidos de um a cinco por cento pela eficiente utilização dos equipamentos.

2.2. Desafios/barreiras à adoção

Um dos desafios à adoção generalizada da RFID é a falta de *standards* e a divergência na alocação dos espectros de ondas de rádio [Wu et al. 2006]. Esta ausência de *standards*, contribui para o surgimento de problemas tecnológicos como interferências de outros aparelhos (telefones, PDA, etc.) pela utilização da mesma frequência [Wu et al. 2006; McFarlane & Sheffi 2003], elevados custos de produção, dificuldades na personalização e integração de sistemas (intra e interorganizacionais) associados à falta de uma arquitetura bem desenvolvida e universalmente aceite [Wu et al. 2006] e um grande volume de dados e a respetiva capacidade de processamento e armazenamento dos mesmos [McFarlane & Sheffi 2003].

Outro dos entraves à adoção generalizada da RFID prende-se com a dificuldade em obter métricas precisas relativas ao impacto do efeito da RFID na organização, levando a questionar se esse impacto se deve ao efeito da RFID ou à melhoria dos processos internos [Lee e Ozer 2007]. Segundo Lee & Lee [2010], apesar de existir uma grande procura de métodos para avaliar estes projetos, estes ainda não foram devidamente desenvolvidos e métricas habituais como ROI, NPV (*Net Present Value*) e *Payback Period* são limitados, à partida, pela difícil quantificação dos benefícios realizados com a utilização da RFID. Estes autores propõem um método de avaliação de investimento em RFID com três vertentes: eficiência das compras,

eficiência JIT (*Just in time*) e eficiência operacional, derivando níveis ótimos de investimento para cada vertente.

2.3. A RFID no setor aeronáutico e na manutenção

De acordo com os testes efetuados nos EUA, a RFID apresenta resultados muito superiores aos códigos de barras na gestão das bagagens de passageiros em aeroportos [Chang et al. 2006]. De facto, a Administração Federal de Aviação Americana aprovou formalmente a utilização de etiquetas passivas de RFID em peças de aviões [Chang et al. 2006; Ngai et al. 2012]. E posteriormente, a Boeing e a Airbus anunciaram planos para desenvolver uma especificação de RFID para ser usada por ambas, com o intuito de padronizar e promover a utilização da RFID na indústria, nomeadamente em peças de aviões comerciais [Chang et al. 2006].

A eficiência dos processos de MRO é diretamente afetada pela gestão inadequada do inventário de peças, dos erros humanos e dos procedimentos de verificação [Ngai et al. 2012], contribuindo assim a RFID para obtenção de benefícios como a diminuição de atrasos, a diminuição de erros humanos, a automatização da documentação e o uso eficiente dos recursos disponíveis [Lampe et al. 2004]. Jimenez et al. [2011] resumem os benefícios que a RFID pode trazer aos processos de manutenção em cinco pontos: (1) Simplificação do processo de configuração de produtos; (2) controlo do inventário; (3) melhor planeamento e antecipação das tarefas de manutenção, traduzindo-se em menores tempos de intervenção e maior disponibilidade das aeronaves; (4) melhorias nas tarefas de manutenção (procura de peças, recolha de dados automática, preenchimento de documentos, etc.); (5) melhorias nos processos de verificação e controlo.

2.4. Modelos e escalas de sucesso, qualidade e usabilidade de sistemas

A medição do sucesso ou da eficácia de um SI é crucial para avaliar os investimentos e as ações da gestão de SI [DeLone & McLean 2003]. Assim, com base no modelo de qualidade da norma ISO/IEC 9126-1, Kang et al. [2013] propuseram um conjunto de seis critérios para a qualidade de um sistema de RFID: funcionalidade, fiabilidade, usabilidade, eficiência, manutenção e conceito de negócio. Os critérios estão subdivididos em subcritérios: a funcionalidade inclui subcritérios ligados à RFID como alcance de leitura e precisão de leitura, a eficiência considera o subcritério capacidade de dados, em termos de armazenamento ou tráfego [Kang et al. 2013].

Relativamente ao modelo proposto por DeLone & McLean [2003], Kang et al. [2013] substituíram as dimensões de qualidade (do sistema, da informação e de serviço) pelos atributos de qualidade de um sistema de RFID, mantiveram a variável satisfação do utilizador, adicionaram a satisfação do programador que substituiu o uso e a intenção de uso, e por fim os

benefícios do sistema em vez de benefícios líquidos. Os atributos de qualidade influenciam a satisfação do utilizador e do programador, até ao ponto em que atingem um nível que permite realizar benefícios do sistema [Kang et al. 2013].

No presente artigo, para avaliar a usabilidade do sistema foi adotada a *Systems Usability Scale* (SUS). Trata-se de um método validado e padronizado que pode ser usado para medir a usabilidade de praticamente qualquer interface de utilizador [Sauro 2011] e já foi amplamente utilizado para testar a usabilidade de *hardware*, *software*, *sites*, telemóveis e até das páginas amarelas [Sauro 2011; Brooke 2013].

3. Metodologia de investigação

Para responder às questões de investigação, este artigo aplica o método de estudo de caso único, porque para além de se enquadrar nos objetivos da investigação, privilegia o contacto direto com o fenómeno/caso em estudo.

Esse método é caracterizado por estudar fenómenos contemporâneos, num contexto real e sem influência ou manipulação, sendo adequado para responder à questões de investigação do tipo "porquê?" e "como?" [Benbasat et al. 1987; Yin 2014]. Esta investigação pode ser considerada um estudo de caso descritivo e exploratório [Yin 2014] visto que se pretende descrever um fenómeno no seu contexto e se trata de um problema pouco conhecido – o impacto que a introdução de RFID teve na empresa e nos seus colaboradores, bem como a avaliação da perceção dos colaboradores acerca da qualidade do sistema.

Escolhemos uma empresa que tivesse adotado RFID nos seus processos, a M&E, uma empresa conceituada de um grupo aeronáutico que adotou RFID para a identificação de peças de motores que dão entrada na sua oficina para serem efetuados trabalhos de MRO.

A recolha de dados foi realizada através de nove entrevistas com colaboradores da organização e de uma organização parceira, escolhidos com base no seu envolvimento no desenho e implementação do sistema ou por serem os atuais utilizadores da aplicação Mobile (*nome fictício que corresponde à aplicação móvel de reparações).* Foram assim inquiridos colaboradores das equipas de Técnico de Preparação, Planeamento e Compras, de Inspeção e o *Project Manager* de uma empresa externa responsável pela implementação, empresa gestora dos SI do grupo. Todas as entrevistas foram gravadas e posteriormente transcritas.

As entrevistas foram semiestruturadas com recurso a guião e a elaboração deste teve por base o modelo de Kang et al. [2013], contemplando os diferentes critérios de qualidade de um sistema de RFID, bem como a SUS, para melhor analisar a Usabilidade do sistema e como a utilização

do mesmo afeta os colaboradores [Sauro 2011; Brooke 2013]. Ao basear os guiões em modelos e escalas amplamente utilizados, é pretendido que as conclusões desta investigação sejam apoiadas em dados consistentes e relevantes para o tema.

Adicionalmente, esta investigação envolve uma triangulação de dados ao contemplar múltiplas fontes de informação [Yin, 2014], nomeadamente diferentes tipos de *stakeholders* do projeto de RFID bem como documentação do mesmo. Esta técnica foi adotada com o intuito de aumentar a validade das conclusões do estudo, confrontando os diferentes dados recolhidos.

4. O Estudo de Caso

A M&E é a organização de manutenção de um grupo aeronáutico e o seu *core business* divide-se em três áreas de manutenção: de aeronaves, de motores e de componentes. Este estudo incide na área de manutenção de motores, presente em quatro edifícios separados. Um motor contém entre quatro a seis módulos, dezassete minimódulos (em média) e aproximadamente cinco mil peças. Com aproximadamente cem motores restaurados/reparados por ano, o volume de peças que é preciso gerir em qualquer momento é significativo.

Os processos da manutenção de motores envolvem aproximadamente quinhentas peças de um motor, sendo geridos por uma combinação de folhas de serviço e um *Maintenance Management Information System* que por razões de confidencialidade, será denominado de MMIS. Este sistema foi desenvolvido internamente e suporta as atividades da oficina através da gestão de componentes e ordens de serviço nas diversas fases de manutenção. O sistema encontra-se na segunda versão e está interligado a outros sistemas como o Financeiro, o de Inventário e Compras e o de Recursos Humanos.

Neste projeto participaram ainda, uma empresa de sistemas de informação no *design*, suporte de TI e gestão do projeto, uma consultora na análise de negócio e definição do caso e na arquitetura da solução e a empresa responsável pelo desenvolvimento da aplicação Mobile e pelo apoio operacional. Este sistema é composto por etiquetas nas folhas de serviço e nas prateleiras (*racks* fixas e móveis), PDA e impressoras de etiquetas, *middleware* e sistemas *backend* como o MMIS. Segundo o PM: "*O objetivo não visou obter ganhos de eficiência por via da automatização de tarefas, mas sim a redução de custos por facilitar a execução de tarefas que eram cem por cento manuais e por isso time consuming* (…) ".

A aplicação Mobile carrega, através de um *script,* as listas de minimódulos e outros dados do inventário armazenados e disponíveis no MMIS nos PDA, que para isso têm de ser conectados diariamente a um computador com ligação à rede interna. Quando os técnicos usam os PDA para efetuar leituras, os dados são processados pela aplicação Mobile e apresentados em tempo

real no visor, permitindo ao utilizador acesso a informação sobre as etiquetas lidas. Quando os PDA são ligados à rede trocam dados com o middleware que por sua vez comunica com o MMIS. O PDA disponibiliza informação, como por exemplo a última localização de uma peça, a comparação de leituras de inventário, sendo os relatórios mais complexos gerados através da aplicação MMIS. Este foi um projeto intraorganizacional pois envolveu a integração da RFID com os SI da empresa [Wamba & Chatfield 2009].

O projeto de implementação durou dez meses, embora desde a conceção até à implementação tenham decorrido três anos. De acordo com informação recolhida nas entrevistas e a documentação do projeto, os principais benefícios esperados do sistema são o aumento da eficiência dos colaboradores (quer nos processos de reparação, quer nos de logística) reduzindo custos com os mesmos, a redução do TAT, a redução de multas por atrasos (com o menor número de peças perdidas, o risco de paragens de trabalhos/produção diminui), a redução do risco de reposição de peças perdidas, incorrendo mais custos, uma melhor experiência dos utilizadores e a possibilidade de potenciar o desenvolvimento de projetos de TI futuros dentro do grupo aeronáutico com recurso a RFID, por se tratar de um projeto inovador.

A fase de implementação e testes teve alguns desafios, nomeadamente em termos do ambiente da oficina. Tendo em conta a grande quantidade de metais (das peças, ferramentas, etc.) existente na oficina de motores e a possibilidade de interferirem com as ondas de rádio, houve necessidade de um processo de análise cuidado para prevenir problemas nas leituras. Por isso mesmo, a conjugação de *hardware*, *software* e etiquetas teve de ser alvo de testes rigorosos para garantir a eficiente utilização do sistema. Para além disso, revelou-se extremamente importante a formação dada aos utilizadores, não só na utilização dos PDA como também em termos da reestruturação de alguns processos, nomeadamente na arrumação e disposição das peças no recinto da oficina. Segundo o PM existiram "*ações de formação para um grupo alargado de utilizadores, onde se incluíram todas as chefias de produção, as chefias de segundo nível e cerca de 30 utilizadores. As ações de formação tiveram uma componente teórica em sala e uma componente prática ao nível da oficina, utilizando dados reais de produção (…)* ". O sistema entrou em produção em Janeiro de 2012.

5. Apresentação e discussão dos resultados

Em termos do processo de leitura através de PDA, todos os entrevistados consideraram uma experiência satisfatória, apresentando uma boa capacidade de leitura. No entanto, há dois fatores que afetam a leitura de etiquetas, nomeadamente a sobreposição de peças ou etiquetas, "*quando estamos a fazer picagem ou nas prateleiras, principalmente em caixas. É preciso ter algum*

cuidado mas com a experiência torna-se fácil". O outro fator é a infraestrutura da oficina que pode ser um fator negativo para o processo de leitura porque " *(…) não está adequada ao sistema, (…)* " tendo um dos colaboradores explicado que obter uma taxa de cem por cento é difícil, especialmente num ambiente oficinal onde predominam peças de metal. Os utilizadores confiam nas leituras do PDA, no entanto, é muitas vezes feita uma confirmação visual: "*faço a leitura com a máquina mas também faço a confirmação visual.*". Isto porque: "*ainda confiamos mais em nós, no nosso know-how (…) até porque pode haver uma troca de etiquetas com peças. Houve um caso com 3 peças trocadas que causou alguns problemas.*" Este exemplo evidencia que apesar de a tecnologia reduzir a probabilidade de erro humano, é necessário formar os utilizadores para promover uma correta utilização da mesma.

De acordo com as entrevistas efetuadas, o sistema foi avaliado positivamente em termos da existência e recuperação de *bugs* (fiabilidade). O único problema reportado foi o facto de ocasionalmente os cartões usados para fazer login no PDA não serem reconhecidos à primeira. Como referido anteriormente, para analisar a capacidade do sistema satisfazer as necessidades dos utilizadores de forma simples e eficaz (usabilidade) foi utilizada a SUS. Um dos entrevistados afirmou que para " *(…) quem está ligado às novas tecnologias isto é muito fácil de utilizar*" e outro referiu que " *(…) para mim é facílimo; aquilo é quase um Windows em ponto pequeno*."

A eficiência foi avaliada através de três critérios, a velocidade de processamento/tempo de resposta, a velocidade de arranque do PDA e a velocidade de sincronização do PDA com o MMIS. Em relação aos dois primeiros critérios, todos os entrevistados avaliaram-nos positivamente. Apenas em relação à velocidade de sincronização a avaliação foi menos positiva.

A manutenção da aplicação Mobile é limitada pela dependência de um fornecedor externo. Quando foi requisitado um *upgrade* funcional " *(…) para acomodar os pedidos efetuados pelos utilizadores durante as ações de formação, (…) o processo revelou-se demorado, muito formal e burocrático, (…) e o resultado final não correspondeu a cem por cento ao pretendido. Por não termos acesso ao código fonte (…) estamos completamente dependentes (…) "*. Em contrapartida, o recurso a um fornecedor externo " *(…) com um profundo know-how e experiência nesta tecnologia aplicada ao mesmo setor de atividade, acabou por alavancar a criação de uma equipa com competências nesta tecnologia."* Em termos de projetos futuros o gestor de projeto disse: "*Relativamente a futuras soluções internas nesta tecnologia, sem dúvida seremos nós a desenvolvê-las, a menos que estas soluções sejam COTS (Commercial Off-The-Shelf) disponibilizadas no mercado para satisfazer requisitos específicos do negócio da aviação, suportadas em standards globais (…)."*

Através da análise da qualidade do sistema foi possível detetar pontos a melhorar e erros cometidos no processo de desenvolvimento e implementação, nomeadamente a falta de arquiteturas disponíveis para sistemas de RFID (tal como referido por McFarlane & Sheffi [2003] e Wu et al. [2006]), que nem o profundo *know-how* dos parceiros conseguiu mitigar.

A qualidade do sistema, ao influenciar a utilização do mesmo, influencia também a capacidade de obtenção de benefícios [DeLone & McLean 2003; Kang et al. 2013]. Apesar das funcionalidades existentes serem adequadas, os aspetos referidos prejudicam a perceção da qualidade do sistema, permanecendo a convicção de que o potencial do sistema ainda não foi totalmente atingido, o que constitui um impacto negativo na realização dos benefícios potenciais.

No processo de desenvolvimento, não foram envolvidas as equipas da oficina de motores e por isso um dos entrevistados referiu que *" (…) devia haver da parte de quem concebe o programa, uma noção mais profunda do que realmente se passa no terreno, porque isto é uma ferramenta de trabalho (…) Não nos sentimos muito integrados*". Este fato pode ser prejudicial para projetos de TI por não adequar a solução aos utilizadores como foi indicado pelo PM: "*Os requisitos de negócio que estiveram na origem da solução [Mobile] não foram objeto de uma participação ativa em termos de discussão e conceção por parte das áreas de produção das oficinas de motores, principais utilizadores da aplicação".* O sistema acabou por não corresponder por completo às expectativas de todos os colaboradores da oficina: " (…) *a imagem que tinha, ao princípio, deste sistema parecia que ia ser muito mais útil. Com o tempo é que se percebeu que de facto não correspondeu às expectativas (…)* ". Foi possível detetar um consenso de que o sistema trouxe vantagens, principalmente às tarefas de reparação, logística e preparação. Um dos colaboradores refere que " *(…) como lidamos com stocks e valores muito altos, pode haver falhas e não sabermos onde estão peças. Através da RFID conseguimos encontrá-las e isso facilita bastante (…)* ". Com algumas modificações e desenvolvimentos, o sistema poderia ter um maior impacto a nível da produtividade e responsabilização dos colaboradores bem como na sua satisfação.

Em termos dos benefícios para os colaboradores, o sistema facilitou o processo de procura de peças, que nalguns casos era bastante frustrante. Este benefício é de natureza operacional e seguindo Roh et al. [2009] pode ser classificado como um benefício direto da implementação de RFID. Em termos das categorias de Tellkamp [2006], pode ser considerada uma transformação dos processos internos. Desta forma, principalmente para os técnicos, o sistema teve um impacto bastante positivo no seu dia-a-dia, evitando erros humanos e eliminando tempos

mortos. É então possível afirmar que em termos do processo de procura de peças, o objetivo de melhorar a experiência dos utilizadores foi, sem dúvida, alcançado. Seria no entanto muito vantajoso para a satisfação dos colaboradores e consequentemente para o desempenho da empresa, que este benefício fosse estendido a outros processos, através de novas funcionalidades. Questões mais humanas como o acompanhamento no processo de transição, " *(…) o cuidado de explicar às pessoas o funcionamento (…) "* e o envolvimento dos superiores " *(…) para passar a mensagem do porquê da mudança e de como utilizar o sistema,*" também tiveram um impacto positivo na satisfação dos utilizadores.

Para analisar o impacto organizacional do projeto Mobile, as respostas dos entrevistados foram comparadas com os benefícios esperados, definidos na fase do *proof of concept* do projeto.

Vários colaboradores da empresa referem ganhos de eficiência nas suas tarefas, por exemplo: *"automatização de tarefas e aumento da eficiência das pessoas*" e *"facilitou imenso na picagem de material e na procura de peças.*" Estes benefícios vão ao encontro dos indicados por Lampe et al. [2004], Jimenez et al. [2011] e Legner & Thiesse [2006]. Com este projeto mitigaram-se as fraquezas dos processos de MRO reportadas por Ngai et al. [2012], entre elas a gestão inadequada do inventário de peças, dos erros humanos e dos procedimentos de verificação. Um colaborador da empresa há trinta e quatro anos, que conhece bem esta indústria referiu: *" (…) a manutenção tem a ver com o TAT e o tempo que é necessário para a realizar. Cada vez mais são esmagados os tempos em termos contratuais. Encontrar ferramentas para aliviar a perda de tempo, é sempre benéfico.*". Legner & Thiesse [2006] referem que a RFID acrescenta valor, simplesmente reduzindo os tempos de procura, indo ao encontro do que o PM referiu: "O *processo manual de procura de peças no circuito oficinal é extremamente moroso e ineficiente e a aplicação [Mobile] veio dar uma ajuda fundamental neste processo."* Este facto permitiu melhorias operacionais, nomeadamente a redução do número de peças perdidas e tempo despendido na sua procura. Estas melhorias permitiram benefícios intangíveis como a redução do risco de paragens dos trabalhos e eventuais custos adicionais para repor uma peça perdida. Todas estas melhorias permitiram uma redução do tempo de entrega e um melhor serviço ao cliente.

O projeto proporcionou também uma *" (…) situação de win-win em termos da parceria*" entre a empresa fornecedora e o cliente. Com a já referida criação do núcleo interno de competência em RFID na empresa implementadora do sistema é possível oferecer mais soluções às empresas do grupo aeronáutico. Nomeadamente a M&E ganhou *" (…) um fornecedor interno para as suas soluções de negócio envolvendo tecnologia RFID, sem necessidade de recorrer a fornecedores*

externos, que representam um custo acrescido sem necessariamente trazerem vantagens óbvias e duradouras para o negócio."

Pode-se afirmar que a utilização de RFID na M&E trouxe diversos benefícios diretos e indiretos. Os primeiros, em termos operacionais como a redução de custos e o aumento da eficiência. Os segundos, em termos do desenvolvimento de competências internas na área da RFID. Apesar das expectativas dos utilizadores não terem sido totalmente conseguidas, os benefícios esperados definidos na fase inicial do projeto foram alcançados, ainda que nem todos ao nível desejado.

6. Conclusões, limitações e estudos futuros

A avaliação global dos utilizadores em relação à qualidade do sistema foi positiva, persistindo a convicção de que o sistema não alcançou o seu potencial máximo e que não foi concebido totalmente de acordo com as necessidades de quem o iria utilizar. Isto prejudicou a adoção do sistema mas, apesar disso, a implementação do projeto Mobile veio afetar positivamente o dia-a-dia dos colaboradores da oficina da M&E, tendo-se verificado melhorias de eficiência no processo de procura de peças e tendo aumentado a satisfação dos utilizadores ao evitar paragens de produção, perdas de peças e redução do tempo despendido à procura das mesmas. A eficiência dos colaboradores aumentou (reduzindo os custos com os mesmos) e foi possível reduzir o TAT. O sucesso do projeto foi alcançado na medida em que todos os benefícios esperados foram atingidos, ainda que aquém do nível pretendido.

O projeto teve alguns desafios, nomeadamente em termos da falta de *standards* e melhores práticas disponíveis para adoção da RFID em empresas de manutenção. Esta realidade obriga a elevados custos de desenvolvimentos à medida, o que somado ao facto do fornecedor desta aplicação ser estrangeiro, causou alguns problemas na comunicação e adequação do sistema às necessidades da empresa e dos seus colaboradores. Foram também cometidos erros durante o desenvolvimento do sistema, mais concretamente o não envolvimento dos utilizadores do sistema na definição dos respetivos requisitos.

Como considerações finais, é pretendido que o presente estudo possa contribuir para empresas que pretendam implementar RFID, ao apresentar um caso real onde são descritos não só os benefícios obtidos como os erros cometidos e técnicas criativas para contrariar problemas do uso da tecnologia.

As principais limitações deste estudo prendem-se com o facto da gestão da oficina da M&E não se ter mostrado disponível para oferecer o seu contributo a este estudo e com o facto do estudo apenas contemplar uma empresa de um setor específico.

Tendo em conta a maturidade da tecnologia, seria importante a realização de estudos futuros em implementações da RFID que serão fulcrais para demonstrar como devem ser conduzidos os processos de análise e definição de um sistema de RFID, implementação e avaliação/manutenção posterior. Poderão ser desenvolvidas *frameworks* genéricas que irão não só fomentar a adoção desta tecnologia como reduzir custos com este processo, ao minimizar os custos de desenvolvimentos à medida.

Referências

Asif Z. e Mandviwalla M. (2005). Integrating the supply chain with RFID: a technical and business analysis. *Communication of AIS* 15 (24), 393–426.

Benbasat, I., Goldstein, D. e Mead. M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly* 11, 369-386.

Brooke, J. (2013). SUS: A Retrospective. *Journal of usability studies* 8 (2), 29-40.

Brown, P. (2003). Companies get creative in their Inventory Management Solution. *Aviation Now,* 15.

Castro L. e Wamba S. (2007). An Inside Look at RFID Technology. *Journal of Technology Management & Innovation*, 2 (1), 128-141.

Chang, Y., Oh, C., Whang, Y., Lee, J., Kwon, J., Kang, M. S., Park, J. e Ung, Y. (2006). Development of RFID enabled aircraft maintenance system, *2006 IEEE International Conference on Industrial Informatics,* Singapura, 16-18 Agosto de 2006.

Chappell, G., Ginsburg, L., Schmidt, P., Smith, J. e Tobolski, J. (2003). Auto-ID on the line: The value of Auto-ID technology in manufacturing. *Auto ID Center, ACN-AutoID-BC-005* 9.

DeLone, W. e McLean, E. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of management information systems* 19 (4), 9-30.

Ferrer, G., Heath, S. K. e Dew, N. (2011). An RFID application in large job shop remanufacturing operations. *International Journal of Production Economics*, 133 (2), 612-621.

Gille, D. e Strüker, J. (2008). *Into the Unknown: Measuring the Business Performance of RFID Applications* [Em linha]. Disponível em: http://is2.lse.ac.uk/asp/aspecis/20080218.pdf [Acesso em 02/02/2014].

Harun, K., Cheng, K. e Wibbelmann, M. (2008). RFID-enabled aerospace manufacturing: theoretical models, simulation and implementation issues. Em: *Industrial Engineering and Engineering Management*. Singapura, Dezembro 2008. IEEE, 1824-1829.

Jimenez, C., Dauzère-Pérèz, S., Feuillebois, C. e Pauly, E. (2011). Simulation of aircraft maintenance improvement using RFID systems. *International Conference on Industrial Engineering and Systems Management.*

Kang, A., Park, J., Barolli, L. e Jeong, H. (2013). A QoS Model for a RFID Enabled Application with Next-Generation Sensors for Manufacturing Systems. *International Journal of Distributed Sensor Networks* 2013.

Lampe, M., Strassner, M. e Fleisch, E. (2004). A ubiquitous computing environment for aircraft maintenance, *2004 ACM symposium on Applied computing*, ACM, Nicósia, 14-17 Março de 2004.

Lee, H. e Ozer, O. (2007). Unlocking the value of RFID. *Production and Operations Management* 16 (1), 40–64.

Lee, I. e Lee, B. (2010). An investment evaluation of supply chain RFID technologies: a normative modeling approach. *International Journal of Production Economics*, 125 (2), 313-323.

Lee, S., Ma, Y., Thimm, G. e Verstraeten, J. (2008). Product lifecycle management in aviation maintenance, repair and overhaul. *Computers in Industry*, 59 (2), 296-303.

Legner, C. e Thiesse, F. (2006). RFID-based maintenance at Frankfurt airport. *IEEE Pervasive Computing* 5 (1), 34-39.

McFarlane D. e Sheffi Y. (2003). The impact of automatic identification on supply chain operations. *International Journal of Logistics Management* 14 (1), 1–17.

Ngai, E., Chau, D., Poon, J., Chan, A., Chan, B. e Wu, W. (2012). Implementing an RFID-based manufacturing process management system: Lessons learned and success factors. *Journal of Engineering and Technology Management* 29 (1), 112-130.

Roh J., Kunnathur A. e Tarafdar M. (2009). Classification of RFID adoption: an expected benefits approach. *Information & Management* 46 (6), 357-363.

Sauro, J. (2011). *Measuring Usability with the System Usability Scale (SUS)*. [Em linha]. Disponível em: http://www.measuringu.com/sus.php [Acesso em 02/02/2014].

Subirana, B., Eckes, C., Herman, G., Sarma, S. e Barrett, M. (2003). *Measuring the Impact of Information Technology on Value and Productivity using a Process-Based Approach: The case for RFID Technologies.*

Tellkamp, C. (2006). The Impact of Auto-ID Technology on Process Performance – RFID in the FMCG Supply Chain, Dissertação [Em linha]. Doutoramento, Universidade de St. Gallen. Disponível em: http://cocoa.ethz.ch/downloads/2014/06/None_STG-AUTOID-WH003_01.pdf [Acesso em 02/02/2014].

Wamba, S. e Chatfield, A. (2009). A contingency model for creating value from RFID supply chain network projects in logistics and manufacturing environments. *European Journal of Information Systems* 18 (6), 615-636.

Wu, N., Nystrom, M., Lin, T. e Yu, H. (2006). Challenges to global RFID adoption. *Technovation* 26 (12), 1317-1323.

Wyld, David C. (2006). RFID 101: the next big thing for management. *Management Research News* 29 (4), 154-173.

Yin, R. (2014). *Case study research: Design and methods*, 5ª Ed. Thousand Oaks, CA, Sage Publications.

Zelbst, P., Green Jr, K. e Sower, V. (2010). Impact of RFID technology utilization on operational performance. *Management Research Review* 33 (10), 994-1004.

O Impacto do Sistema de Workflow no Licenciamento de Redes e Serviços de Radiocomunicações – Estudo de caso na ANACOM

Manuel Gomes[1], Luisa Domingues[2]

1) ISCTE-IUL, Portugal

maneljrgomes@gmail.com

2) ISCTE-IUL, Portugal
luisa.domingues@iscte.pt

Resumo

O objetivo deste estudo visa avaliar o impacto de um sistema de workflow no principal processo da ANACOM, i.e., no Licenciamento de Redes e Serviços de Radiocomunicações. Neste contexto o desafio colocou-se em determinar o impacto em termos de eficiência operacional, eficácia organizacional e satisfação dos recursos humanos.

De forma a atingir os objetivos propostos foram realizadas entrevistas semi-estruturadas aos responsáveis dos processos, bem como aos recursos humanos que operam diretamente com o sistema. A eficiência organizacional obtida foi aferida pela recolha de elementos quantitativos relativos a tempo de trabalho e a custos envolvidos, num cenário pré e pós implementação do sistema. Relativamente à avaliação da eficácia fez-se o confronto de medidas de eficiência versus objetivos estratégicos da organização. Para validação do impacto na satisfação dos recursos humanos e qualidade percebida, realizou-se um questionário interno anónimo e de resposta fechada que cobriu todo o universo.

A conclusão a que se chegou foi a de que a adoção do sistema de workflow teve um impacto generalizado na gestão documental da organização.

Palavras chave: gestão da mudança, impacto organizacional, caso de sucesso, workflow

1. Introdução

O sistema de workflow, designado na organização por my work space (MyWs), em 2008, veio substituir os sistemas de Classificação, Circulação e Arquivo de Documentos (CCAD) e o Registo de Entradas e Saídas (RES). O MyWs substituiu estes sistemas, designadamente ao nível procedimental, tanto no que diz respeito à gestão da correspondência - digitalização de documentos -, como no que diz respeito a base de dados de conhecimento sustentada por dimensões de classificação, nomeadamente pelo tipo de ações (pareceres, expedição, autorização, etc.). Todos os documentos digitais, existentes no CCAD, e o histórico de circulação desses documentos, registado no RES, foram migrados para o MyWs. Por exemplo, o MyWs passou a garantir o registo de circulação de documentos equivalente ao RES, mas com a diferença de que a um determinado documento (com identificação e classificação) RES estava inerente a criação

de um template desse documento que deixou de existir uma vez que a informação de circulação designadamente, no que diz respeito ao tipo de documento, à data em que é criado, ao assunto que trata e a quem o criou é registada no fluxo de trabalho, como metadados do documento "informação".

O MyWs, enquadrado como a aplicação substituto das aplicações CCAD e RES teve um impacto, na generalidade, na organização, designadamente por criar um novo conceito – conceito de grupo (tecnicamente 'pool') de recursos – que estabeleceu a comunicação electrónica entre as diversas Unidades Organizacionais (UO) e com o contexto tecnológico exterior à aplicação. Entende-se como exterior uma entidade externa ou outro sistema aplicacional. Foram assim disponibilizados grupos administrativos, um por cada UO para comunicação interna – são exemplo os serviços administrativos de cada UO ou a secretaria; e grupos funcionais, associados ou não à conta de correio electrónico, para comunicação com o exterior. No que respeita a processos transversais, cabe à UO responsável por esse processo a definição de novos procedimentos suportados pelo workflow. Por exemplo, caberá à divisão de licenciamentos de redes e serviços radioelétricos definir os procedimentos de circulação electrónica do documento "licença radioelétrica".

Evidencia-se a mais-valia que o MyWs trouxe em relação ao CCAD e ao RES, sobre a qual se enumeram alguns benefícios como:

1. a desmaterialização dos documentos;
2. a eliminação da necessidade de assinatura em documentos internos;
3. a centralização de um serviço de expediente e arquivo;
4. a gestão de ações efectuadas no documento;
5. a gestão de processos pendentes;
6. a gestão do estado dos documentos, com controlo de versões;
7. a eliminação da redundância dos arquivos locais;
8. a diminuição da lentidão em encontrar, aprovar e partilhar documentos;
9. a redução do tempo de tomada de decisão, por melhor estruturação da informação;
10. o aumento da segurança – os documentos não se perdem;
11. a redução dos gastos operacionais e o aumento da produtividade, por conjugação dos pontos anteriores;
12. a uniformização de conceitos, processos e procedimentos.

2. Metodologia

O presente trabalho é o estudo de um caso, cujo objeto é sobre a implementação do processo de Licenciamento de Redes e Serviços de Radiocomunicações e consequentemente aferir sobre o impacto organizacional com a adoção, por parte da organização, de um Workflow. A metodologia para a análise desse objeto de estudo enquadra-se numa abordagem de estudo de caso por ser especialmente adequada quando se procura compreender, explorar ou descrever acontecimentos e contextos complexos, nos quais estão simultaneamente envolvidos diversos factores, designadamente quando se procuram respostas para o "como?" e "porquê?" e/ou quando o objetivo é descrever ou analisar um fenómeno de forma a apreender a sua dinâmica e o seu processo (Yin, 1994).

O estudo de caso, como metodologia utilizada, é definido com base nas características do fenómeno em estudo e com base num conjunto de características associadas ao processo de recolha de dados e às estratégias de análise dos mesmos (Yin, 1994). Este método de pesquisa utiliza, geralmente, dados qualitativos que são recolhidos a partir de acontecimentos reais, com o objetivo de explicar, explorar ou descrever fenómenos enquadrados e contextualizados no seu próprio meio ambiente.

A escolha da abordagem metodológica depende do objetivo e, consequentemente, das questões que o investigador quer ver respondidas (Branski, et al.). Há três categorias básicas em que o objetivo de uma pesquisa se caracteriza: exploratória, explicativa e descritiva. A exploratória visa compreender um fenómeno pouco estudado ou aspetos específicos de uma teoria; a explicativa identifica os fatores que determinam ou contribuem para a ocorrência dos fenómenos, explicando as suas causas; e por último, a descritiva que descreve determinada população ou fenómeno.

Para se desenvolver uma pesquisa com utilização do método de estudo de caso é preciso cumprir com cinco etapas: delimitar, desenhar, preparar e recolher dados, analisar os casos entre os casos e concluir (Branski, et al.).

A começar pela etapa da delimitação da pesquisa, definiram-se, na presente pesquisa, os seus objetivos, que são determinar o impacto de um Workflow num processo de negócio em termos de eficiência operacional, eficácia organizacional e satisfação dos recursos humanos. Assim, trata-se de uma pesquisa explicativa e descritiva, adequada à utilização do método de estudo de caso na medida em que o fenómeno é contemporâneo e investigado no seu contexto real, visa entender o porquê de decisões tomadas, como foram implementadas e os resultados obtidos.

Definido o método, realizou-se uma revisão bibliográfica que permitiu estabelecer os fundamentos da pesquisa, identificando as teorias e conceitos relevantes para o trabalho. A revisão bibliográfica permitiu, também, a definição dos elementos que deveriam ser levantados na

pesquisa de campo para investigar as proposições relacionados com a identificação de aplicações de Tecnologias de Informação (TI) baseadas em workflow e utilizadas por organizações em processos de negócio e para analisar o seu impacto na eficiência, na eficácia (missão) e na satisfação dos utilizadores dessas TI.

A atividade seguinte foi a definição do caso. Foram identificados na organização os intervenientes internos ao funcionamento do processo (processo de Licenciamento de Redes e Serviços de Radiocomunicações). Foi estabelecido o primeiro contato para entrevista e selecionado o universo para o lançamento do questionário que visava a recolha de informação sobre a satisfação dos utilizadores.

Relativamente aos instrumentos para levantamento dos dados, utilizou-se a Internet, publicações sobre Workflows e processos de negócio, apresentações e relatórios internos da organização. Sobre a análise de dados, foram realizadas entrevistas semi-estruturadas, questionário interno anónimo e de resposta fechada que cobriu todo o universo e observação-direta.

Quanto à preparação e recolha de dados, os entrevistados foram contactados pessoalmente e foi-lhes explicado resumidamente os propósitos da pesquisa. Foram analisados casos de sucesso em organizações que utilizaram workflows nos seus processos de negócio, para além de procurar outras fontes com o objetivo de captar o máximo de informação possível sobre implementação de processos de negócio com a adoção de Workflow e aferir do impacto.

As entrevistas foram agendadas previamente no local de trabalho e sem tempo definido de duração. O roteiro de entrevista foi composto por perguntas abertas (conforme se mostra em anexo B) e com liberdade para os entrevistados de discorrerem sobre os aspectos que consideravam mais importantes. Foram consultados informantes priveligiados: um executivo da área de Tecnologias de Informação (TI), outro da área responsável pelo processo de negócio e outro da área de recursos humanos. No final da entrevista, o roteiro era consultado e as eventuais lacunas preenchidas. Os entrevistados concordaram com o pedido de consulta de informação sobre as suas áreas. Finalmente, todo o material recolhido foi organizado.

3. Estado da Arte

As mudanças organizacionais podem ser mais ou menos planeadas, antecipadas ou próativas, podem ser efetuadas de forma radical ou episódica, segundo a definição dos autores (Cunha, et al., 2007), pois de acordo com os mesmos, a mudança radical ou episódica, ao contrário da incremental ou contínua, acontece mediante um episódio de forte impacto na organização; isto é, mudanças organizacionais que tendem a ser pouco frequentes, descontínuas e intencionais e

ocorrem em períodos de divergência, quando a organização está a afastar-se das suas condições de equilíbrio.

Uma mudança tecnológica, numa organização, implica com os modos de realizar o trabalho e, apesar do objetivo dessa mudança, por vezes, não ser tornar o processo de negócio mais eficiente, torna-o implícito (Kotter, 1996), são os níveis mais baixos da organização que ajustam essa mudança, canalizando-a para os níveis hierárquicos mais altos para aprovação, pois são os níveis hierárquicos mais baixos que compreendem a tecnologia e têm o *skill* para propor mudanças.

Por outro lado, a compreensão da tecnologia aumenta, conforme nos demonstram os modelos de maturidade da gestão de Sistemas de Informação (SI), com a premissa de que as pessoas, organizações, áreas funcionais, processos, etc. evoluem através de um processo de desenvolvimento, crescimento ou maturação em direção a uma maturidade mais avançada, atravessando um determinado número de estádios distintos.

A teoria de estádios de Nolan (1979) demonstra os estádios até se atingir a maturidade: Iniciação, Contágio, Controlo, Integração, Gestão de Dados e Maturidade com um ponto de transicção entre os estádios Controlo e Integração (figura 1). O ponto de transicção define o fim da primeira curva em "S" e o início da segunda curva em "S" de crescimento explosivo da tecnologia usada, custos e aprendizagem organizacional.

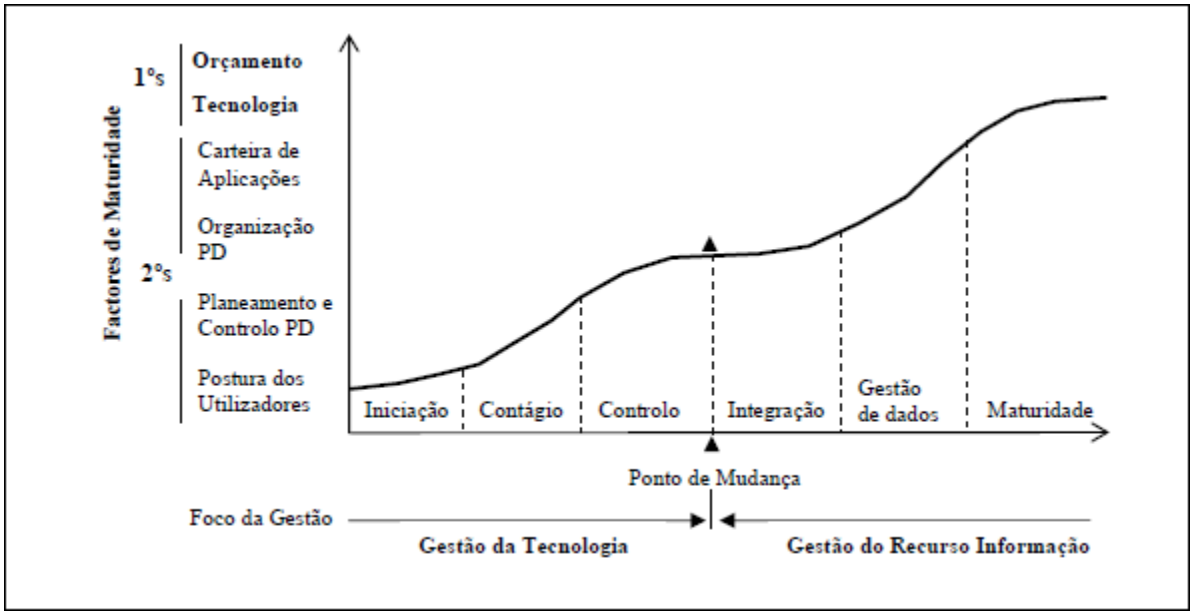


Figura 1- Modelo de seis estádios de crescimento de Nolan

Associada às mudanças organizacionais está sempre a satisfação dos recursos humanos e no enquadramento à questão desta investigação (Qual o impacto do Sistema de Workflow no processo de Licenciamento de Redes e Serviços de Radiocomunicações da ANACOM?), um dos objetivos é medi-la face ao uso do Sistema de Informação (SI) MyWs e o *Technology Acceptance Model* - TAM de Davis (1989) permitiu medir os efeitos da utilidade percebida e a intenção relativamente ao uso de um SI, verificados nos utilizadores. Os efeitos na eficiência e eficácia no processo de negócio levou também à questão de analisar a evolução deste modelo teórico – TAM – na inclusão de aspetos adicionais relativos aos efeitos da utilidade percebida e da intenção do uso serem influenciados ou alterados com o aumento da experiência do utilizador em relação ao sistema que utiliza.

A questão das dificuldades na utilização de um Sistema de Informação (SI) é enquadrado teoricamente por Davis (1989), através do conceito, desenvolvido pelo autor, da utilidade percebida, que significa que as pessoas tendem a usar ou não um SI com o objectivo principal de melhorar o seu trabalho ou, por outro lado, da facilidade de uso percebida, que significa que mesmo que a pessoa veja utilidade numa dada tecnologia, esta se for de difícil utilização não lhe dedica esforço para a utilizar.

Por último, evidencia-se no impacto dos sistemas de workflow a variável produtividade e com fortes implicações nesta encontra-se a eficiência que segundo Viana (2010) revela a relação entre os recursos consumidos e os resultados que se pretendem alcançar com vista à otimização dos meios em função dos fins. Afirmando que o conceito de eficiência está associado ao conceito de eficácia, relacionando os fins ou resultados alcançados com os objetivos previamente definidos.

A produtividade é definida pela diferença entre os resultados obtidos e a quantidade de recursos utilizados, ou então, "o ratio entre o output e o input", compondo a eficiência o conceito de produtividade.

4. Recolha de dados

Os dados recolhidos, por consulta a base de dados de indicadores de gestão, responderam às questões de tempo de trabalho e custos envolvidos no cenário pré e pós implementação do Workflow. Recolheram-se dados relativos ao processo desde 2009 a 2013 sobre: (i) número de colaboradores; (ii) horas trabalhadas anualmente; (iii) custos com salários e benefícios dos trabalhadores; (iv) pedidos recebidos anualmente; e (v) pedidos de licenciamento concluídos anualmente.

Na base de dados de reporte de horas foram retiradas as horas trabalhadas dos colaboradores e o número de recursos afetos ao processo desde 2009 a 2013. Neste sentido, recolheram-se dados

segundo as seguintes variáveis: mês; ano; a Direção a que os colaboradores pertenciam no sentido de perceber se trabalhavam para o processo, a atividade ou atividades que efetuavam; o serviço a que essa atividade pertenciam, se relativas ao processo ou não; e as horas efetivas trabalhadas.

Relativamente à verificação do objetivo satisfação dos recursos humanos, realizou-se um questionário interno anónimo e de resposta fechada que cobriu todo o universo de colaboradores do processo. O questionário foi distribuído a um universo de 20 colaboradores, todos eles pertencentes à divisão de consignações de frequências e licenciamentos, que é o serviço de que é "dono" do processo em análise, o qual foi criado na ferramenta GoogleDocs.

O questionário utilizado começou por se iniciar com uma caracterização do inquirido, onde o mesmo se identificou quanto ao género, às habilitações académicas, ao grupo profissional, ao nível de conhecimentos em tecnologias de informação e quanto ao ter trabalhado anteriormente com uma aplicação com workflow.

Após a identificação do inquirido, o questionário dividiu-se em 5 secções de A a D (anexo B). Estas secções foram identificadas relativamente à utilização do workflow, designadamente quanto à formação, à utilidade e facilidade percebida e à intenção de usar.

Estas secções foram preenchidas com 20 afirmações apresentadas conforme os assuntos dessas secções e foi usada a escala de Likert de 1 a 7, variando de (1) discordo plenamente e (7) concordo plenamente.

5. Análise de Dados

À questão sobre o impacto de um sistema de workflow no processo de licenciamento de redes e serviços de radiocomunicações – um estudo de caso na Autoridade Nacional de Comunicações, levanta-se a hipótese desse impacto ter aumentado a eficiência e a eficácia do processo com o aumento da produtividade, sendo que este aumento se deveu à intenção dos colaboradores que trabalham para o processo em utilizarem o sistema de workflow no futuro.

A começar pela análise da eficiência operacional do processo, apresentam-se na tabela 1 os seguintes dados: i) horas trabalhadas; ii) número de colaboradores; iii) e salários e benefícios auferidos por estes colaboradores; iv) as solicitações recebidas e v) as solicitações fechadas.

Anos	Recursos Humanos [a)]	Horas [b)]	Monetários e Outros (m.o.) [c)]	Solicitações Recebidas [d)]	Solicitações Fechadas [e)]
2009	35	14478:35	92 723,50 €	439	439
2010	29	13830:02	88 730,90 €	1856	1752
2011	30	14313:08	95 135,80 €	2364	2315
2012	30	14136:44	94 445,50 €	2292	2269
2013	17	8874:30	50 428,10 €	1400	1397

Tabela 1 - Eficiência do processo

a) Colaboradores que trabalham para o processo de licenciamento de redes e serviços de radiocomunicações.
b) Horas trabalhadas anualmente pelas pessoas que trabalham para o processo.
c) Custos com salários e benefícios dos trabalhadores que trabalham para o processo.
d) Pedidos de licenciamento de redes e serviços de radiocomunicações recebidos anualmente.
e) Pedidos de licenciamento de redes e serviços de radiocomunicações concluídos anualmente.

Os dados dizem respeito a 2009, ano em que o workflow entrou em produção, até 2013. Destaca-se, numa primeira análise, nesses dados, o número de colaboradores que trabalhava no processo em 2009 ser maior por comparação com 2010, 2011, 2012 e 2013. A este facto deve-se o impacto do workflow começar a sentir-se mais nos anos posteriores a 2009, tal como é demonstrado na teoria de estádios de maturidade de Nolan (1979).

Verifica-se na investigação em curso (gráfico 6) um ponto de transição que define o fim de um ciclo de controlo da tecnologia para dar lugar ao início de um outro ciclo, o da integração e do crescimento explosivo do sistema de workflow e da aprendizagem organizacional, tal como é enquadrado na teoria de estádios de maturidade de Nolan, quando o autor se refere ao fim da primeira curva em "S" como o controlo, dando-se de seguida o início da segunda curva em "S", a da integração.

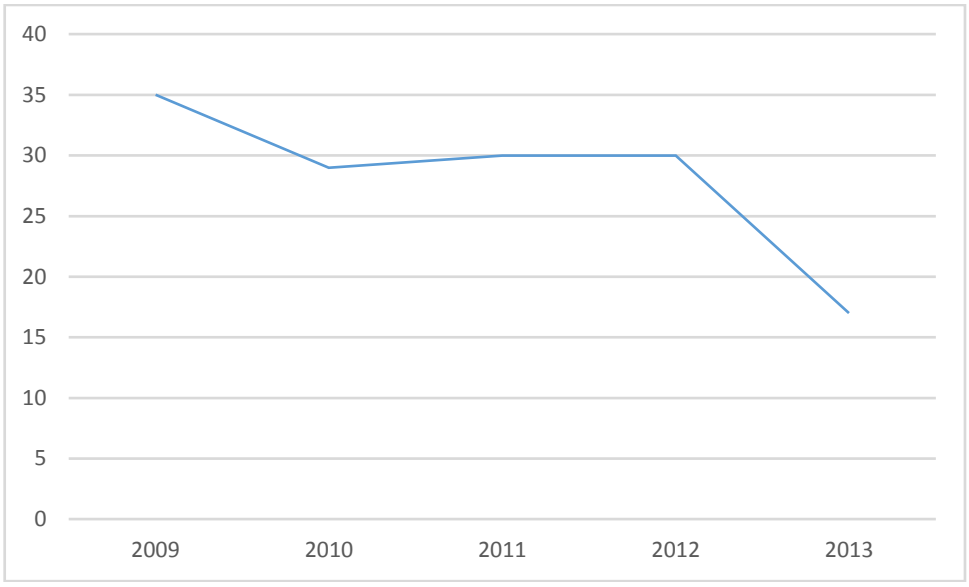


Gráfico 1 - Evolução do número de colaboradores que trabalham para o processo de licenciamento de redes e serviços de radiocomunicações

O início da integração do sistema de workflow começou em 2010 com as tarefas relativas ao processo a deixarem de ser suportadas em papel. O fluxo do processo foi repensado apesar da lógica das tarefas se manter igual ao do suporte de papel.

No presente estudo de caso, a diminuição do número de colaboradores desde 2009 até 2013 deveu-se essencialmente às alterações ao nível da coordenação e controlo das tarefas pela implementação do workflow. Houve uma centralização da coordenação do processo e uma diminuição de trabalho administrativo que se encontrava disperso. Este controlo e coordenação, tal como é adiantado pela tese de Sarmento (2005), no estudo de caso em análise verifica-se no facto do sistema de workflow avisar o utilizador de quando é que ele tem de tomar alguma iniciativa ou decisão, o que não acontecia aquando do suporte em papel.

Verifica-se, conforme ilustra o gráfico 2, que para cada solicitação em 2009 foram precisas 32 horas e 58 minutos com custos monetários e outros (m.o.) de 211,22 €/ano. Em 2013 foram gastas 6 horas e 20 minutos de trabalho com cada solicitação fechada e com um custo m.o. de 36,07€. E comparando 2009 com 2013, verifica-se uma diferença substancial tanto ao nível das horas trabalhadas como dos custos m.o. com cada solicitação fechada.

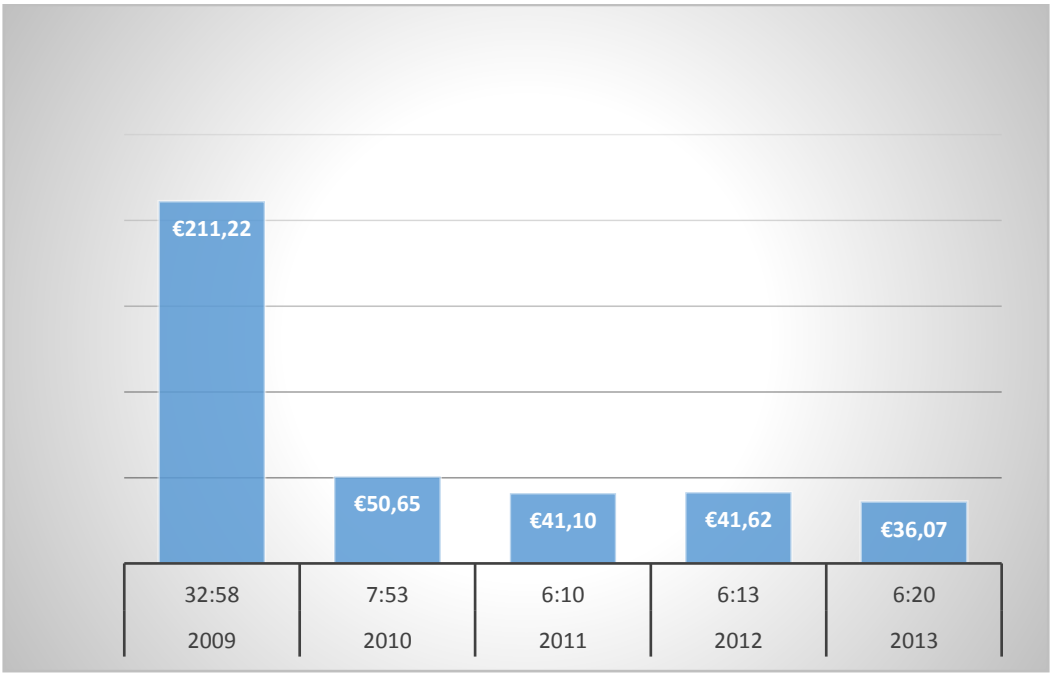


Gráfico 2 - Custo e tempo de trabalho anuais por cada solicitação

A eficiência foi aumentado gradualmente ao longo dos cinco anos, conforme ilustra o gráfico 3, verificando-se que em função dos resultados obtidos (solicitações fechadas) – os Outputs - houve um menor dispêndio de recursos, designadamente os recursos humanos foram diminuindo, registando-se em 2009 um número de 35 colaboradores que trabalhavam para o processo e em 2013 apenas 17, uma redução de 48,5% que se deve a um ajuste dos fluxos de trabalho no que diz respeito à coordenação e controlo das tarefas.

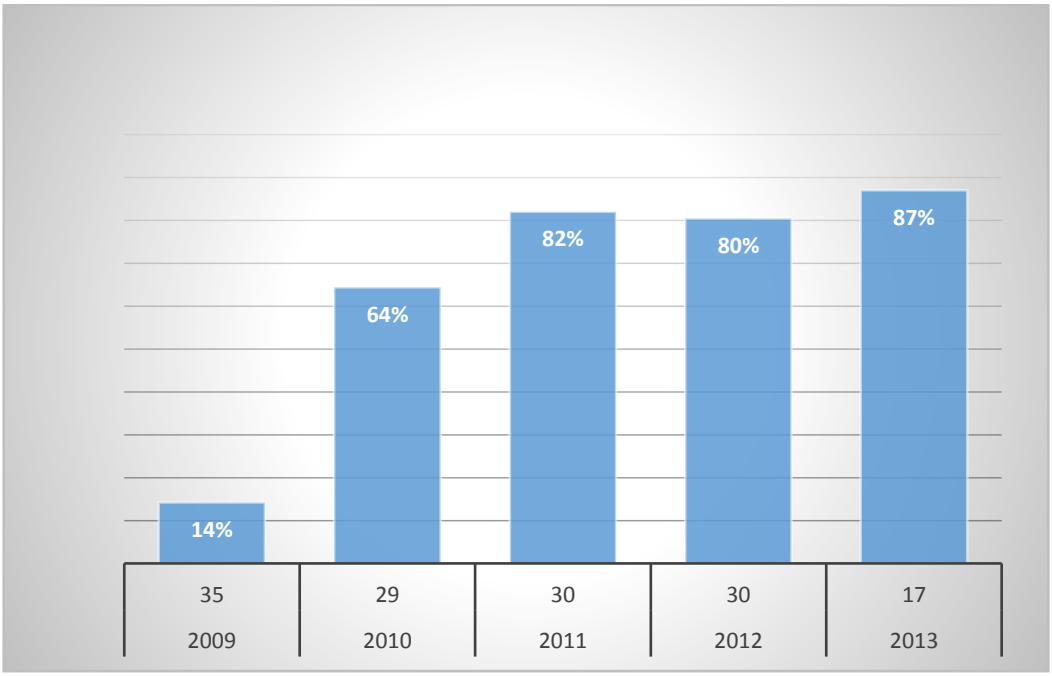


Gráfico 3 - Evolução da Produtividade e recursos humanos

Portanto, a eficiência operacional é aqui entendida, como faz referência Silvestre (2009), pela diferença entre os resultados obtidos e a quantidade de recursos utilizados, isto é, o ratio entre o Output e o Input, no que resulta na análise aos dados os seguintes factos:

- triplicação de pedidos e na sua satisfação;
- redução do número de colaboradores que trabalham para o processo;
- diminuição do número de horas trabalhadas;
- diminuição dos custos monetários e outros.

Concluindo-se desta forma que o ratio entre o Output (triplicação de pedidos e na sua satisfação) e os vários Inputs (redução do número de colaboradores que trabalham para o processo; diminuição do número de horas trabalhadas; e diminuição dos custos monetários e outros.) resultaram num aumento da produtividade tal como ilustra o gráfico 3, logo um aumento da eficiência, devendo-se este facto à adoção do sistema de workflow, cujo impacto se faz sentir na alteração de fluxos de trabalho, os quais se traduzem em encurtar o tempo de encaminhamento de uma determinada tarefa que passa pela desmaterialização de documentos.

A adoção do sistema de workflow e o seu impacto no processo, designadamente no aumento gradual da produtividade que implica o aumento da sua eficiência poderá estar implícito a este fenómeno o efeito de "curva de aprendizagem", conceito, apresentado por Oliveira (2012), que

estima a taxa de aprendizagem que permite, através da experiência acumulada, a realização de tarefas num período mais curto ou com um custo médio decrescente.

A eficiência do processo traduziu-se num aumento do índice de produtividade que resultou também na superação dos objetivos tornando-o também mais eficaz e a causa são os novos modelos de trabalho que por consequência surgiram com a adoção do sistema de workflow.

A eficácia organizacional foi medida através dos tempos médios de resposta às solicitações recebidas e tratadas desde 2009 a 2013, estes tempos foram retirados da aplicação de indicadores de gestão da unidade organizacional da ANACOM que gere o processo de licenciamento de redes e serviços de radio comunicações. A eficácia vai aumentando desde 2010, quando os fins ultrapassam os objetivos propostos, verificando-se, designadamente que em 2013 o objetivo era de 27 dias úteis para resposta a solicitações recebidas e tratadas e o resultado apresentado foi de 21 dias úteis.

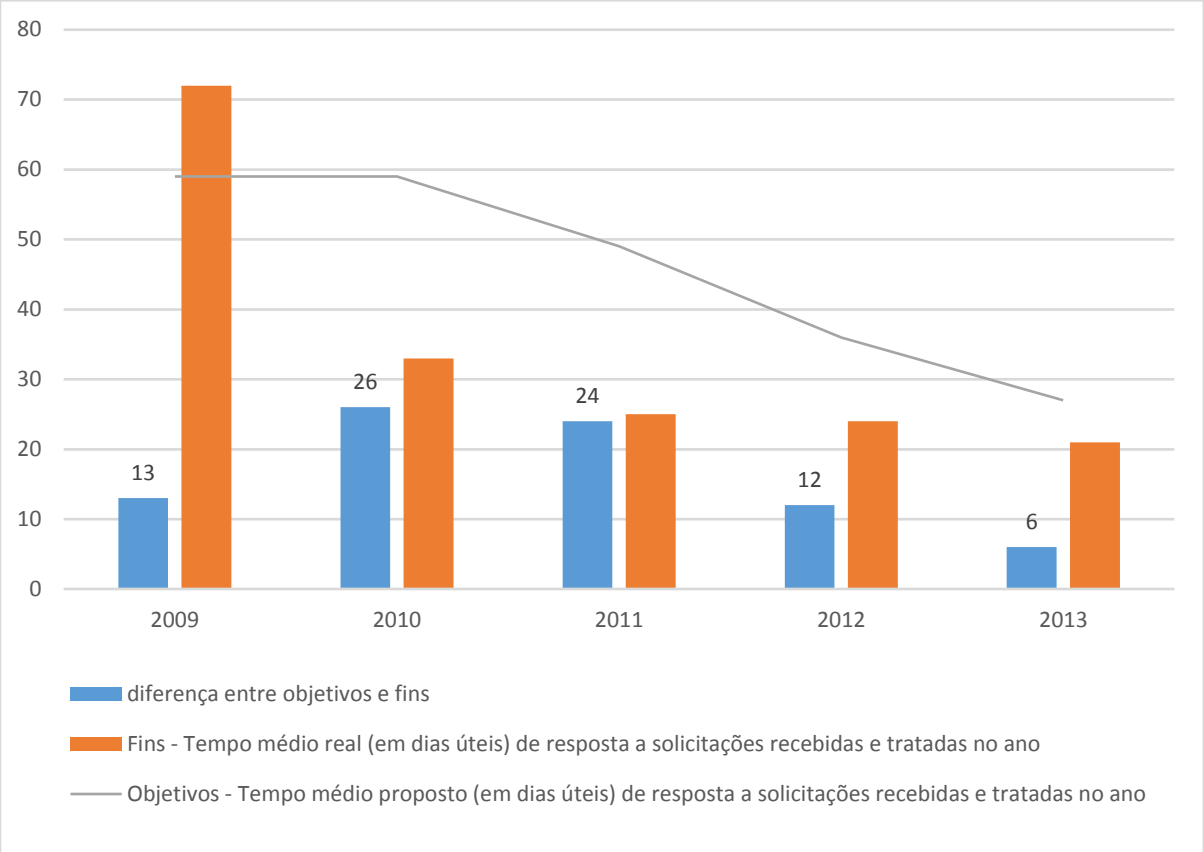


Gráfico 4 - Eficácia do Processo de Licenciamento de Redes e Serviços de Radiocomunicações

Os colaboradores que trabalham para o processo têm intenção de usar o sistema de worflow no futuro?

A questão é levantada no âmbito do objetivo sobre a satisfação dos recursos humanos que trabalham para o processo em análise, designadamente quanto à perceção, à facilidade e à utilidade que os colaboradores vêm no uso do sistema de workflow para o processo. A medição destas três variáveis é enquadrada teoricamente no Technology Acceptance Model (TAM) de Davis (1989).

Para verificação desta questão foi lançado um questionário que, segundo o coeficiente alfa de Cronbach, demonstrou ter um índice de confiabilidade excelente ($\alpha = 0{,}9569$). Este questionário foi dividido em três variáveis (utilidade do workflow, a facilidade de uso do workflow e a intenção de usar o sistema de workflow no futuro) que responderam à questão.

A análise relativa às variáveis levam à conclusão que há uma concordância de 50% na intenção de uso do sistema de workflow no futuro, apesar da concordância na sua utilidade ser de 67%.

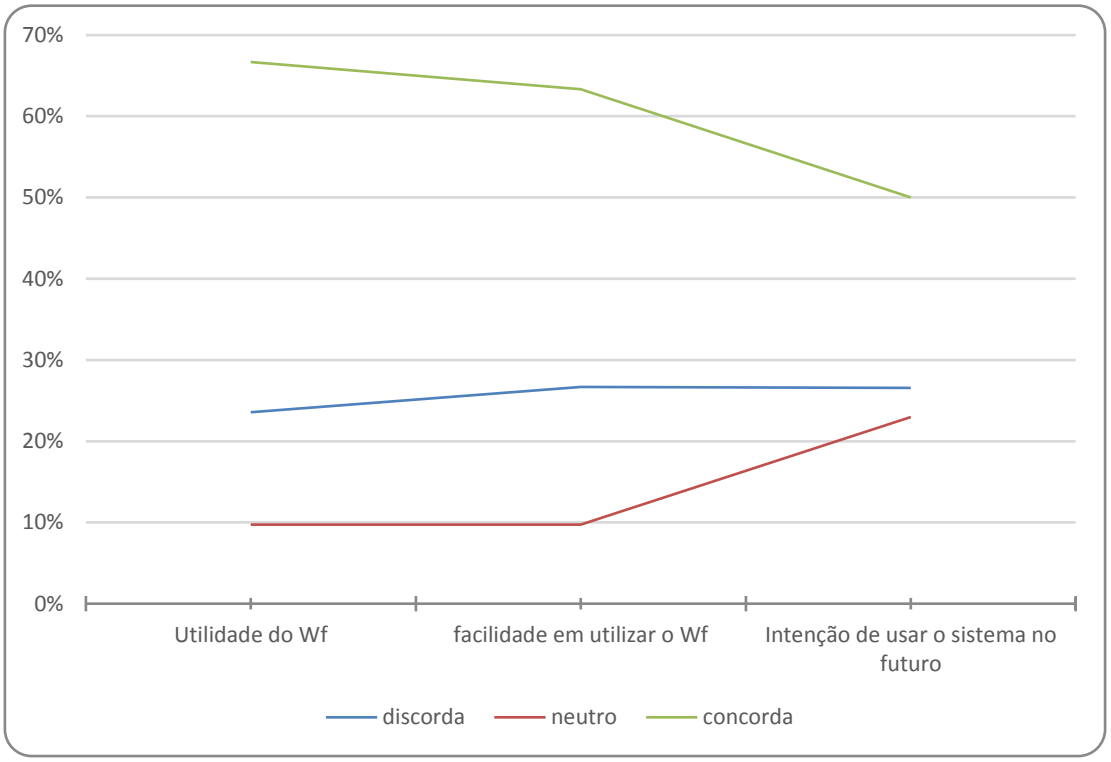


Gráfico 5 - facilidade e utilidade percebida na intenção de usar o sistema de workflow

A análise à utilidade do sistema de workflow (MyWs) foi feita através das seguintes questões:

- aprender a usar o MyWs foi fácil;
- foi necessário muito tempo para aprender a usar o MyWs;
- O MyWs é intuitivo;
- O MyWs promove a harmonização de procedimentos no processo;
- O uso do MyWs melhorou a eficiência do processo de trabalho;
- O MyWs acrescenta valor à ANACOM;
- A utilização do MyWs facilita a circulação de documentos;
- A utilização do MyWs facilita o acesso à informação.

E, foram evidenciadas as questões da eficiência do processo e do valor acrescentado para a ANACOM, na medida em que respondem mais diretamente à questão fulcral, que é a de saber sobre a intenção dos colaboradores do processo utilizarem o sistema de workflow no futuro, mas também servirá para cruzar com o objetivo da eficiência operacional relativamente ao impacto do sistema de workflow no processo.

Verifica-se então que, da utilidade percebida dos colaboradores que trabalham para o processo, há uma concordância de 80% (gráficos 6 e7) de que a eficiência do processo aumentou com a adoção do sistema de workflow e que este acrescenta valor à organização.

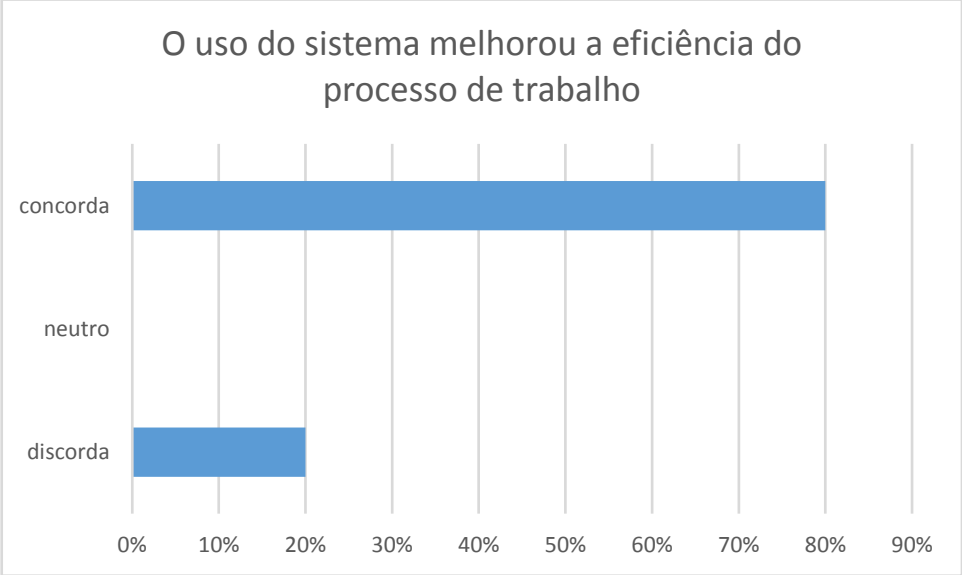


Gráfico 6 - utilidade percebida do sistema de workflow pelos colaboradores do processo

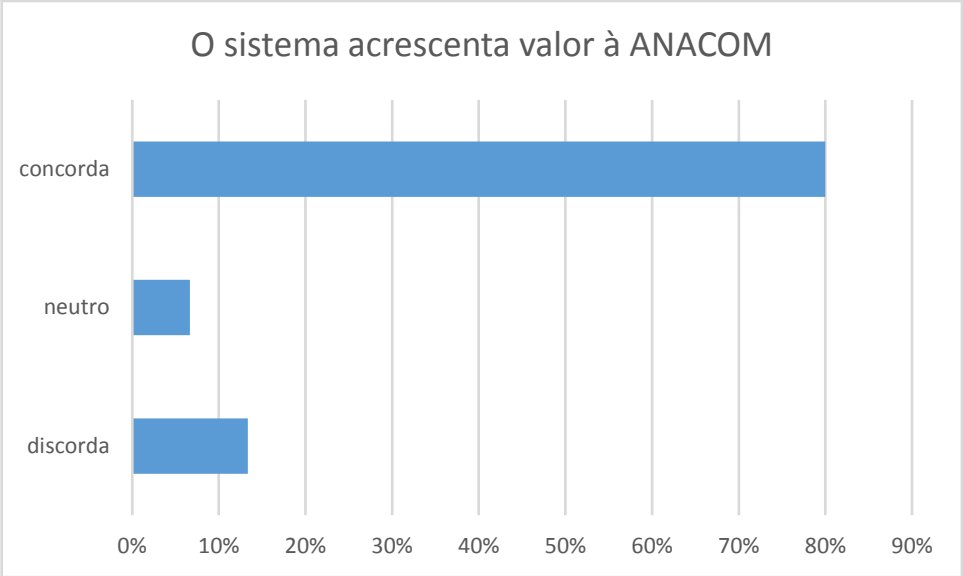


Gráfico 7 - utilidade percebida do sistema de workflow pelos utilizadores do processo

Por comparação com o estudo da Sardinha (2011), gráfico 8, há um aumento de 21% de concordância na utilidade do sistema de workflow, sendo que a intenção de o utilizar no futuro teve apenas um aumento de 3%, que é muito pouco significativo.

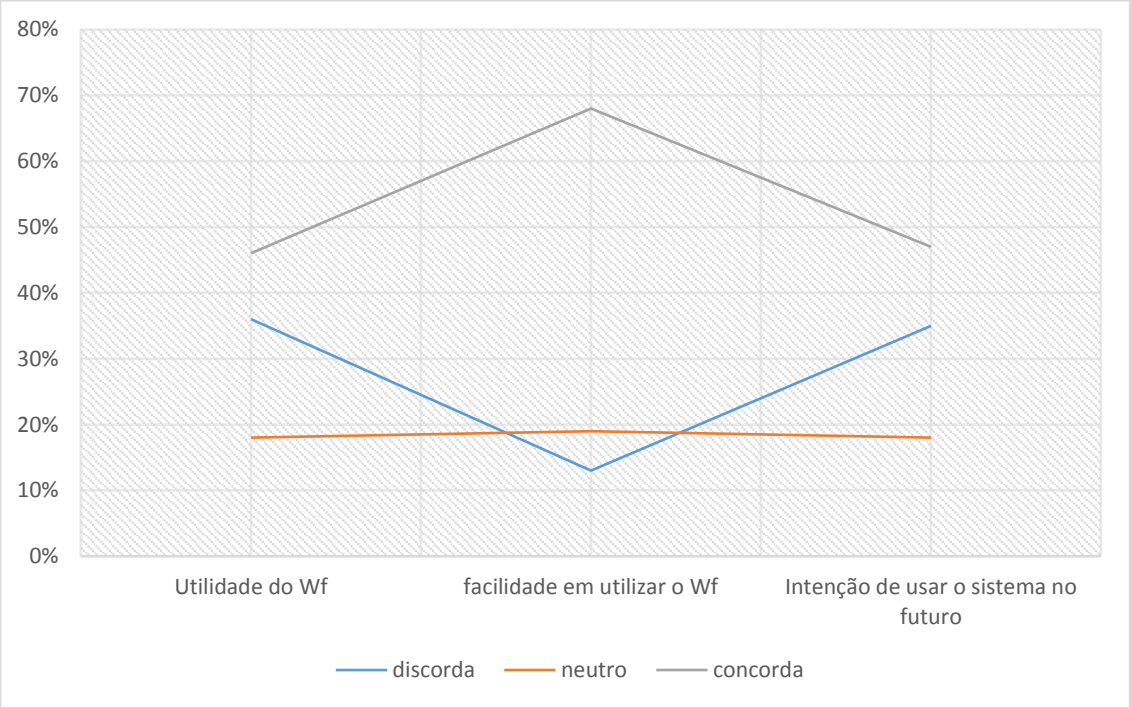


Gráfico 8 - facilidade e utilidade percebida na intenção de usar o sistema de workflow

6. Conclusão e trabalho futuro

Foram definidos objetivos que visaram determinar o impacto do workflow no processo em termos de eficiência operacional, de eficácia organizacional e determinar a intenção dos colaboradores que trabalham para o processo em utilizarem o sistema de workflow no futuro.

O resultado do impacto no processo, com a adoção de um sistema de workflow, foi o aumento da sua eficiência e da sua eficácia devido ao aumento da produtividade. O aumento da produtividade, por sua vez, deveu-se à intenção dos colaboradores que trabalham para o processo em utilizarem o sistema de workflow no futuro.

A justificação do resultado do impacto foi devidamente verificado no capítulo anterior, tendo em conta os objetivos definidos e o seu enquadramento teórico, a determinação da eficiência e da eficácia através das variáveis: i) as horas trabalhadas para o processo; ii) o número de colaboradores que trabalham para o processo; iii) os salários auferidos pelos colaboradores que trabalham para o processo; iv) e o número de pedidos de licenças que deram entrada anualmente e foram concluídos.

Foram também analisadas outras variáveis que serviram o objetivo de saber sobre a intenção dos colaboradores que trabalham para o processo em utilizarem o sistema de workflow no futuro, designadamente quanto à perceção, à facilidade e à utilidade que os colaboradores vêm no uso do sistema de workflow para o processo.

Em virtude do que foi mencionado conclui-se que o resultado, por exemplo, da análise dos dados sobre a eficiência e a eficácia do processo de licenciamento de redes e serviços de radiocomunicações deu que o processo ficou mais eficiente e eficaz, demonstrado pelo índice de produtividade ter aumentado de 14% em 2009 para 87% em 2013, período em que foi adotado o workflow, sendo que se verifica a posição de Khoshafian (1995) e Jablonski (1996) sobre os sistemas de informação, designadamente workflows, serem a solução capaz de melhorar a eficiência e a gestão dos processos organizacionais.

O índice de produtividade aumentou porque houve uma gestão eficiente dos recursos – o ratio entre o output e o input – que se deveu à adoção do sistema de worflow por este ter alterado fluxos de trabalho, os quais se traduziram em encurtar o tempo de encaminhamento das tarefas pela desmaterialização de documentos.

Os workflows melhoram a eficiência e a gestão dos processos organizacionais relativamente à capacidade que estes têm de fornecer as ferramentas que facilitam a automatização e a redução de tempo de realização das tarefas, tal como referem Khoshafian (1995) e Jablonski (1996). Sendo que para cada pedido de licença em 2009 foram precisas 32 horas e 58 minutos com custos

monetários e outros (m.o.) de 211,22 €/ano e em 2013 foram gastas 6 horas e 20 minutos de trabalho com cada pedido com um custo m.o. de 36,07 €.

Portanto, a eficiência do processo de licenciamento de redes e serviços de radiocomunicações aumentou devido a adoção do sistema de workflow, concluindo-se desta forma que o houve uma diminuição de dias úteis de resposta a pedidos de licenças e também um aumento da eficácia do processo.

A intenção de usar o sistema de worflow no futuro pelos colaboradores que trabalham para o processo foi a questão que se pretendeu saber na perspetiva de perceber sobre a satisfação dos recursos humanos.

A conclusão que se chegou foi a de haver uma concordância de 50% na intenção de usarem o sistema de workflow no futuro, apesar de 67% concordarem com a utilidade do mesmo. Por comparação com estudo da Sardinha (2011), houve um aumento de 21% de concordância na utilidade do sistema de workflow e a intenção de utilizarem o sistema no futuro já se cifrava nos 50%.

Relativamente ainda à satisfação dos recursos humanos, evidencia-se o facto de haver uma concordância de 80% de que a eficiência do processo aumentou com a adoção do sistema de workflow e que este acrescenta valor à organização, sendo que há coerência com os resultados sobre o índice de produtividade.

O aumento da eficiência do processo foi efetivo, a mudança foi do tipo tecnológica e como tal relativa ao processo de como a organização realiza o seu trabalho, foi uma mudança que não teve o objetivo de tornar mais eficiente o processo de negócio, apesar de essa mudança o tornar implícito conforme Kotter (1996).

Fica a questão de como interpretar o sucesso do sistema de workflow no processo sem arquitectura de empresa? A questão coloca-se na medida em que não há o garante do alinhamento entre a missão da ANACOM (objetivos corporativos e orientação estratégica) e as tecnologias, apesar de haver uma resposta concreta positiva aos objetivos delineados na estratégia da organização, sobre as respostas às solicitações de pedidos de licenciamento de redes e serviços de radiocomunicações, cujo resultado se traduziu num efetivo aumento da eficácia do processo com a adoção do sistema de workflow.

Contudo, evidencia-se que há alinhamento da missão da ANACOM com as tecnologias da informação apenas no que diz respeito ao processo. Há um desenvolvimento, por parte do departamento que gere o processo, de sistemas de informação que servem apenas esse

departamento e o sistema de workflow, com uma arquitetura de dados pensada em função das funcionalidades desse sistema.

Portanto, não basta às organizações apostarem apenas no desenvolvimento e implementação rápida de sistemas de informação, mas também na modelação dos processos de negócio, de forma a construir um modelo que reflita as necessidades Lousã (2000). Os sistemas de workflow, definidos por Lousã (2000), são sistemas de gestão proativos que gerem os fluxos de trabalho entre os participantes (utilizadores ou outros sistemas), de acordo com procedimentos pré definidos que constituem as tarefas. Estes sistemas coordenam os participantes e recursos, de acordo com os objectivos pretendidos e dentro do tempo previsto, mas não coordenam os dados e as entidades representados na especificação dos processos, continuando estes dados a serem isolados.

Haveria espaço para uma outra investigação se a procura fosse sobre como arquitectar e organizar os dados dispersos e desconectados de modo a criar uma Arquitectura de Informação que alimentasse uma Arquitectura de Aplicações que poderia muito bem ser uma agregação – workflow.

7. Referências

Attinger, Monique. 1997. *Is quality an art form?* New York : John Wiley & Sons, Inc., 1997.

Bilhim, João. 2013. *Teoria organizacional: Estruturas e pessoas*. Lisboa : ISCSP (7.ª edição), 2013.

Boyle, R. D.. 1995. *Avoiding common pitfalls of reengineering*. Montvale : Management Accounting, 1995.

Branski, Regina Meyer, Franco, Raul Arellano Caldeira e Lima Jr., Orlando Fontes. *METODOLOGIA DE ESTUDO DE CASOS APLICADA À LOGÍSTICA*.

Coghlan, D.. 1993. *A person-centred approach to dealing with resistance to change*. s.l. : Leadership and Organization Development Journal, 1993.

Cunha, Miguel Pina e al, et. 2007. *Manual de comportamento organizacional e gestão*. s.l. : RH Editora, 2007.

Davis, F.D.. 1989. *Perceived usefulness, perceived ease of use, and user acceptance of information technology*. s.l. : MIS Quarterly, 1989.

Denning, Peter J. and al, et. 1989. *Computing as a discipline*. s.l. : Communications of the ACM, 1989.

Fishbein, M. & Ajzen, I.. 1975. *Belief, attitude, intention, and behavior: An introduction to theory and research*. s.l. : Reading, MA: Addison-Wesley, 1975.

Georgakopoulos, D., Hornick, M and Shet, A. 1995. *An Overview of Workflow Management: from process modeling to workflow automation infrasestructure*. s.l. : Distributed and Parallel Databases, 1995.

Hales, Keith. 1997. *Workflow in context*. New York : John Wiley & Sons, Inc, 1997.

Iskatt, G. J. and Liebowitz, J.. 1996. *What to do when employees resist change*. s.l. : Supervision, 1996.

Jablonski, Stefan and Bussler, Christoph. 1996. *Workflow Management Systems*. *Modelling Concepts, Architecture and Implementation*. s.l. : Cengage Learning EMEA, 1996.

Jordan, Patrick W.. 1998. *An Introduction to Usability*. s.l. : Taylor & Francis, 1998.

Khoshafian, Setrag. 1995. *Introduction to Groupware, Workflow and Workgroup Computing*. New York : John Wiley & Sons, Inc, 1995.

Kotter, J. P. and Schlesinger, L. A.. 1979. *Chossing strategies for change*. Boston : Harvard Business Review, 1979.

Kotter, John P. 1996. *Leading Change*. s.l. : Library of Congress Cataloging-in-Publication Data, 1996.

Kurtz, P. and Duncan, A.. 1998. *Shared service centres: overcoming resistance to implementation of a shared service centre*. Montvale : Management Accounting, 1998.

Lewin, Kurt. 1947. *Frontiers in group dynamics*. *Human Relations*. New York : s.n., 1947.

Loureiro, João Carlos Simões Gonçalves. 1995. *O procedimento administrativo entre a eficiência e a garantia dos particulares*. Coimbra : Coimbra Editora, 1995.

Lousã, Mário, Sarmento, Anabela e Machado, Altamiro. 2000. Sistema de Automatização de Processos de Negócio (Workflow e Systems): Considerações sobre o contexto organizacional e proposta de estrutura de análise do seu impacto nas organizações. 2000.

Nielsen, Jacob. 2003. *Usability 101: Introduction to usability*. s.l. : Alertbox, 2003.

Nolan, R. 1973. *Managing the computer resource: a stage hypotesis*. s.l. : communications of the ACM, 1973. Vol. 16, n.º 7, pp. 399-405.

Norman, D. A.. 2004. *Emotional Design: Why we love (or hate) everyday things*. New York, NY : Basic Books, 2004.

Oliveira, Luis Manuel Botelho. 2012. Professor adjunto do ISCAL. *Projeto Académico Individual - Componente de Investigação - Curvas de experiência, alterações tecnológicas e vantagens competitivas*. Lisboa : ISCAL, outubro de 2012.

Poe, C. D. and Viator, R. E.. 1990. *What to do when employees resist automation*. s.l. : Journal of Accounting and EDP, 1990.

Porras, Jerry I. e Robertson, Peter J. 1991. *Organization Development: Theory, Research, and Practice*. Stanford : Graduate School of Business, Stanford University, 1991.

Preece, J, Rogers, Y and Sharp, H.. 2002. *Interaction design – beyond human-computer interaction*. s.l. : John Wiley & Sons, Inc, 2002.

Sardinha, A. Filipa. 2011. *A Formação em Tecnologias de Informação e as Características da Interface na Aceitação da Tecnologia - uma extensão ao modelo TAM*. Lisboa : Dissertação submetida ao Instituto Universitário de Lisboa (ISCTE-IUL) como requisito parcial para obtenção do grau de Mestre em Gestão de Sistemas de Informação, 2011.

Sarin, K. Abbott & S. 1994. *Experiences with Workflow Management: Issues for the Next Generation*. North Carolina : Proceedings of CSCW 94, 1994.

Sarmento, Anabela e Gonçalves, Paulo. 2005. *Os sistemas de gestão de processos de negócio e o seu impacto nas organizações – apresentação de um caso*. 2005. pp. 1-10.

Sarmento, Anabela. 2002. *Impacto dos Sistemas Colaborativos nas Organizações: Estudo de Casos de Adopção e Utilização de Sistemas Workflow - Tese de Doutoramento*. Braga : Universidade do Minho, 2002.

Shneiderman, B.. 1999. *Designing the User Interface, Strategies for Effective Human-Computer Interaction*. s.l. : Addison Wesley, 1999.

Silvestre, Hugo Consciência. 2009. *Gestão pública: Modelos de prestação no serviço público*. Lisboa : Almedina, 2009.

Stark, Heather. 1997. *Understanding workflow*. NY : Workflow handbook 1997, 1997.

Venkatesh, A. and Davis, F.D. 2000. *A theoretical extension of the technology acceptance model*. s.l. : Management Science, 2000.

Venkatesh, V., et al. 2003. *User acceptance of information technology: Toward a unified view*. s.l. : MIS Quarterly, 2003.

Viana, Cláudia. 2010. O Princípio da eficiência: A eficiente eficácia da administração pública. 2010. pp. 301-311.

Vlachantonis, Nikolaos. 1997. *Workflow Applications within Bussiness Organizations*. s.l. : Lawrence (Ed.), Workflow Handbook 1997, Chichester: John Wiley & Sons., 1997.

WfMC. 2004. Workflow Management Coalition. *Workflow Management Coalition*. [Online] 2004. http://www.wfmc.org .

Yin, Robert K. 1994. *Case Study Research: Design and Methods*. 1994.

O *Twitter* como agente facilitador de recolha e interpretação de sentimentos: Exemplo na escolha da palavra do ano [1]

Carina Andrade [1], Maribel Yasmina Santos [2]

1) Centro de Investigação ALGORITMI,
Universidade do Minho, Portugal
a61575@alunos.uminho.pt

2) Centro de Investigação ALGORITMI,
Universidade do Minho, Portugal
maribel@dsi.uminho.pt

Resumo

Desde 2009 que a Porto Editora elege em "infopédia.pt" a palavra que melhor representa os anos que terminam. Este trabalho apresenta uma forma alternativa a essa eleição, substituindo a votação dos cidadãos pela recolha de dados da rede social *Twitter* ao longo do ano, e procedendo à análise dos mesmos em substituição da votação. Assim sendo, foram recolhidos dados associados às dez palavras finalistas incluídas no conjunto da palavra do ano 2014, os quais foram armazenados em ambiente *Hadoop* para seguidamente e recorrendo a dois *lexicons* ser possível a classificação dos *tweets*. Os *lexicons* utilizados incluem, por um lado, a lista de palavras positivas e negativas e, por outro, as polaridades associadas às palavras em conjugação com o top vinte e cinco de *emoticons* utilizados no *Twitter*. Os resultados obtidos permitem identificar a palavra mais referida e o sentimento, positivo ou negativo associado à mesma.

Palavras-chave: *Análise de Sentimentos*, *Tweets*, *Big Data*

1. Introdução

Desde sempre que no processo de tomada de decisão existe a tendência de se ter em consideração a opinião de terceiros. Segundo [Pang & Lee 2008], esta informação que é procurada nas opiniões dos outros, inicialmente era baseada em conhecidos ou indivíduos especializados no tema mas, depois da expansão da internet, essa recolha de informação deixou de ser tão pessoal e passou a basear-se em opiniões de desconhecidos.

Atualmente, as *reviews* livremente partilhadas por vários indivíduos em redes sociais, *blogs* e canais do *Youtube*, são uma fonte de dados essencial no apoio à tomada de decisão. Contudo, devido à rápida divulgação dessas opiniões, as organizações veem a sua reputação social influenciada e por isso têm cada vez mais interesse em analisá-las.

[1] Trabalho realizado com o apoio da FCT – Fundação para a Ciência e Tecnologia, no âmbito do projeto UID/CEC/00319/2013, e da Cloud365, Lda.

Porém, este sistema de aconselhamento baseado nas opiniões divulgadas na internet por si só não aparenta ser suficiente, essencialmente devido à quantidade de informação disponível. Assim sendo, apesar de teoricamente ser positivo analisar as opiniões disponibilizadas pelos internautas (sobre qualquer que seja o tema), é impossível conseguir-se uma visão geral do assunto visto que podem existir milhões de opiniões sobre o mesmo tópico e, a amostra selecionada pelos interessados pode não se revelar adequada.

Desta forma surge a oportunidade de utilização de técnicas de *Text Mining* e/ou *Sentiment Analysis* de maneira a possibilitar a visão geral (referida acima) sobre determinado assunto. Neste caso em concreto, o objetivo deste trabalho passa por analisar os sentimentos associados às publicações existentes na rede social *Twitter* (os *tweets*) e perceber assim, de entre dez palavras-chave finalistas, quais as mais mencionadas e que sentimentos lhes estão associados.

Este artigo encontra-se organizado em oito secções: a introdução ao trabalho (secção 1), apresentação dos trabalhos relacionados identificados (secção 2), a abordagem metodológica seguida, explicação das características dos dados e arquitetura proposta (secção 3) e exposição do que foi efetuado para o tratamento de dados (secção 4), bem como a explicação da técnica de *Sentiment Analysis* utilizada e sua avaliação (secções 5 e 6), os resultados obtidos e conclusões, secções 7 e 8, respetivamente.

2. Trabalho Relacionado

Os termos *Sentiment Analysis* e *Opinion Mining* são explorados por [Liu 2012] como representando o mesmo campo de estudo considerando que, independentemente de se mencionar um ou outro termo, se pretende referenciar o estudo das opiniões que expressam um sentimento, quer ele seja positivo ou negativo. Por outro lado, [Pang & Lee 2008] referem os dois termos como sendo do mesmo campo de estudo mas podendo ser considerados uma subárea da análise da subjetividade.

Assim sendo, e apesar de não se debruçar sobre este aspeto, desde 2009 que a Porto Editora elege em ["Infopédia" 2015], no final de cada ano civil, a palavra que melhor define o ano que termina. O conceito passa por, de um conjunto de dez palavras consideradas finalistas e que representam os assuntos que mais se pronunciaram durante o ano, ser possível aos internautas interessados acederem ao portal e votarem no termo que para eles mais caracteriza o ano que termina.

No entanto, facilmente se percebe que recorrendo a redes sociais como o *Twitter*, se pode chegar às mesmas conclusões sobre o tema sem que os internautas tenham obrigatoriedade em apresentar o seu voto. Sendo o objetivo analisar as publicações dos utilizadores da rede social,

verifica-se a existência de diversas publicações com variados objetivos mas, tendo em conta essa mesma diversidade e os limites de escrita associados a este artigo, são destacadas apenas quatro publicações descritas de seguida:

- Segundo o trabalho de [Asur & Huberman 2010], através da análise dos *tweets* que vão sendo gerados consegue perceber-se os efeitos que os mesmos terão no mundo real. No caso específico deste trabalho, pela análise dos dados do *Twitter* relacionados com as estreias de filmes, os autores conseguem um modelo que prevê as receitas de bilheteira desses mesmos filmes: durante três meses os autores recolheram *tweets* relacionados com vinte e quatro filmes diferentes, recorrendo à pesquisa dos mesmos por palavras relacionadas com os seus títulos. Os autores não colocaram de parte o facto de antes das estreias, as próprias produtoras fazerem campanhas de *marketing* com lançamento de vídeos, fotografias ou mesmo declarações de atores considerando que estas campanhas de *marketing* provocam também um aumento dos *retweets*, ou seja, partilhas dos *tweets* publicitários. No que respeita à análise de sentimentos dos *tweets* recolhidos, [Asur & Huberman 2010] utilizaram dois pontos de comparação de forma a perceber o sucesso ou insucesso do filme: os valores da polaridade e a subjetividade associada aos *tweets*, isto é, a relação entre os *tweets* positivos ou negativos de determinado filme e a relação entre os *tweets* classificados como positivos ou negativos e os que foram classificados como neutros (sem sentimento explícito associado ao mesmo). A análise destes valores levou os autores a concluir que tendencialmente, a subjetividade associada aos filmes aumenta depois da estreia do mesmo, devido claro, à formação de uma opinião mais sólida por parte dos espectadores depois de assistir aos filmes e, consequentemente, classificando-os mais facilmente com sentimentos positivos ou negativos (sucesso ou insucesso do filme).

- [Gebremeskel 2011] apresenta outra perspetiva da utilização dos *tweets* para análise: as notícias. O autor recolheu, recorrendo a *APIs*, dois conjuntos de dados diferentes: os neutros (sem qualquer tipo de sentimento associado ao *tweet* sendo essencialmente recolhidos de agências de notícias) e os dados subjetivos (*tweets* com sentimento positivo, negativo ou ambos) apoiados por dois conjuntos de *emoticons* (positivos e negativos). Este último *dataset* foi explorado de forma a ser dividido por idiomas isto é, dos dados recolhidos definiu um conjunto de dados de *tweets* escritos em inglês e outro com os *tweets* escritos noutro qualquer idioma. Partindo dos dados recolhidos, o autor testou *unsupervised approach* e *supervised approaches*: abordagem baseada em palavras-chave (utilizando um dicionário de palavras positivas e outro de palavras negativas, foram contadas as palavras positivas e negativas em cada *tweet* e esse valor

definia a positividade ou negatividade do mesmo) e algoritmos de *machine-learning* (como o *Naive Bayes*) concluindo o autor que esta última abordagem supera a primeira mencionada.

- Num outro trabalho os autores debruçaram-se sobre a capacidade de perceber informação implícita em linguagem informal (e por vezes criativa) utilizada frequentemente em redes sociais e *blogs*. Para tal, [Kouloumpis, Wilson & Moore 2011] utilizaram três *datasets* diferentes: o primeiro, um conjunto de *tweets* com *Hashtags* - *tweets* recolhidos do *Edinburgh Twitter Corpus*; o segundo *dataset*, *tweets* com *emoticons* positivos ou negativos (descartando os que contêm ambos); por último, o terceiro conjunto de dados utilizado, denominado de *iSieve* com aproximadamente 4000 *tweets*. Os dados recolhidos foram tratados tendo em conta três objetivos: identificar *emoticons* e abreviaturas (substituindo estas últimas pelo seu significado), identificar intensificadores de sentimentos (utilização de *Caps Lock* por exemplo) e por último, identificar termos especiais do *Twitter: Hashtags*, identificadores de utilizadores ou *URLs*. Conjugando os três objetivos, os autores conseguiram concluir que para a análise de publicações no *Twitter* o mais útil e que revela melhores resultados passa pela utilização de *lexicons* em conjunto com a análise de características típicas de escrita em redes sociais (abreviaturas, *emoticons* e presença de intensificadores em palavras).
- [Kumari, Singh, More, Talpade & Pathak 2015] analisaram também no seu trabalho os sentimentos associados ao *tweets*: a recolha de dados foi efetuada recorrendo a *streaming* sem ter sido aplicada nenhuma restrição (geográfica ou idioma). Posteriormente, os *tweets* recolhidos foram traduzidos para inglês recorrendo ao *Google Translate* e pré-classificados um conjunto de *tweets* como positivos, negativos ou neutros de forma a serem utilizados para a classificação do conjunto de teste recorrendo ao *Naive Bayes*. Os resultados obtidos demonstram, segundo os autores, que este método de classificação teve uma boa performance no que diz respeito à classificação dos *tweets* consoante a sua polaridade.

3. Metodologia e Características dos Dados e da Arquitetura Proposta

O *Twitter* é considerado uma das maiores redes sociais e com maior potencial para a partilha de informação e criação de informação viral na internet. O funcionamento desta rede social é relativamente simples:

- Existem utilizadores que fazem publicações sobre assuntos que achem interessantes – os *tweets* (máximo 140 caracteres);

- Esses *tweets* podem ser partilhados por outros utilizadores, seguidores ou não do autor. Quando isso acontece a partilha é classificada como um *retweet*;
- Por outro lado, é possível dirigir mensagens a outros utilizadores: por forma a perceber-se que o conteúdo do *tweet* é direcionado a alguém específico basta identificar o utilizador usando "*@user id*".

Tal como foi referido na secção 2, a Porto Editora elege no final de cada ano e recorrendo aos votos dos internautas, a palavra que melhor o define. Utilizando a rede social *Twitter* e os dados nela partilhados, tendo em conta as dez palavras consideradas finalistas do ano 2014 em ["Infopédia" 2015] e transferindo o conceito temporalmente para o mês de Maio de 2015, pretende-se perceber qual a palavra mais mencionada no *Twitter* e que sentimento está associada à mesma. Pretende-se demonstrar com este trabalho que, tendo em conta a utilização cada vez mais frequente das redes sociais por parte da população, torna-se desnecessário abordar a mesma pedindo-lhe um voto para o assunto, sendo possível recolher os dados das publicações produzidas ao longo do ano sobre variados temas que vão surgindo na sociedade portuguesa e no final do ano, perceber qual a palavra que mais foi mencionada e associar-lhe um sentimento.

No que diz respeito à metodologia seguida, o CRISP-DM [North 2012], foram tidas em conta as fases consideradas pelo modelo: compreensão do negócio (perceção do funcionamento da eleição atual levada a cabo pela Porto Editora), compreensão dos dados (análise dos dados recolhidos do *Twitter*), preparação dos dados (tratamento efetuado sobre os dados), modelação (definição da técnica de *Sentiment Analysis* a utilizar), avaliação (avaliação do desempenho da técnica definida) e desenvolvimento (aplicação da técnica de *Sentiment Analysis* definida e análise dos resultados obtidos).

Colocando agora o foco na recolha de dados e tendo em conta que a rede social é conhecida também pela elevada utilização de *Hashtags*, para além das palavras em si foram pesquisados *tweets* que contivessem as respetivas *Hashtags* das dez palavras finalistas. A Tabela 1 representa os termos de pesquisa utilizados para a recolha de dados contendo a duplicação do termo "Legionela" devido à consideração por parte da Porto Editora como escrevendo-se a palavra com apenas um "l" mas que, facilmente depois de uma rápida pesquisa na internet, se percebe que também os portugueses a escrevem com dois "l" ("Legionella"). Assim, de forma extraordinária, os *tweets* do mesmo termo são recolhidos recorrendo a duas palavras sendo ambas consideradas "Legionela".

Palavra Finalista da Porto Editora	*Hashtag* da Palavra
banco	#banco
basqueiro	#basqueiro
cibervadiagem	#cibervadiagem
corrupção	#corrupção
ébola	#ébola
gamificação	#gamificação
jihadismo	#jihadismo
legionela	#legionela
selfie	#selfie
xurdir	#xurdir
legionella	#legionella

Tabela 1 - Termos de Pesquisa Utilizados

A recolha dos dados foi efetuada recorrendo à ferramenta *Palladian do KNIME* [*KNIME* 2015], uma ferramenta dedicada a variadas funções de análise de texto (recolha de informação, classificação de texto, reconhecimento de entidades, entre outros) possibilitando entre elas, a extração de informação de redes sociais de uma forma bastante simples (motivo pelo qual foi utilizada). Definindo uma conta do *Twitter* como conta de programador é possível através desta ferramenta, recolher dados associados aos *tweets* publicados. A Figura 1 representa o fluxo de recolha de dados do *Twitter* que, com três passos, permite extrair a informação que se procura:

1. Inicialmente definem-se os termos de pesquisa (presentes na Tabela 1);
2. No componente "*WebSearcher*" existe a possibilidade de definir o idioma de pesquisa dos *tweets* e tendo em conta o contexto do trabalho, foi definida a língua Portuguesa para a pesquisa;
3. Os dados recolhidos (exemplo ilustrado na Tabela 2) são depois guardados num ficheiro .csv (utilizado apenas como intermediário neste processo de recolha de dados visto que a extensão do *KNIME* para *Big Data* requer uma licença comercial), que integra os seguintes atributos:
    a. "*Query*" – identificação do termo associado à pesquisa que retornou o registo;
    b. "*Title*" – O *tweet* que contém o termo pesquisado;
    c. "*URL*" – O *URL* que dá acesso ao *tweet*;
    d. "*Date*" – A data a que o *tweet* foi publicado;

e. "*Coordinate*" – As coordenadas associadas ao dispositivo aquando da publicação do *tweet* (frequentemente campo nulo devido à localização desligada nos dispositivos de publicação);

f. "*Time*" – A hora (hora, minuto e segundo) a que foi efetuada a publicação.

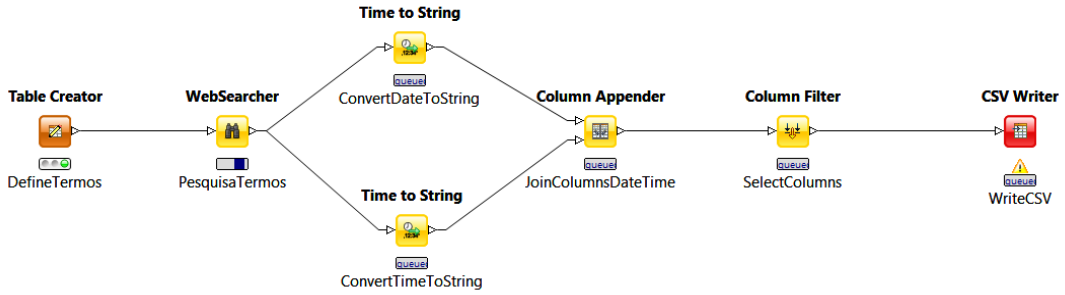


Figura 1 - Exemplo de Fluxo do *KNIME* para Recolha de Dados do *Twitter*

Query	basqueiro
Title	Estas aulas de historia e só basqueiro
URL	*HTTP://Twitter*.com/pipa_daisy21/status/596638928943321088
Date	08/05/2015
Coordinate	(valor nulo)
Time	12:33:00

Tabela 2 – Exemplo de Dados Recolhidos do *Twitter*

No que diz respeito à arquitetura tecnológica proposta para a análise de *tweets*, conforme presente na Figura 2, o *KNIME* foi utilizado para a recolha dos dados que, devido à sua quantidade e variedade, foram carregados para o *HBase* tendo sido utilizada a máquina virtual [*Cloudera* 2015]. Este passo revelou-se fundamental para a gestão dos dados recolhidos visto que o *HBase*, para além de se mostrar adequado ao armazenamento de grandes quantidades e variedades de dados, é caracterizado pela relação chave-valor que permitiu a eliminação de *tweets* repetidos com a utilização do campo "*Title*" (o *tweet*), como chave no *HBase*. Estando este passo concluído e depois do *dataset* completamente tratado recorrendo ao *Talend Open Studio for Big Data* [*Talend* 2015] e das polaridades devidamente atribuídas em *Java*, os dados foram armazenados no *Hive* para facilitar o acesso a partir do *Microsoft Excel Power BI* onde foram elaboradas as análises dos dados.

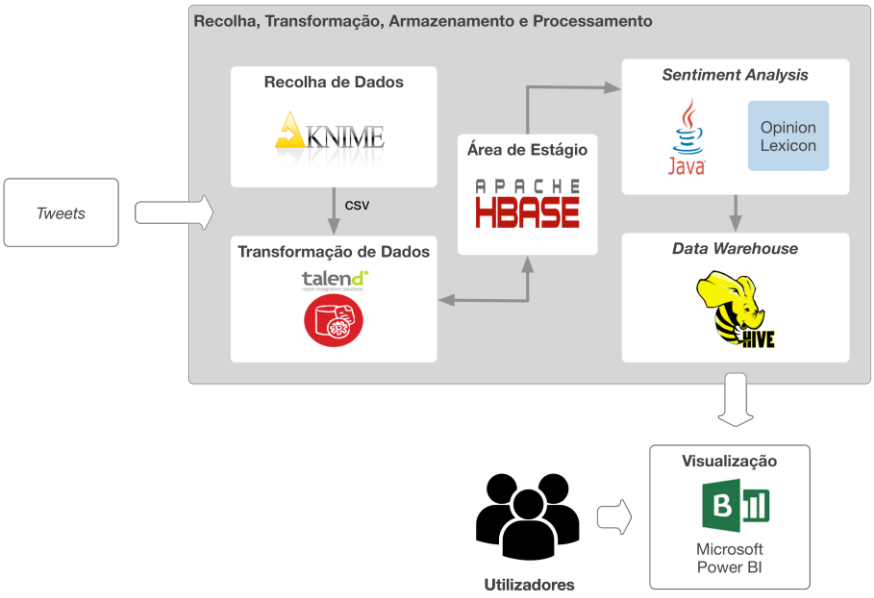


Figura 2 - Arquitetura Tecnológica Proposta para a Análise de *Tweets*

4. Tratamento de Dados

Os dados foram recolhidos entre 1 de Maio de 2015 e 25 de Maio de 2015, totalizando um valor de cerca de 500.000 *tweets*. Uma vez recolhidos e tendo em conta que os *tweets* são frases que muitas vezes contêm caracteres especiais, próprios da rede social em questão, considerou-se necessário proceder ao tratamento dos mesmos nas condições apresentadas na Tabela 3. Para além disso, conforme confirma a Figura 3, recorrendo ao [*Talend* 2015], foram também retiradas as aspas a todo o conjunto de dados recolhido.

Caso	Ação sobre o *tweet*	Ação extra
Contém *subString* "*HTTP*"	Remoção do *URL* completo associado à *subString* "*HTTP*"	Campo "*News*" = *true*
Contém *String* "RT"	Remoção da *String* "RT"	Campo "*Retweet*" = *true*
Contém caractere "#"	Remoção do caractere "#"	Campo "*Hashtag*" = *true*
Contém aspas (" ")	Remoção das aspas	
Todo o *tweet* está escrito em letras maiúsculas	Transformação do *tweet* em letras minúsculas	Campo "*CapsLock*" = *true*
Contém o caractere "@"	Remoção do caractere "@" e substituição do mesmo pela palavra "*user*"	Campo "*User*" = *true*

Tabela 3 – Ações de Tratamento de Dados

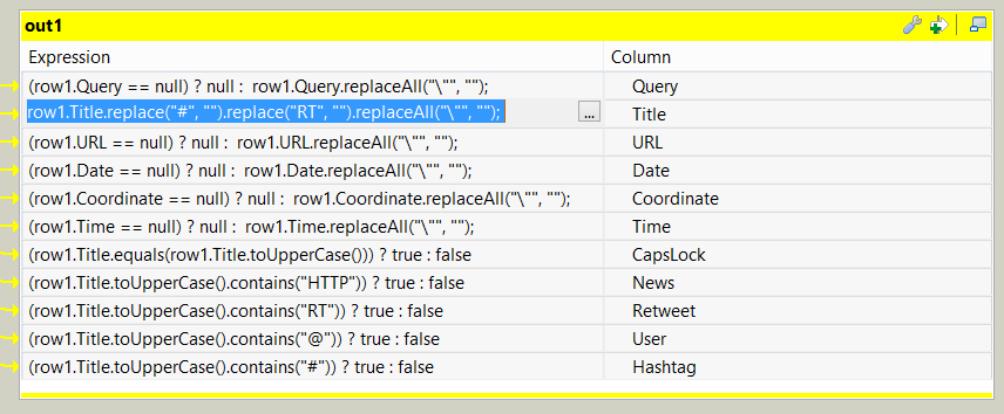


Figura 3 - Tratamento de Dados no *Talend Open Studio for Big Data*

Estas alterações aos *tweets* foram efetuadas para evitar que caracteres especiais prejudicassem a associação de polaridade às palavras. Usando como exemplo a palavra "corrupção", presente nos dicionários, pesquisando a palavra "#corrupção" não seria retornado nenhum resultado pelo facto da palavra com "#" não existir nos *lexicons*. Um outro exemplo das vantagens associadas a este tipo de tratamento de dados passa pela eliminação de *tweets* repetidos: um dos casos comuns nos dados era a repetição de frases acrescida de um *URL* que variava. Estes casos são típicos de partilha de notícias em que a descrição é a mesma mas o *URL* gerado para apresentar a notícia varia. Eliminando o *URL* dos *tweets* estes permanecem apenas com a frase do utilizador pelo que, sendo iguais, é considerado apenas um. Este facto por si só levou a uma redução de mais de 253.000 registos para um total final de 185.936, tendo sido o valor inicial de dados recolhidos de cerca de 500.000 (reduzidos a metade quando carregados para o *HBase* pela primeira vez, devido à análise dos pares chave-valor e à remoção das chaves repetidas).

Para além do tratamento já referido, foram ainda criados novos atributos que têm como objetivo caracterizar os *tweets* recorrendo à informação presente nos mesmos e que, ao ter sido eliminada seria perdida. Desta forma, e para tirar partido destas informações, as mesmas podem ser utilizadas na análise de dados e consideradas para a conclusão que se pretende retirar deste trabalho:

- Atributo "*Retweet*": um *retweet* pressupõe a existência de um *tweet*, o original, e existindo o *retweet* está-lhe associado a divulgação da informação original;
- Atributo "*User*": o mesmo pressuposto aplicado ao atributo *retweet* é aplicado ao *user* visto que, a identificação de utilizadores num *tweet*, é considerado partilha de informação;
- Atributo "*Hashtag*": sendo essa a sua principal característica, a presença de *Hashtags* num *tweet* propaga a informação presente no mesmo visto que é criado,

automaticamente, um *link* que uma vez acedido apresenta todos os *tweets* existentes com a mesma *Hashtag*;

- Atributo "News": através dos *URLs* presentes por vezes nos *tweets*, o mesmo é classificado como notícia sendo assim considerada maior a sua divulgação;
- Atributo "*CapsLock*": a existência de *tweets* escritos em letras maiúsculas remetem a uma tentativa por parte do utilizador de dar ênfase ao que escreveu, quer seja ênfase com sentimento positivo ou negativo.

5. Desenvolvimento da Técnica de *Sentiment Analysis*

Para dar início à atribuição de polaridades aos *tweets* (já devidamente tratados) e porque os mesmos estão escritos na Língua Portuguesa, devido ao tema em análise, e os *lexicons* disponíveis são maioritariamente ingleses, foi necessário proceder à tradução das palavras disponíveis num destes dicionários. Assim sendo, o *discriminatory-word lexicon* - *Opinion Lexicon for English* de [Liu & Hu 2004], constituído por cerca de 6000 palavras utilizadas com frequência *online* e dividido num conjunto de palavras com sentimento negativo e outro conjunto com sentimento positivo, foi traduzido automaticamente para português recorrendo ao *Google Translate* [Google 2015]. Esta verificou-se ser a melhor opção devido ao elevado número de palavras e ao tempo necessário à tradução manual das mesmas.

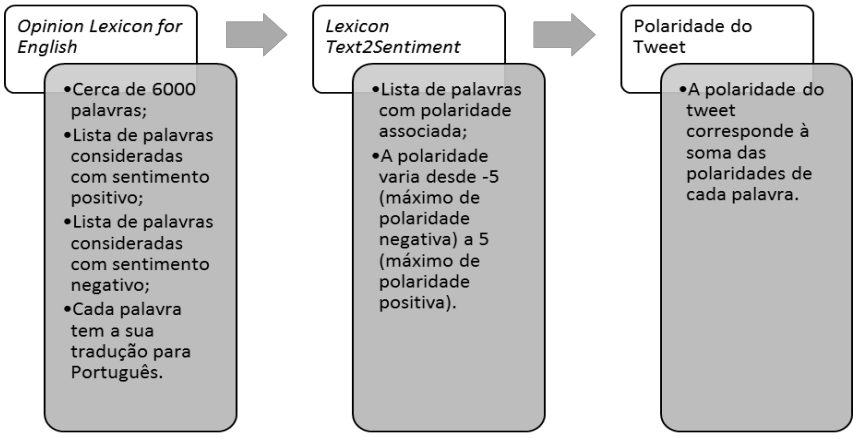


Figura 4 - Características dos *Lexicons* utilizados

Dada a lista de palavras e o sentimento associado às mesmas, o passo seguinte passa por atribuir uma polaridade a cada palavra dos *tweets* representando assim, o sentimento associado a cada uma delas. Posteriormente, e com base na polaridade das palavras, é calculada a polaridade do *tweet*. Este processo é sistematizado na Figura 4, que apresenta as principais características dos *lexicons* utilizados e auxilia a explicação do processo de atribuição de polaridade aos *tweets* (codificado em *Java*):

1. Verifica-se palavra a palavra presente no *tweet* se a mesma se encontra na lista de palavras portuguesas do *Opinion Lexicon for English*;
2. Se a palavra se encontrar na lista de palavras negativas é-lhe atribuída a polaridade -1, caso contrário, é-lhe atribuída a polaridade 1; É contabilizado o facto de ter sido encontrada a palavra (quer seja na lista de palavras negativas ou positivas);
3. A palavra correspondente em Inglês é utilizada para verificar se esta se encontra no *lexicon Text2Sentiment* de [Warden 2011]. Caso esteja presente é retornada a polaridade associada à mesma (quer seja negativa ou positiva); As polaridades encontram-se entre os valores -5 (máximo de negatividade) e o 5 (máximo de positividade);
4. A polaridade do *tweet* corresponde à soma das polaridades individuais de cada palavra retornada do *lexicon Text2Sentiment*;
5. Caso não tenha sido retornada nenhuma polaridade deste *lexicon* mas tenham sido encontradas palavras na lista de negativas ou positivas do *Opinion Lexicon for English*, são somadas essas polaridades atribuídas à palavra na verificação do *lexicon* (ponto 2), sendo esse o valor da polaridade do *tweet*.
6. De forma a manter a coerência entre as polaridades das palavras e a total dos *tweets*, sempre que este último valor ultrapassa os limites do *lexicon Text2Sentiment* (-5 e 5), o mesmo é substituído por esses limites.

Por forma a melhorar a performance de classificação do sentimento de cada *tweet* verificou-se que era essencial, tendo em conta que se fala de dados de redes sociais, incluir também *emoticons* para associar polaridade aos *tweets*, uma vez que são frequentemente utilizados para expressar sentimentos. Desta forma, foram utilizados os vinte e cinco *emoticons* mais usados no *Twitter* segundo [Berry 2012], sendo atribuída uma polaridade de 1 valor aos positivos e -1 aos *emoticons* negativos.

Para além desses, e tendo em conta que os termos de pesquisa estão relacionados com assuntos muito variados, achou-se relevante associar a cada termo de pesquisa uma polaridade e assim incluí-los no *lexicon*. Devido à falta de termos relacionados com os mesmos no *Opinion Lexicon Text2Sentiment*, algumas das palavras foram classificadas com uma polaridade tendo em conta a conotação do seu significado. A Tabela 4 apresenta a polaridade atribuída a cada um dos termos e a respetiva justificação para o valor atribuído sendo que, mais à frente, será avaliado o impacto da atribuição destes valores a cada uma das palavras.

Termo	Polaridade	Fundamentação
banco	-1	Atual conotação negativa da palavra em Portugal;
basqueiro	-1	Palavra "*noise*" presente na lista de negativas no *Opinion Lexicon for English;*
cibervadiagem	-1	Conotação negativa associada ao significado da palavra;
corrupção	-3	Conotação negativa associada à palavra; Palavra "*corruption*" presente na lista de negativas no *Opinion Lexicon for English;*
ébola	-3	Conotação muito negativa associada à doença;
gamificação	1	Conotação positiva associada ao significado da palavra;
jihadismo	-5	Associada a terrorismo revela-se a palavra com conotação mais negativa;
legionela	-2	Conotação negativa associada à doença;
selfie	1	Conotação positiva associada ao significado da palavra;
xurdir	2	Significado relacionado com: trabalho exaustivo, "lutar pela vida"; Tendo em conta a presença da expressão "*not working*" com polaridade -2 no *lexicon Text2Sentiment*, "xurdir" foi classificado com +2.

Tabela 4 - Definição de Polaridades para os Termos de Pesquisa

Os *tweets* foram então classificados com uma polaridade tendo sido criados os seguintes campos para cada *tweet*:

- "*Polarity*" = Total de polaridades negativa (valor negativo) + total de polaridades positiva (valor positivo);
- "PosNegCount" = Total de palavras positivas – total de palavras negativas.

6. Avaliação da Técnica de *Sentiment Analysis* Implementada

Depois de classificados os *tweets* e de forma a validar a classificação dos mesmos, foram escolhidos aleatoriamente de entre os 185.936, vinte *tweets* que foram fornecidos a 4 pessoas[2] sendo pedido que cada uma delas classificasse cada *tweet* como positivo (1 a 5), negativo (-1 a -5) ou neutro (0). Produzidas as médias das 4 opiniões pessoais dos colaboradores e comparadas com a classificação pelos *lexicons*, concluiu-se que este método se revelou adequado na classificação dos *tweets* como positivos, negativos ou neutros numa ordem de 75%, tal como se pode observar na Tabela 5: a vermelho os *tweets* em que não coincide a classificação humana com a classificação segundo os *lexicons* e a verde os *tweets* que coincidem.

[2] Colegas de formação com atuação em diferentes áreas de estudo.

Tweet	**Média Polaridades Colaboradores**	**Polaridade *Lexicons***
O MARXISTA FACHIN NO SUPREMO: Alô, senadores tucanos, os brasileiros já estão com o saco cheio de traidores!	-2	-4
Piscinas de Arouca encerradas devido a legionela de diminuta perigosidade	-3	-4
user Depois eu sentei no lugar dela e a migs e a elvira sentaram num banco juntas	0	-1
user Libéria está livre da epidemia de ebola, diz OMS	4	-2
user Partiu selfie com a urna.	-1	1
user essa é a verdadeira selfie que a hayley postou no tumblr e depois apagou	0	0
Odieo. Ebola	-4	-3
O que me leva a xurdir é o medo de regredir.	0	0
O basqueiro que tao a fazer em Aveiro. Ide pra casa ..	-1	-1
ACABOU O ENCANTO! Este é um sinal CLARO de q movimentos contra corrupção estão tendo sucesso	3	4
Banco de Portugal aperta vigilância	2	0
Cadê a expulsão d membros d PT envolvidos em corrupção, ou será mais uma mentira d MOLUSCO? VemPraRua17Maio	-3	-3
Dólares deixam o país na parcial de maio, mostra Banco Central: US$ 796 milhões deixaram a economia no acumula...	0	-1
Esquema de corrupção na Fifa envolvia direitos de transmissão de TV | Portal Fórum	-4	-2
Gostei de um vídeo *user* Selfie por dia durante 365 dias (1 ano)!	2	3
Modernidade....	0	0
Núcleo Educar de Tecnologias Digitais da Unimontes promove oficinas gratuitas de Gamificação	2	1
O poder sobe a cabeça até no Jihadismo hahahaha	-2	-5
OdeioQuando vou a um restaurante e aquilo tem bue miúdos a fazer bue basqueiro. Pqp.	-3	-1
porque é que se voltou a falar da legionella	-1	-2

Tabela 5 - Comparação de Classificação dos *Tweets*

De forma a sustentar a importância do tratamento efetuado sobre os dados, procedeu-se a uma experiência que consistiu em atribuir polaridade aos *tweets* tratados e aos não tratados e com os resultados obtidos (presentes na Tabela 6), percebe-se que a subjetividade é maior nos dados que receberam o tratamento apresentado na secção 4 aumentando a perceção dos sentimentos associados aos *tweets*.

Dados Tratados	
Total *Tweets*	185.936
Total *Tweets* Positivos e Negativos	163.288
Total *Tweets* Neutros	22.648
Subjetividade = Positivos&Negativos/Neutros	**7,21**
Dados não tratados	
Total *Tweets*	253.305
Total *Tweets* Positivos e Negativos	213.984
Total *Tweets* Neutros	39.321
Subjetividade = Positivos&Negativos/Neutros	**5,44**

Tabela 6 - Subjetividade nos Dados

7. Análise de Dados

A análise de dados foi efetuada recorrendo ao *Microsoft Excel Power BI* tendo sido criado um conjunto de *dashboards* que auxiliam nas conclusões a tirar sobre o tema. A primeira análise que se destaca como necessária passa pela observação da coerência entre a polaridade ("*Polarity*") atribuída aos *tweets* e a relação da quantidade das palavras positivas e negativas presentes nos mesmos ("PosNegCount"). Se a polaridade atribuída a um *tweet* é negativa, de forma a ser coerente, o campo "PosNegCount" deve também ser negativo e vice-versa. Na Tabela 7 podemos constatar esse facto observando um subconjunto aleatório do *dataset*, evidenciando os *tweets* já classificados com a sua polaridade.

Polarity (total polaridades negativas + total polaridades positivas)	**PosNegCount** (total palavras positivas – total palavras negativas)
0 + 3 = 3	1 – 0 = 1
-1 + 4 = 3	4 – 1 = 3
-2 + 1 = -1	1 – 2 = -1
-4 + 2 = -2	1 – 4 = -3
-7 + 0 = -5 (substituição do -7 pelo mínimo do *lexicon* → -5)	0 – 5 = -5
-3 + 1 = -2	1 – 1 = 0

Tabela 7 - Exemplos de Cálculo da "*Polarity*" e "PosNegCount"

Apesar disso, como se observa na Figura 5, existem algumas exceções em que existindo uma polaridade negativa, o campo "PosNegCount" é positivo e vice-versa. Este facto, assinalado na figura (*) acontece em casos particulares onde são contabilizadas as palavras existentes no *Opinion Lexicon for English* mas que, porque é retornada pelo menos uma polaridade do *Opinion Lexicon Text2Sentiment*, não são acumuladas as polaridades (o valor -1 ou 1 mencionado na secção 5). Apesar disso, confirma-se na imagem que são reduzidos os casos em

que se verifica esta situação e por isso, revelou-se favorável a adoção de contabilização de palavras existentes no *Opinion Lexicon for English*.

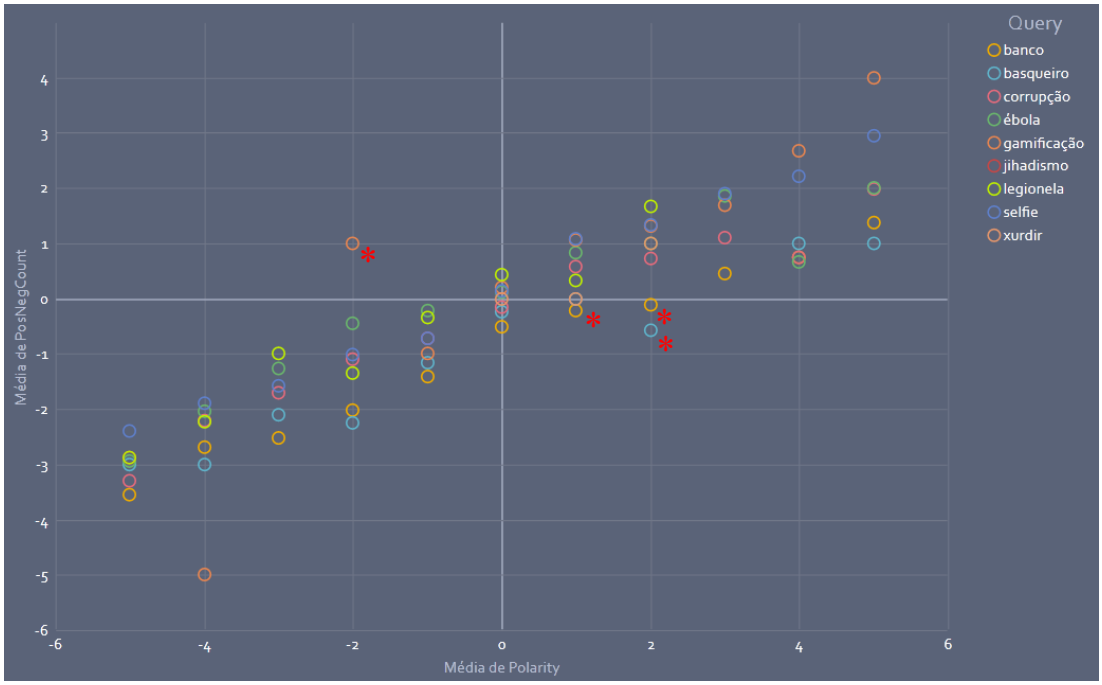


Figura 5 - Relação entre Polaridade e Contagem de Palavras do *Tweet* por Termo

No seguimento do que já foi exposto, é também interessante perceber os picos associados à polaridade e quantidades de palavras classificadas nos *tweets*. Esses mesmos picos (máximos e mínimos) podem ser observados na Figura 6 e na Figura 7, constatando-se que a polaridade e a relação das quantidades de palavras positivas e negativas frequentemente seguem o mesmo padrão destacando-se o termo "Jihadismo" com os seus máximos em valores negativos (Figura 6) e o termo "Xurdir" com os seus mínimos no valor zero (Figura 6).

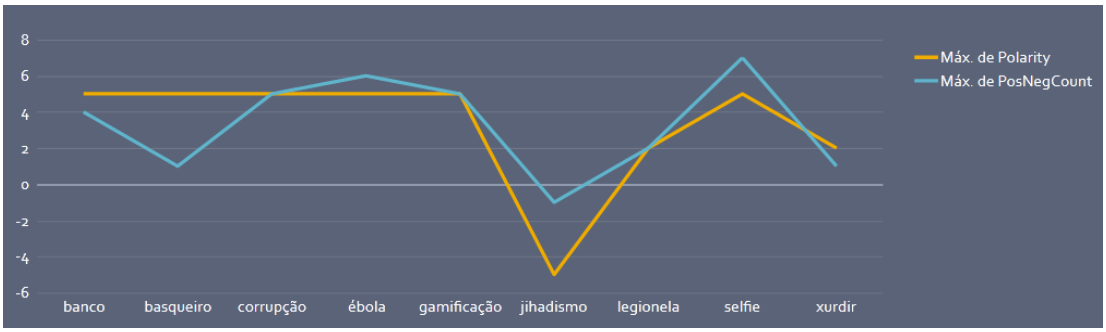


Figura 6 - Máximos de Polaridades e Quantidades de Palavras

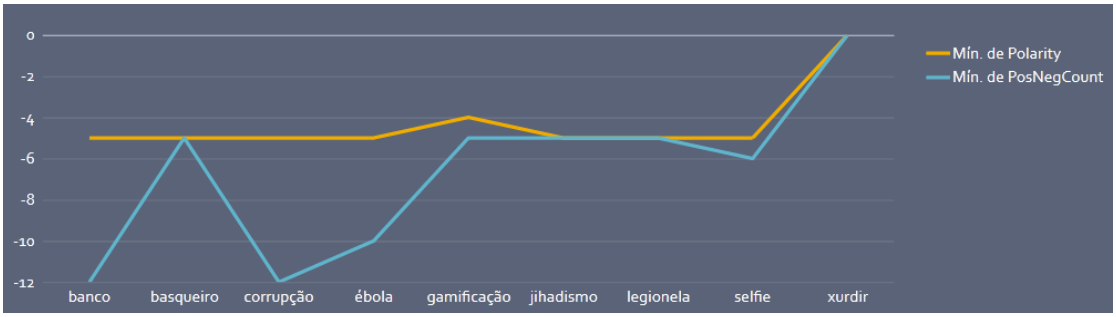


Figura 7 - Mínimos de Polaridades e Quantidades de Palavras

No que diz respeito aos atributos encontrados nos *tweets* e mencionados na secção 4, a Figura 8 e a Figura 9 apresentam uma visão geral dos mesmos quando associados a cada termo de pesquisa[3]. Assim sendo, na Figura 8 é possível observar a quantidade de *retweets* associados a cada termo e na Figura 9 a distribuição entre os termos de:

- *Tweets* que continham *Hashtags* (tipicamente conhecidos como propagadores de informação);
- *Tweets* classificados como notícias (continham um *URL* sendo considerados partilhas de notícias);
- *Tweets* que continham a identificação de outros utilizadores do *Twitter* (partilha de informação);
- *Tweets* completamente escritos em letras maiúsculas.

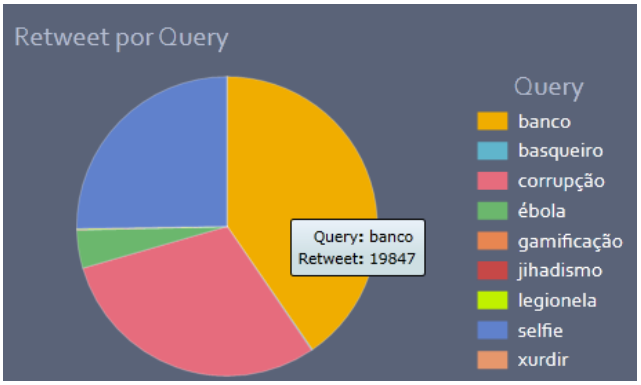


Figura 8 - Quantidade de *Retweets* por Termo

[3] Verificou-se não ter sido retornado nenhum *tweet* associado ao termo "Cibervadiagem" estando em análise as restantes 9 palavras.

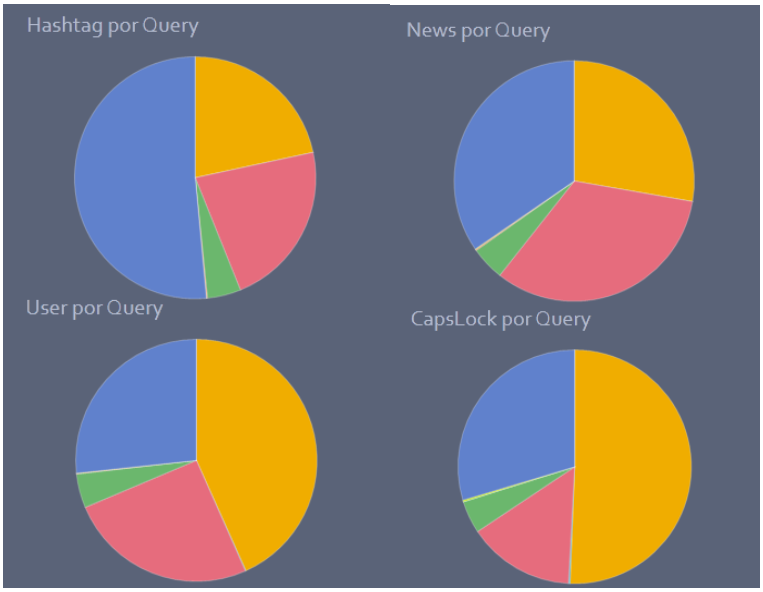


Figura 9 - Análise de *Tweets* por Termo

Estes dados mencionados podem por si só influenciar a popularidade dos termos em análise, isto é, além de se considerar o maior número de *tweets* recolhidos dentro dos dez termos pesquisados, poderá dar-se mais ou menos importância a *retweets* ou *tweets* com notícias ou até *tweets* que contêm *Hashtags* e por isso propagam informação. Todos estes pontos podem ser considerados para a "eleição" da palavra mais mencionada no *Twitter*.

Indo agora de encontro à questão principal: qual a palavra mais mencionada na rede social *Twitter* e que sentimento está associada à mesma, a Figura 10 representa no tamanho dos círculos a quantidade de *tweets* associada à palavra e no quadrante em que se encontra, o sentimento positivo ou negativo que lhe está inerente tendo em conta a média de todos os *tweets* existentes da palavra. Conclui-se, assim, que a palavra mais mencionada é "banco" com 86.269 do total de 185.936 *tweets* existentes no *dataset* e com um sentimento negativo associado. Segue-se a palavra "selfie" com mais de 48.000 registos e sentimento positivo associado e "corrupção" com pouco mais que 42.500 e uma conotação negativa.

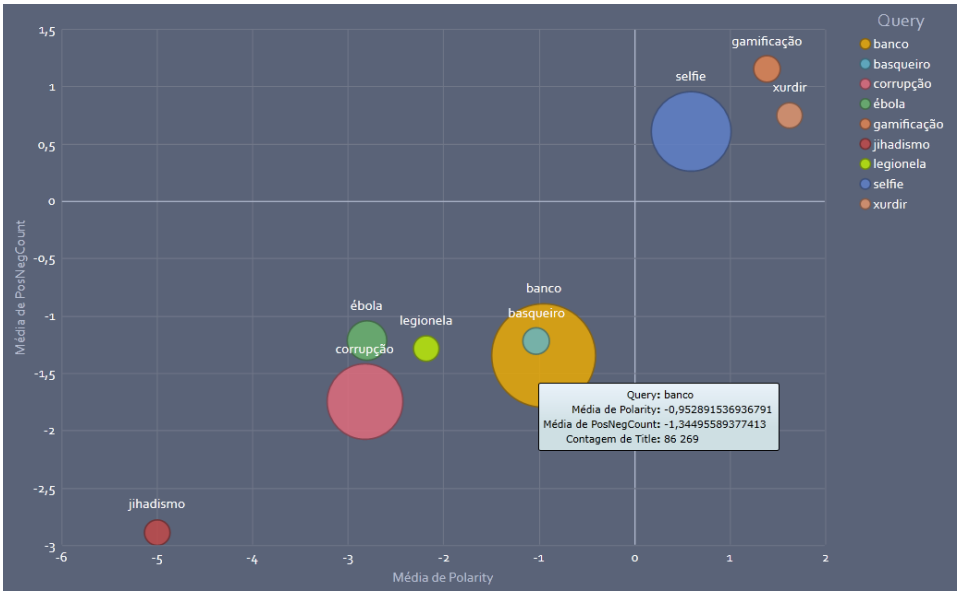


Figura 10 - Palavras mais Mencionadas e seu Sentimento

No que diz respeito à palavra "banco", esta também apresenta aspetos mencionados na Figura 8 e na Figura 9 a seu favor:

- O maior número de *retweets* entre as dez palavras em análise (19.847) – confirmando que é a palavra mais mencionada;
- O maior número de *tweets* com identificação de outros utilizadores (29.399);
- O terceiro maior número de notícias associadas aos *tweets* (20.894);
- O terceiro maior número de *tweets* com *Hashtags* (7.565);
- O maior número de *tweets* escritos completamente em letras maiúsculas (920).

Considerando que na Tabela 4 foram apresentadas as polaridades para cada um dos termos pesquisados, a Figura 11 apresenta os resultados obtidos quando se retiram os termos de pesquisa dos dicionários utilizados, isto é, quando não é associada nenhuma polaridade a cada um desses termos. Nesta situação, pode-se verificar que, retirando os termos dos dicionários ou atribuindo-lhes o valor neutro (zero), o sentimento atribuído a cada termo tende a ser o mesmo diminuindo apenas o enfase negativo ou positivo associado a cada um.

Assim sendo, constata-se que "jihadismo" continua a apresentar uma conotação muito negativa e que "banco", "corrupção" e "selfie" continuam a ser as palavras mais referidas tendo as duas primeiras uma conotação negativa. As palavras com diferenças significativas de sentimento em relação aos resultados apresentados anteriormente são "selfie" e "xurdir" devido à falta de semântica associada aos termos na língua portuguesa.

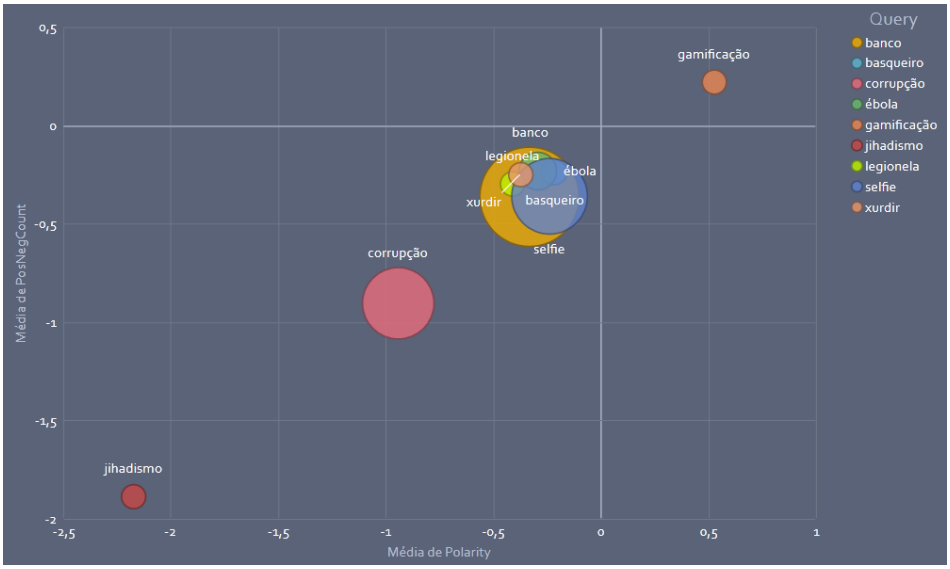


Figura 11 – Resultados Face à Inexistência dos Termos de Pesquisa nos *Lexicons*

8. Conclusões

Neste artigo foi apresentada uma alternativa para a identificação da palavra do ano, votação introduzida pela Porto Editora. A solução passa pela recolha de dados do *Twitter* tendo em conta os temas que vão surgindo na sociedade portuguesa e assim, utilizando as publicações partilhadas pelos utilizadores ao longo do ano, evita-se a necessidade de uma votação.

Os dados foram recolhidos recorrendo à ferramenta *KNIME* e seguidamente armazenados em ambiente *Hadoop* utilizando o *HBase*, adequado para este tipo de dados não estruturados e ajudando na eliminação de registos repetidos.

O tratamento de dados que foi efetuado sobre os *tweets* revelou-se mais tarde essencial para a classificação dos mesmos utilizando os dois dicionários de palavras. Esta técnica de atribuição de polaridade às palavras presentes no *tweet* foi avaliada recorrendo a quatro colaboradores que classificaram um conjunto de vinte *tweets*, escolhidos aleatoriamente do *dataset* já classificado. As médias das avaliações atribuídas aos *tweets* pelas quatro pessoas foram comparadas com as polaridades atribuídas recorrendo aos dois *lexicons*, sendo verificado que a classificação dos *tweets* foi adequada na ordem dos 75%.

Partindo então desses dados, recorrendo ao *Microsoft Excel Power BI*, foram elaborados conjuntos de análises e concluiu-se sobre a principal questão: qual a palavra mais mencionada e que sentimento lhe está associado? A resposta a esta questão é a palavra "banco" como sendo a mais mencionada e tendo uma conotação negativa associada aos seus *tweets*.

Como trabalho futuro, verifica-se a necessidade de incluir no componente de *Sentiment Analysis* o estudo de expressões contraditórias visto que atualmente este ponto não foi abordado e

verificou-se ser uma das situações presentes na Tabela 5 para o erro de classificações de *tweets* em comparação com a classificação humana. Por outro lado, seria interessante também incluir a utilização de um dicionário de abreviaturas, tendo em conta que estas são frequentemente utilizadas nas redes sociais devido à limitação de caracteres nas publicações. Para além disso, sendo aspetos que dão seguimento ao trabalho já realizado deve garantir-se, em primeiro lugar, que não são utilizadas palavras pré-selecionadas para a recolha de dados da rede social mas, pelo contrário, é elaborada uma recolha gradual, ao longo do ano, sobre variados temas que vão surgindo no país permitindo assim, dependendo do período do ano, recolher informação sobre diferentes assuntos. Em segundo lugar, apesar de mencionada a importância da contagem de *retweets* e *hashtags* associadas aos *tweets*, os mesmos valores não foram considerados para a identificação da palavra do ano, tendo estes valores apenas sido considerados na análise dos dados, facto que se deve alterar.

9. Referências

Asur, S. e Huberman, B. A., Predicting the Future with Social Media, In 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) (Vol. 1, pp. 492–499), 2010.

Berry, N., Emoticon Analysis in Twitter [Blog], http://www.datagenetics.com/blog/october52012/index.html, (Maio de 2015), 2012.

Cloudera, http://www.Cloudera.com, (Maio de 2015), 2015.

Gebremeskel, G., Sentiment Analysis of Twitter posts about news, Masters Thesis - University of Malta, 2011.

Google, Google Translate, https://translate.google.pt/, (Maio de 2015), 2015.

Infopédia, Porto Editora, http://www.infopedia.pt/palavra-do-ano/, (Maio de 2015), 2015.

KNIME, https://www.KNIME.org/, (Maio de 2015), 2015.

Kouloumpis, E., Wilson, T. e Moore, J., Twitter Sentiment Analysis: The Good the Bad and the OMG!, Presented at the Fifth International AAAI Conference on Weblogs and Social Media, 2011.

Kumari, P., Singh, S., More, D., Talpade, D., e Pathak, M., Sentiment Analysis of Tweets, International Journal of Science Technology & Engineering (Volume-1, Issue-10), 2015.

Liu, B., Sentiment Analysis and Opinion Mining, Synthesis Lectures on Human Language Technologies, 5(1), 1–167, 2012.

Liu, B., e Hu, M., Opinion Mining, Sentiment Analysis, and Opinion Spam Detection, 2004.

North, M., Data Mining for the Masses, 2012.

Pang, B., e Lee, L., Opinion Mining and Sentiment Analysis. Found. Trends Inf. Retr., 2(1-2), 1–135, 2008.

Talend, Talend Open Studio for Big Data, https://www.Talend.com/products/big-data/, (Maio de 2015), 2015.

Warden, P., Text2sentiment Words, https://github.com/petewarden/dstk/blob/master/text2sentiment.rb, (Maio de 2015), 2011.

Os perfis de infraestrutura e a partilha do conhecimento CAPSI/2015

Mírian Oliveira [1], Antonio C. G. Maçada [2], Carla Curado [3], Felipe Nodari [1]

1) Pontifícia Universidade Católica do Rio Grande do Sul, Brasil
miriano@pucrs.br

2) Universidade Federal do Rio Grande do Sul, Brasil
acgmacada@ea.ufrgs.br

3) Universidade de Lisboa, Portugal
ccurado@iseg.utl.pt

1) Pontifícia Universidade Católica do Rio Grande do Sul, Brasil
fnodari@gmail.com

Resumo

A tecnologia da informação e o suporte da gestão de topo são considerados como infraestrutura para a partilha do conhecimento. Esta investigação tem como objetivo identificar os perfis de infraestrutura e a relação com a partilha de conhecimento. O método adotado foi um inquérito, os 231 respondentes trabalham em equipas de tecnologia da informação no Brasil. Na análise foram identificados quatro perfis (foco na tecnologia da informação, foco no suporte da gestão de topo, à deriva e orientado à partilha do conhecimento), sendo que o mais efetivo é o perfil orientado à partilha do conhecimento, não havendo diferença significativa entre os demais clusters em relação a partilha do conhecimento.

Palavras chave: tecnologia da informação, suporte da gestão de topo, partilha de conhecimento, infraestrutura

1. Introdução

O conhecimento é um ativo intangível relevante para a obtenção de vantagem competitiva sustentável [Wang e Noe 2010; Tan e Wong 2015]. A gestão do conhecimento tem o intuito de contribuir para que a empresa atinja seus objetivos de negócio [Lee e Yang 2000]. A partilha do conhecimento (PC) é um dos processos mais relevantes para a gestão do conhecimento [Yu *et al*. 2010; Iskoujina e Robberts 2015], e pode ser entendida como o grau em que os indivíduos doam ou coletam conhecimento tácito ou explícito na equipa [Huang 2009]. A partilha do conhecimento é importante para evitar a perda do conhecimento quando um funcionário deixa a

empresa, assim como para aumentar a produtividade e a inovação [Xue *et al*. 2009; Liao *et al*. 2007].

A tecnologia da informação (TI) e o suporte da gestão de topo (SGT) são antecedentes da partilha do conhecimento [Lee *et al*. 2012; Chennamaneni *et al*. 2012], ou seja, são parte da infraestrutura necessária para que a partilha do conhecimento ocorra [Lee *et al*. 2012]. A TI permite que o conhecimento possa ser partilhado independentemente do local, os funcionários podem estar ou não dispersos geograficamente, e do tempo, podendo a partilha do conhecimento ser síncrona ou assíncrona [Choi *et al*. 2008]. Enquanto a TI pode ser associada à forma como a partilha ocorre, o suporte da gestão de topo pode ser relacionado com o motivo pelo qual os indivíduos partilham. A gestão de topo influencia o comportamento dos funcionários, sinalizando as prioridades, tanto através de exemplos quanto financiando as condições para que a partilha do conhecimento ocorra [Lin 2007].

As investigações até o momento destacam a relevância da TI e do SGT para a partilha do conhecimento. No entanto, não deixam clara a relação entre os perfis de infraestrutura e a ocorrência da partilha do conhecimento. Para preencher esta lacuna, esta investigação tem como objetivo analisar a relação entre os diferentes perfis de infraestrutura (TI e SGT) e a partilha do conhecimento.

Este artigo apresenta a seguir: a revisão da literatura sobre tecnologia da informação e suporte da gestão de topo como infraestrutura para a partilha do conhecimento (seção 2); os procedimentos metodológicos adotados nesta investigação (seção 3); a análise e discussão dos resultados (seção 4); as conclusões, limitações e sugestões de estudos futuros (seção 5).

2. Infraestrutura para a partilha do conhecimento

Tanto o conhecimento explícito, que pode ser documentado, quanto o conhecimento tácito, que reside na mente dos indivíduos devem ser considerados na partilha do conhecimento [Johannessen *et al*. 2001]. A partilha do conhecimento contribui para que o funcionário possua o conhecimento que necessita para as suas atividades, assim como para a retenção do conhecimento na empresa. A TI é considerada um suporte para a partilha do conhecimento, principalmente quando existir dispersão geográfica dos funcionários e um grande volume de conhecimento disponível. A TI permite que o conhecimento seja partilhado "em qualquer momento, de qualquer lugar, de um modo securo e amigável" [Lee 2006 p. 58]. As soluções WEB 2.0 têm obtido destaque pela facilidade de uso, potencial para colaboração, entre outros aspectos [Grace 2009].

A TI pode facilitar a partilha do conhecimento verticalmente e horizontalmente nas empresas

[Yeh 2006], permitindo o armazenamento, acesso e recuperação do conhecimento, assim como o contato entre os indivíduos [Alavi e Leidner 2001; Yeh *et al*. 2006]. Isto faz com que a TI seja relevante tanto para a partilha do conhecimento explícito quanto para a do conhecimento tácito [Choi *et al*. 2008]. Tecnologias de colaboração estão associadas à estratégia de personalização (foco no conhecimento tácito), enquanto os repositórios estão relacionados com a estratégia de codificação (foco no conhecimento explícito) [Saito *et al*. 2007].

A TI pode contribuir e em alguns casos ser mesmo necessária para a partilha do conhecimento, no entanto, não é suficiente para que ela ocorra [Choi *et al*. 2008], pois a sua existência não garante o seu uso pelos indivíduos [Chennamaneni *et al*. 2012]. A gestão de topo tem um papel relevante no sucesso de qualquer atividade numa empresa [Tan e Wong 2015], por exemplo, a gestão de topo pode incentivar os indivíduos a partilharem o conhecimento [Ragab e Arisha 2013; Lin e Tseng 2005; Cavaliere e Lombardi 2015;].

O comprometimento da gestão de topo com a partilha do conhecimento inflluencia o comportamento dos funcionários em relação a este processo [Chennamaneni *et al*. 2012]. Lin [2007] identificou que o SGT à partilha do conhecimento influencia positivamente a inclinação do funcionário a partilhar o seu conhecimento. De acordo com Ruggles [1998], uma barreira para a partilha do conhecimento é a falta de sinalização da gestão de topo sobre a importância dada a este processo. Lee [2006] verificou que o SGT influencia tanto a quantidade quanto a qualidade da partilha do conhecimento.

A gestão de topo facilitar a interação social e promover atividades para a partilha do conhecimento é mais relevante do que motivações extrínsecas como recompensas monetárias [Lin 2007]. Lee, Kim e Kim [2012, p. 200] consideram que o suporte da gestão de topo pode ter menor importância quando a tecnologia da informação estiver desenvolvida, pois para estes autores, “a TI é o principal componente da infraestrutura da gestão do conhecimento.”.

O suporte da gestão de topo para o partilha do conhecimento é fundamental, uma vez que os funcionários se interessam em agir de acordo com as diretrizes dos gestores [Connelly e Kelloway 2003]. Chennamaneni *et al*. [2012] identificaram que os funcionários se sentem mais inclinados a partilharem conhecimento quando percebem o suporte da gestão de topo à partilha do conhecimento. Os gestores de topo podem suportar a partilha de conhecimento convencendo os funcionários sobre a sua importância [Lee e Wong 2015].

O suporte da gestão de topo, além de influenciar o posicionamento dos funcionários perante os processos de gestão do conhecimento [Chennamaneni *et al*. 2012] também molda as ações estratégicas [Lee *et al*. 2012], as quais incluem o desenvolvimento da TI voltada à partilha de

conhecimento. Assim, pode-se supor que existam perfis distintos de empresas, no que tange ao SGT e à TI, com diferentes níveis de comportamento de partilha do conhecimento.

3. Método

O método adotado nesta investigação foi o inquérito do tipo corte-transversal. O instrumento utilizado para coleta de dados está dividido em duas partes: itens dos três construtos; e, caracterização dos respondentes. Os itens dos construtos foram medidos por meio de uma escala Likert, onde as opções variam de 1 (discordo fortemente) até 7 (concordo totalmente). Os três construtos (tecnologia da informação, suporte da gestão de topo e partilha do conhecimento) foram operacionalizados com escalas previamente validadas.

Suporte da gestão de topo foi medido aplicando a escala utilizada por Lin [2007], a qual foi originalmente desenvolvida por Tan e Zhao [2003]. Tecnologia da informação foi medida através da escala utilizada por Lin [2007], a qual foi originalmente desenvolvida por Lee e Choi [2003]. A escala adotada por Huang [2009] foi utilizada para medir a partilha do conhecimento, e foi originalmente proposta por Bock *et al*. [2005].

A tradução reversa, a validação de conteúdo e a validação de face foram os procedimentos adotados para aumentar a qualidade da investigação. A versão final dos itens utilizados na coleta de dados no Brasil é apresentada na Tabela 1.

Construto	Itens
PC	PC1. Eu compartilho frequentemente os relatórios e os documentos oficiais do meu trabalho com os membros da minha equipe. PC2. Eu sempre compartilho meus manuais, metodologias e modelos com os membros da minha equipe. PC3. Eu frequentemente compartilho minha experiência ou know-how (conhecimento sobre como fazer algo) de trabalho com os membros da minha equipe. PC4. Eu sempre compartilho meu *know-where* (conhecimento sobre onde está armazenado determinado conhecimento) ou *know-whom* (conhecimento sobre quem possui determinado conhecimento) quando solicitado pelos membros da minha equipe.
TI	TI1. Minha equipe usa armazenamento eletrônico para acessar conhecimento. TI2. Minha equipe usa redes de conhecimento (por exemplo, comunidades virtuais, *groupware*) para se comunicar. TI3. A empresa usa tecnologia que permite que minha equipe compartilhe conhecimento com pessoas de outras empresas. TI4. Minha empresa usa tecnologia que permite que minha equipe compartilhe conhecimento com outras pessoas da minha empresa
SGT	SGT1. A alta administração acredita que estimular a partilha de conhecimento entre membros da equipe é benéfico.

	SGT2. A alta administração sempre suporta e encoraja os funcionários a compartilharem seus conhecimentos com os outros membros da equipe. SGT3. A alta administração fornece a maioria do auxílio e dos recursos necessários para que os funcionários compartilhem conhecimento com os outros membros da equipe. SGT4. A alta administração é empolgada para ver que os funcionários estão felizes em compartilhar seus conhecimentos com os membros de sua equipe.

Tabela 1 – Construtos e itens

Uma mensagem com o convite para participação na investigação foi enviada para os respondentes. Nesta mensagem foi apresentado o objetivo da investigação, o link e a senha para acessar o instrumento, o qual foi disponibilizado através do Qualtrics®.

Os 231 respondentes são funcionários que trabalham em equipas de tecnologia da informação no Brasil, selecionados por conveniência. A idade média dos respondentes é de 33,1 anos; o tempo médio de experiência é de 12,8 anos; 74,5 % são homens; e, 78,4 % possuem no mínimo graduação completa. Os dados foram analisados através da Análise Fatorial Exploratória, Análise Descritiva e Análise de Cluster, com o suporte do SPSS 17.0®.

4. Análise e discussão dos resultados

A Análise Fatorial Exploratória utilizou a Análise de Componentes Principais com rotação Varimax [Hair *et al*. 2005]: 1) a carga fatorial de cada item foi superior ou igual ao valor recomendado de 0,7; 2) a variância explicada pelos fatores foi de 68,4 %, o que está acima do valor recomendado de 60%. A média para os itens dos construtos é apresentada na Tabela 2, considerando os 231 respondentes.

Construto	Média	Desvio Padrão	Alfa de Cronbach
PC	5.81	1.03	0.83
PC1	5.39	1.50	
PC2	5.60	1.46	
PC3	5.96	1.15	
PC4	6.29	1.03	
TI	4.97	1.41	0.76
TI1	5.89	1.37	
TI2	4.85	2.02	
TI3	4.04	2.07	
TI4	5.09	1.85	
SGT	5.16	1.50	0.92

SGT1	5.66	1.59	
SGT2	5.41	1.64	
SGT3	4.80	1.73	
SGT4	4.79	1.75	

Tabela 2 – Análise descritiva e confiabilidade

Ao observar a diferença nas médias dos itens do construto partilha do conhecimento, percebe-se que aquelas associadas aos itens que tem foco no conhecimento tácito são maiores, o que sugere que a partilha do conhecimento é mais positiva em relação ao conhecimento tácito do que em relação ao conhecimento explícito. Isto pode representar uma opção estratégica das empresas, de priorizar a personalização em relação à codificação, ou uma maior dificuldade de adoção de práticas associadas à gestão do conhecimento explícito. Neste caso, o conhecimento está mais protegido, considerando que o conhecimento codificado pode mais facilmente ser copiado pela concorrência [Chilton e Bloodgood 2008]. Por outro lado, o conhecimento está mais associado aos indivíduos do que à empresa, e a saída de funcionários da organização pode levar a perda do conhecimento [Ragab e Arisha 2013].

A média do construto partilha do conhecimento é superior à média do construto tecnologia da informação. Isto pode indicar que parte da partilha do conhecimento não está associada ao uso de tecnologia da informação, ocorrendo através de interações presenciais, o que pode ser parte da explicação para a personalização ter sido privilegiada em relação à codificação. Isto porque a codificação é mais dependente da tecnologia da informação do que a personalização. Embora ambas as estratégias possam se beneficiar do uso da tecnologia da informação [Ragab e Arisha 2013; Saito *et al*. 2007].

Não existe um conjunto de tecnologias que seja mais apropriada para a partilha do conhecimento. No entanto, é preciso contemplar tecnologias que privilegiam a personalização e a codificação, pois é necessário contemplar os dois tipos de estratégias para que o desempenho organizacional tenha melhor resultado, como identificado por Choi e Lee [2003].

Em relação aos itens do construto tecnologia da informação, verifica-se que a menor média está associada à partilha do conhecimento com o ambiente externo. Em longo prazo, isto pode não ser suficiente para alavancar a inovação, uma vez que o conhecimento interno pode tornar-se homogêneo com o tempo.

As médias dos itens do construto suporte da gestão de topo indicam que: a gestão de topo considera que a partilha do conhecimento agrega valor; a percepção sobre o aporte de recursos para a partilha do conhecimento é inferior à percepção de encorajamento para a partilha do

conhecimento. De acordo com Lin [2007], o suporte da gestão de topo se manifesta tanto pelo encorajamento quanto pelo aporte de recursos. Isto evidencia que o suporte da gestão de topo deve considerar aspectos humanos, viabilizando um clima propício para a partilha do conhecimento assim como aspectos tecnológicos, onde o investimento de recursos é necessário. O aporte de recursos pode estar relacionado com a adoção de tecnologia da informação para o partilha do conhecimento, assim como reconhecer que para a partilha do conhecimento ocorrer é necessário que os funcionários dediquem parte do seu tempo para atividades de codificação do conhecimento, assim como para interação com indivíduos dentro e fora da organização. No caso destes respondentes, a existência de espaço para incrementar a adoção de tecnologia da informação parece estar alinhada com a percepção de que existe menor aporte de recursos do que reconhecimento do valor que o partilha do conhecimento pode trazer.

A análise de cluster foi utilizada para identificar os perfis de infraestrutura (TI e SGT) para a partilha do conhecimento. Os respondentes foram agrupados pela técnica não hierárquica K-means. Os quatro clusters são (Tabela 3): Cluster 1 – foco na TI (51 respondentes); Cluster 2 – foco no SGT (53 respondentes); Cluster 3 – à deriva (25 respondentes); Cluster 4 – orientado à partilha do conhecimento (102 respondentes).

	Foco na TI	Foco no SGT	À deriva	Orientado à PC
TI	5,15	3,76	2,71	6,06
SGT	3,97	5,36	2,52	6,31
PC	5,44	5,69	5,32	6,18

Tabela 3 – Clusters

O método Duncan [Choi e Lee 2003] foi utilizado, sendo os resultados apresentados na Tabela 4, o que indica que o cluster "Orientado à PC" é o mais efetivo. A diferença entre os outros três clusters ("foco na TI", "foco no SGT" e "à deriva") não é significativa.

Perfis	Número de casos	1	2
Foco na TI	51	5.44	
Foco no SGT	53	5.69	
À deriva	25	5.32	
Orientado à PC	102		6.18

Tabela 4 – Teste de comparação entre os clusters

A análise de cluster identificou quatro grupos de respondentes considerando a TI e o SGT. Apenas o cluster em que a PC é significativamente diferente, apresenta valores maiores tanto para a tecnologia da informação quanto para o suporte da gestão de topo. Isto sugere que o suporte da gestão de topo e a tecnologia da informação são pré-requisitos para o comportamento da partilha do conhecimento, ou seja, os dois precisam estar presentes para que a PC se diferencie das demais situações. A TI não facilitará a partilha se a gestão de topo não incentivar esta prática e, da mesma forma, somente o SGT não influenciará a partilha do conhecimento, se as ferramentas de TI não a facilitarem. Este resultado não está alinhado com Lee, Kim e Kim [2012], pois o cluster "foco na TI" não possui melhor resultado para a partilha do conhecimento do que o cluster "foco no SGT".

5. Conclusões

Este estudo avaliou empiricamente os perfis de infraestrutura para a partilha do conhecimento, considerando a tecnologia da informação e o suporte da gestão de topo. Isto contribui com os gestores que queiram alavancar a partilha do conhecimento em suas empresas, e com a academia ao identificar que a partilha de conhecimento apresenta melhores resultados somente quando os dois componentes da infraestrutura (tecnologia da informação e suporte da gestão de topo) são considerados.

Esta investigação identificou quatro perfis de infraestrutura relacionados com a partilha de conhecimento (foco na TI, foco no SGT, à deriva e orientada à PC), como mostra a Figura 1. Testes empíricos sugerem que a partilha do conhecimento somente é diferente e com melhor resultado no cluster "orientada à PC". Isto evidencia que somente a disponibilidade da TI ou o SGT não é suficiente para que o funcionário partilhe seu conhecimento. No caso do perfil Foco na TI, o investimento realizado pode ter um custo alto para a empresa sem gerar o retorno esperado, ou seja, a partilha do conhecimento. Por outro lado, a sinalização da gestão de topo de que a partilha do conhecimento é prioridade para a empresa necessita vir acompanhada do investimento em condições para que isto ocorra.

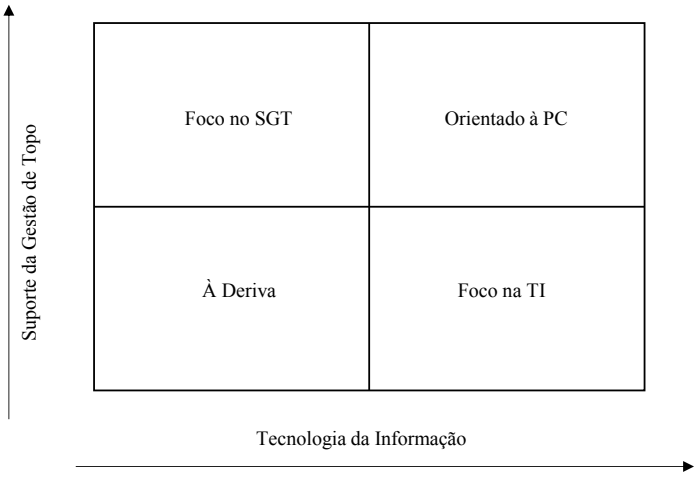


Figura 1 – Perfis de infraestrutura de PC

Esta investigação confirma a importância do suporte da gestão de topo e da tecnologia da informação para a partilha do conhecimento, como previamente identificado na literatura. No entanto, agrega um novo conhecimento ao verificar que o incremento na partilha do conhecimento somente ocorre na presença dos dois fatores da infraestrutura. A tecnologia da informação e o suporte da gestão de topo possuem um papel de suporte em relação à partilha do conhecimento, e devem ser considerados em conjunto. Além disto, considera-se uma contribuição desta investigação, o fato de que o cluster denominado "à deriva", porque nenhum dos dois fatores foi observado, teve efeito semelhante em relação à partilha do conhecimento aos clusters onde somente um dos fatores foi identificado.

Como investigação futura, sugere-se considerar equipas de outros setores, pois esta investigação tratou apenas com equipas de TI. Outro aspecto a ser investigado é se a dispersão geográfica dos componentes das equipas pode ter efeito na relação estudada.

6. Referências

Alavi, M. e Leidner, D.E., "Review: Knowledge management and knowledge management systems: conceptual foundations and research issues", MIS Quarterly, 25(1), 2001, pp. 107-136.

Beckett, A.J., Wainwright, C.E.R. and Bance, D. "Knowledge management: strategy or software?", Management Decision, 38(9), 2000, pp. 601-606.

Bock, G.-W., Zmud, R.W., Kim, Y.-G. e Lee, J.N. "Behavioral intention formation in knowledge sharing: examining the roles of extrinsic motivators, social-psychological forces, and organizational climate", MIS Quarterly, 29(1), 2005, pp. 87-111.

Cavaliere, V. e Lombardi, S. "Exploring different cultural configurations: how do they affect subsidiaries' knowledge sharing behaviors", Journal of Knowledge Management, 19 (2), 2015, pp. 141-163.

Chennamaneni, A., Teng, J.T.C. and Raja, M.K. "A unified model of knowledge sharing behaviours: theoretical development and empirical test", Behaviour and information technology, 31(11), 2012, pp. 1097-1115.

Chilton, M. A. e Bloodgood, J. M. "The dimensions of tacit and explicit knowledge: a description and measure", International Journal of Knowledge Management, 4(2), 2008, pp. 75-91.

Choi, B e Lee, H. "An empirical investigation of KM styles and their effect on corporate performance", Information & Management, 40(5), 2003, pp. 403-417.

Choi, S.Y., Kang, Y.S. e Lee, H. "The effects of socio-technical enablers on knowledge sharing: an exploratory examination", Journal of Information Science, 34(5), 2008, pp. 742-754.

Connelly, C.E. e Kelloway, E.K. "Predictors of employees' perceptions of knowledge sharing culture", Leadership & Organization Development Journal, 24(5), 2003, pp. 294-301.

Grace, T. "Wikis as a knowledge management tool", Journal of Knowledge Management, 13(4), 2009, pp. 64-74.

Hair, J. F., Anderson, R. E., Tatham, R. L., e Black, W. C. 2005. Multivariate data analysis, Upper Saddle River: Prentice-Hall.

Huang, C.-C. "Knowledge sharing and group cohesiveness on performance: an empirical study of technology R&D teams in Taiwan", Technovation, 29(11), 2009, pp. 786-797.

Johannessen, J., Olaisen, J. e Olsen, B. 2001. "Mismanagement of tacit knowledge: the importance of tacit knowledge, the danger of information technology, and what to do about it", International Journal of Information Management, 21(1), 2001, pp. 3-20.

Iskoujina, Z. e Robberts, J. "Knowledge sharing in open source software communities: motivations and management", Journal of Knowledge Management, 19 (4), 2015, pp. 791-813.

Lee, H. e Choi, B. "Knowledge management enablers, processes and organizational performance: an integrative view and empirical examination", Journal of Management Information Systems, 20(1), 2003, pp. 179-228.

Lee, J.-H. "Effects of managerial drivers and climate maturity on knowledge management performance: empirical validation", Information Resources Management Journal, 19(3), 2006, pp. 48-60.

Lee, S., Kim, B.G. and Kim, H. "An integrated view of knowledge management for performance", Journal of Knowledge Management, 16(2), 2012, pp. 183-203.

Lee, C.S. e Wong, K.Y. "Development and validation of knowledge management performance measurement constructs for small and medium enterprises", Journal of Knowledge Management, 19 (4), pp. 711-734.

Li, Y., Tarafdar, M. e Rao S.S. 2012. Collaborative knowledge management practices: theoretical development and empirical analysis. International Journal of Operation & Production Management 32 (4), 398-422.

Liao, S.-H., Fei, W.-C. e Chen, C.-C. "Knowledge sharing, absorptive capacity, and innovation capability: an empirical study of Taiwan´s knowledge-intensive industries", Journal of Information Systems, 33 (3), 2007, pp. 340-359.

Lin, C. e Tseng, S-M. "The implementation gaps for the knowledge management system", Industrial Management and Data Systems, 105(2), 2005, pp. 208–222.

Lin, H.-F. "Knowledge sharing and firm innovation capability: an empirical study", International Journal of Manpower, 28(3/4), 2007, pp. 315-332.

Ma, Z., Qi, L. e Wang, K. "Knowledge sharing in Chinese construction project teams and its affecting factors", Chinese Management Studies, 2(2), 2008, pp. 97-108.

Martins, E.C. e Meyer, H.W.J. "Organizational and behavioral factors that influence knowledge retention", Journal of Knowledge Management, 16(1), 2012, pp. 77-96.

Oliveira, M., Caldeira, M. e Romão, M. "Knowledge Management Implementation: An Evolutionary Process in Organizations', Knowledge and Process Management 19 (1), 2012, pp.17-26.

Ragab, M.A.F. e Arisha, A. "Knowledge management and measurement: a critical review", Journal of Knowledge Management, 17(6), 2013, pp. 873-901.

Ruggles, R. “The state of the notion: knowledge management in practice”, California Management Review, 40(3), 1998, pp. 80-89.

Saito, A, Umemoto, K e Ikeda, M. “A strategy-based ontology of knowledge management technologies”, Journal of Knowledge Management, 11(1), 2007, pp. 97-114.

Tan, L.P. e Wong, K. Y. “Linkage between knowledge management and manufacturing performance: a structural equation modeling approach”, Journal of Knowledge Management, 19 (4), 2015, pp. 814-835.

Tan, H.H. e Zhao, B. “Individual and perceived contextual level antecedents of individual technical information inquiry in organizations,” The Journal of Psychology, 37(6), 2003, pp. 597-621.

Teh, P.-L. and Yong, C.-C. “Knowledge sharing in IS perspective organizational behavior’s perspective”, Journal of Computer Information Systems, 51(4), 2011, pp. 11-21.

Wang, S. and Noe, R.A. “Knowledge sharing: a review and directions for future research”, 20(2), 2010, pp. 115-131.

Wang, W.-T. e Wei, Z.-H. “Knowledge sharing in wiki communities: an empirical study”, Online Information Review, 35(5), 2011, pp. 799-820.

Xue, Y., Bradley, J. e Liang, H. “Team climate, empowering leadership, and knowledge sharing”, Journal of Knowledge Management, 15(2), 2011, pp. 299-312.

Yeh, Y.-J., Lai, S.-Q. e Ho, C.-T., (2006) "Knowledge management enablers: a case study", Industrial Management & Data Systems, 106(6), 2006, pp. 793-810.

Yu, T.-K., Lu, L.-C. e Liu, T.-F. “Exploring factors that influence knowledge sharing behaviour via weblogs”, Computers in Human Behavior, 26(1), 2010, pp. 32-41.

Outsourcing de Serviços de Sistemas de Informação na Banca em Portugal

Cidália Pereira[1], João Varajão[2], Luis Amaral[3], Delfina Sá Soares[4], António Cunha[5]

1) Universidade de Trás-os-Montes e Alto Douro – Mestrado em Informática, Portugal
ccrisgp@hotmail.com

2) Universidade do Minho – Departamento de Sistemas de Informação – Centro ALGORITMI, Portugal
varajao@dsi.uminho.pt

3) Universidade do Minho – Departamento de Sistemas de Informação – Centro ALGORITMI, Portugal
amaral@dsi.uminho.pt

4) Universidade do Minho – Departamento de Sistemas de Informação – Centro ALGORITMI, Portugal
dss@dsi.uminho.pt

5) Universidade de Trás-os-Montes e Alto Douro – Departamento de Engenharias, Portugal
acunha@utad.pt

Resumo

Ao longo das últimas décadas a função sistemas de informação (SI) evoluiu e tornou-se verdadeiramente fundamental no seio das organizações. Apesar de esta evidência, nem sempre as organizações dispõem dos recursos internos e do *know-how* necessário para a desenvolver. O outsourcing de serviços de SI tem vindo a apresentar-se como uma alternativa e opção estratégica no contexto da gestão desta função em muitas organizações e em vários setores de atividade. Desenvolver uma melhor compreensão sobre as principais práticas de outsourcing de serviços de SI no setor da banca, dadas as particularidades que apresenta (por exemplo relativamente aos requisitos de confidencialidade e sigilo), constitui o objetivo central deste trabalho. Para tal, foi realizado um inquérito baseado em questionário com o intuito de caraterizar, entre outros aspetos, os serviços sujeitos a outsourcing, as principais motivações, riscos, barreiras e impactos associados ao outsourcing, bem como aspetos de contratação, relação cliente-fornecedor e satisfação. Os principais resultados obtidos são apresentados neste artigo.

Palavras chave: *Outsourcing*, sistemas de informação, motivações, riscos, entidades bancárias.

1 Introdução

A globalização dos mercados, com a consequente intensificação da competição e o crescente nível de exigência relativamente a produtos e a serviços, leva a acreditar que, num futuro

próximo, todos os aspetos da organização vão influenciar o seu posicionamento competitivo, muito particularmente o seu Sistema de Informação (SI) [Amaral e Varajão 2007].

A Função Sistemas de Informação (FSI), entendida como um conjunto de atividades que visa a otimização de um sistema de informação, é fundamental para o sucesso das organizações, dada a relevância da informação como o recurso que mais pode potenciar ou condicionar a sua gestão [Varajão 2002; Varajão 2005].

A forma como a FSI tem evoluído ao longo dos anos tem sido influenciada pela própria evolução das Tecnologias da Informação (TI) e pelo modo como estas são utilizadas e geridas na organização. A este nível, uma das tendências que mais se tem vindo a evidenciar refere-se à adoção de práticas de outsourcing de serviços de sistemas de informação.

O outsourcing consiste em transferir atividades/serviços internos para uma entidade externa, especializada nesse tipo de serviços [Varajão 2001; Uçar e Bilgen 2013]. Quando o outsourcing é feito recorrendo a organizações localizadas noutros países designa-se por offshore outsourcing [Click e Duening 2005; Dolgui e Proth 2013].

A crescente sofisticação e maturidade dos fornecedores de serviços de SI, bem como a evolução das TI, têm acelerado a aceitação do outsourcing de serviços de SI e a sua consideração como um dos instrumentos de gestão mais utilizado pelas organizações [Leeney et al 2011; Dibbern et al. 2012].

As organizações recorrem ao outsourcing em maior ou menor nível [Uçar e Bilgen 2013]. No passado, recorriam ao outsourcing principalmente pequenas e médias empresas, com falta de recursos ou com problemas na gestão das TI. Por volta de 1990, principalmente após a decisão da Kodak de recorrer ao outsourcing, esta prática tornou-se uma das principais opções estratégicas de gestão da FSI das grandes organizações [Varajão et al. 2009b; Dibbern et al. 2012].

O outsourcing surge também como uma forma das organizações poderem concentrar a sua atenção, esforço e *know-how* nas áreas que constituem o seu *core business*, melhorando o seu posicionamento competitivo no mercado, e deixando que outros se encarreguem dos elementos "acessórios" e de suporte ao seu negócio [Liston et al. 2007]. Uma maior concentração na atividade principal, possibilita a otimização da organização, melhorando o seu desempenho e proporcionando um melhor atendimento aos clientes [Gonzalez et al. 2013].

Nos últimos anos, o outsourcing de SI teve uma evolução extraordinária, prevendo-se que o seu crescimento continue e se acentue ainda mais nos próximos anos [Varajão 2001; Gorla e Somers 2014]. O outsourcing de serviços de SI é atualmente visto pelas organizações como

fundamental para a transformação dos seus negócios, com os fornecedores a focarem a sua oferta ao nível da capacidade de inovação e associação a práticas sofisticadas de *Business Process Outsourcing* [Gonzalez et al. 2013].

A investigação sobre outsourcing de SI tem igualmente evoluído ao longo dos últimos anos, refletindo mudanças na sua prática e um melhor entendimento sobre o seu impacto no negócio [Uçar e Bilgen 2013]. Alguns dos tópicos mais investigados estão relacionados com motivações e riscos da adoção de outsourcing, relação entre cliente e fornecedor, e benefícios e aplicação de teorias económicas a este fenómeno [Gonzalez et al. 2006].

Embora o outsourcing de SI seja um tema abordado por investigadores de todo o mundo, em Portugal são ainda limitados os estudos nesta área, principalmente em setores específicos, como é o caso do setor bancário. Este setor é particularmente interessante, dado que questões como a confidencialidade e o sigilo assumem particular relevância. Considera-se, por isso, importante caracterizar as práticas de outsourcing a nível das entidades bancárias portuguesas, identificando e documentando a sua realidade e estabelecendo uma base de partida que permita o acompanhamento da sua evolução. O presente artigo foca-se, como tal, no estudo do outsourcing na banca em Portugal. A Figura 1 sistematiza as áreas de estudo sobre as quais incide este trabalho.

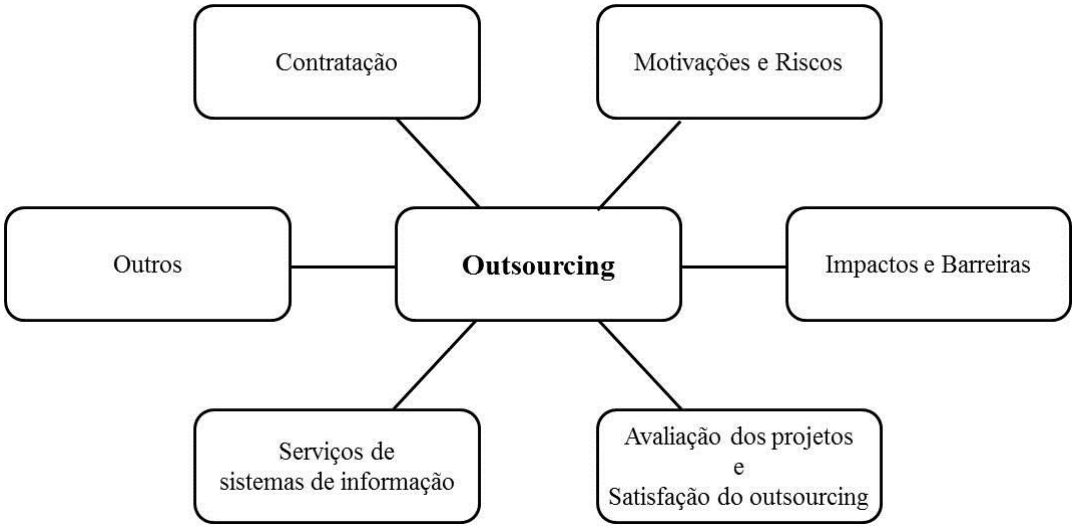


Figura 1 – Áreas abordadas no estudo

Com vista a efetuar a caracterização pretendida, foi realizado um inquérito baseado em questionário procurando identificar os serviços de SI sujeitos ao outsourcing, as motivações que levam as organizações a aderir ao outsourcing, os riscos e barreiras existentes, os impactos

organizacionais resultantes do outsourcing, os aspetos considerados na elaboração de contratos de outsourcing e na seleção dos fornecedores. Procurou-se também identificar como é feita a avaliação/monitorização dos projetos de outsourcing, o grau de satisfação relativamente ao outsourcing, e se é frequente o recurso ao outsourcing offshore. Os principais resultados obtidos são apresentados neste artigo.

Este documento encontra-se organizado em cinco secções. De seguida, efetua-se, na secção 2, um enquadramento concetual do tema. Na secção 3 descreve-se o método de investigação seguido e efetua-se a caracterização das entidades bancárias participantes no estudo. Na secção 4 analisam-se e discutem-se os dados recolhidos. Finalmente, na secção 5, tecem-se considerações finais sobre o trabalho desenvolvido.

2 Outsourcing de sistemas de informação – breve enquadramento

Nesta secção são abordadas as principais motivações subjacentes à adoção do outsourcing, bem como os principais riscos, barreiras e impactos associados à sua prática. São ainda apresentados alguns aspetos fundamentais relativos à contratação.

2.1 Serviços de sistemas de informação

O outsourcing de serviços de SI tornou-se desde há alguns anos uma das ferramentas de gestão mais utilizada pelas organizações na construção e na manutenção das suas capacidades competitivas [Dolgui e Proth 2013].

Alguns dos serviços de SI mais sujeitos a outsourcing incluem [Reid 1996; Uçar e Bilgen 2013]:

- Sistemas e aplicações de software, ferramentas e equipamento;
- Integração de sistemas;
- Desenvolvimento de novos programas e sistemas;
- Segurança de dados e programas e recuperação de falhas;
- Digitação de dados;
- Manutenção de redes e aplicações;
- Manutenção de PCs (Computadores Pessoais);
- Suporte técnico (help desk).

Para além desses, outros serviços têm vindo a ser alvo de outsourcing nomeadamente [Varajão 2001; Frost e Sullivan 2005]:

- Formação de recursos humanos;
- Gestão de centros de dados;
- Instalação de redes de comunicação de dados;
- Gestão de configurações;
- Acesso a bases de dados externas;
- Instalação e manutenção de equipamento;
- Migração de sistemas;
- Gestão de projetos;
- Serviços de Internet.

Adicionalmente, como reflexo da crescente propagação da internet, as empresas têm vindo a recorrer ao outsourcing de serviços de alojamento Web, de plataformas e serviços de comunicações, de desenvolvimento e manutenção de aplicações Web, de serviços de mensagens e-mail e de gestão de redes [Varajão et al. 2007a].

2.2 Motivações

As motivações que levam as organizações a recorrer ao outsourcing são muito diversificadas.

Num estudo desenvolvido pelo The Outsourcing Institute (EUA), foram identificadas dez razões para a adoção do outsourcing [Faulhaber 2005], nomeadamente:

- **Concentração no negócio** – o outsourcing de serviços de SI permite às organizações concentrarem-se na atividade principal do seu negócio, melhorando a qualidade dos serviços e o atendimento aos clientes;

- **Melhoria de processos** – recorrendo a fornecedores externos as organizações são capazes de alcançar mais rapidamente os benefícios pretendidos com a reorganização dos processos;

- **Partilha de riscos** – ao proceder ao outsourcing de um serviço, a organização passa a partilhar os riscos inerentes a esse serviço;

- **Libertação de recursos e capital** – ao libertar recursos estes podem ser redirecionados para áreas diferentes da organização;

- **Aumento das capacidades em TSI** – o outsourcing permite o acesso a recursos, serviços e tecnologias mais qualificados e especializados, não disponíveis internamente;

- **Transferir uma função difícil de gerir** – muitas das atividades da FSI são complexas e difíceis de gerir, podendo assim beneficiar-se da sua transferência para parceiros exteriores.

Adicionalmente, são motivações para o outsourcing a redução de custos, a criação de valor e a obtenção de vantagens competitivas [Uçar e Bilgen 2013], bem como a melhoria dos serviços, maior atualização das TI e maior nível de flexibilidade das organizações para responder às exigências do mercado [Dibbern et al. 2012; SI 2009].

Para Lacity e Hirschheim [1995] as motivações para a adoção do outsourcing podem ser agrupadas em quatro categorias: financeiras; negócio; técnicas; e políticas.

As motivações financeiras são as mais referidas na literatura, havendo, no entanto, especialistas que defendem que estas não devem ser as principais razões para o outsourcing. Apesar disso, a redução de custos continua a ser um dos principais atrativos e a assumir um grande peso na decisão de outsourcing [Uçar e Bilgen 2013]. Um dos pressupostos básicos do outsourcing é o facto de o fornecedor poder disponibilizar os mesmos serviços a um custo mais baixo do que o departamento interno do cliente consegue, graças a eficiências resultantes da produção em massa e da especialização do trabalho [Varajão et al. 2006].

As motivações relacionadas com o negócio são também relevantes. Com efeito, ao permitir a concentração no negócio, a versatilidade e a partilha de riscos, o outsourcing conduzirá à melhoria de processos e à alteração das estruturas organizacionais [Valença e Barbosa 2002].

As motivações técnicas são também reconhecidas, dado que os fornecedores possuem competências técnicas e tecnológicas que uma organização pode não ter, oferecendo acesso a novas tecnologias, a ferramentas e a técnicas melhoradas [Gonzalez et al. 2013; Varajão et al. 2006].

Por vezes a gestão de SI pode tornar-se complexa, conflituosa e difícil de controlar [Varajão 2002]: quando a função interna de SI é percebida como ineficiente, ineficaz ou tecnicamente incompetente; quando o departamento de SI se torna incómodo a diversos níveis de gestão; ou quando se tornam obscuras as relações de autoridade e de responsabilidade. Quando a gestão procura uma solução para problemas deste tipo, o outsourcing apresenta-se como uma opção

viável, dado que a gestão de topo parece ser muitas vezes atraída pela ideia de que o outsourcing irá permitir assumir o controlo dos seus gastos em TI [Varajão 2001].

O otimismo presente na literatura e imprensa relativamente aos resultados do outsourcing é outro fator motivador importante, bem como as pressões colocadas por *stakeholders* e o efeito "moda", ou seja, a vontade de seguir a tendência de adoção evidenciadas pelos pares [Varajão 2002; Dibbern et al. 2012; Gonzalez et al. 2013].

2.3 Riscos

Os riscos do outsourcing são um componente vital da decisão de obtenção de serviços [Sá-Soares et al. 2014]. Embora já constitua uma prática frequente nas organizações, e não obstante as muitas vantagens que lhe são apontadas, é importante realçar que se trata de uma atividade que pode comportar riscos significativos [Varajão 2001].

Alguns dos riscos associados ao outsourcing podem ser [Varajão 2001; Frost & Sullivan 2005]:

- **Seleção de um fornecedor sem competências adequadas** – com a enorme oferta existente no mercado, a organização pode acabar selecionando um fornecedor sem competências adequadas para a prestação do serviço;

- **Custos não previstos** – embora teoriacamente os custos sejam especificados nos contratos, podem emergir custos não previstos, fruto da inexperiência da organização no estabelecimento de contratos de outsourcing ou da falta de conhecimento sobre o serviço que pretende contratar;

- **Quebra de segurança em informação confidencial ou dados do negócio** – no processo de outsourcing, o fornecedor tem acesso a dados da organização, procedimentos de negócio, assim como, em alguns casos, a informação dos setores estratégicos da organização;

- **Ficar preso aos fornecedores, quando há um mercado dinâmico** – devido a contratos de longa duração que podem condicionar a empresa na sua ação;

- **Outsourcing de projetos de serviços de TI sem objetivos claros** – passar a responsabilidade de um projeto de TI para entidades externas sem ter bem claro o objetivo e finalidade do projeto pode levar a conflitos;

- **Incompatibilidade cultural e problemas com o fornecedor** – devido a diferentes culturas organizacionais, falta de cumprimento do contrato e de entendimento entre cliente e fornecedor;

- **Perda de controlo e flexibilidade nos processos vitais para o negócio –** a transmissão de alguns processos de negócio para uma entidade externa pode restringir o modo como uma organização reage ou se adapta à mudança do ambiente de negócio;

- **Incumprimento do fornecedor** – os fornecedores podem apresentar um desempenho inferior, por falta de conhecimento do negócio e de ambientação com a organização e seus clientes, ou por se encontrarem em situação económica que os impeça de ter todos os recursos necessários para prestar o serviço;

- **Resistência à mudança por parte dos funcionários** – potenciais transferências ou rescisões de contratos podem causar desmotivação dos funcionários.

Podem também constituir riscos do outsourcing [Varajão 2001; Tho 2005; Aubert et al. 2005]:

- **Demasiada dependência do fornecedor e ameaça de oportunismo** – não é fácil mudar de fornecedor ou terminar contrato, tornando assim a organização refém do fornecedor; podem ser necessários grandes investimentos para reconstruir serviços e recontratar profissionais em SI;

- **Resistência interna** – a desconfiança entre organizações, o choque de diferentes culturas, o desempenho inadequado por parte do fornecedor;

- **Conflito de interesses entre fornecedor e cliente** – têm metas divergentes: fornecedores visam maximizar os lucros; cliente visam reduzir custos.

Segundo Varajão [2001], constituem também riscos significativos do outsourcing:

- **Perda de capital intelectual** –um dos riscos mais sérios e menos compreendidos pelos gestores que tomam a decisão de outsourcing, dado que o outsourcing pode levar a uma perda de conhecimento nas organizações;

- **Perda de competências e capacidades em sistemas de informação** – com a transferência de atividades de TSI para o fornecedor, a oportunidade para a utilização inovadora de TI pode ficar comprometida. Após esta transferência

provavelmente a organização reduz a equipa de SI e o conhecimento e capacidade existente;

- **Qualidade de serviços –** não existe certeza do cumprimento do nível de qualidade acordado;

- **Custos associados ao ciclo de vida dos contratos –** os contratos de outsourcing podem não ser cumpridos, obrigando a uma renegociação ou até mesmo cancelamento do contrato, o que acarreta consequências (por exemplo, financeiras);

- **Exposição aos riscos de instabilidade do fornecedor** – com os contratos de outsourcing uma organização fica exposta aos riscos do fornecedor;

- **Subcontratação de serviços por parte do fornecedor** – muitas vezes quem fornece os serviços não é o fornecedor contratado mas empresas terceiras, subcontratadas pelo fornecedor.

2.4 Barreiras

No decorrer de um processo de outsourcing são também várias as barreiras que podem surgir.

Um dos grandes fatores de resistência na decisão pelo outsourcing é o receio da perda de controlo [CPW 2005]. Gonzalez et al. [2013] refere que os serviços não devem ser todos entregues ao mesmo fornecedor, sendo necessário selecionar os produtos por áreas de competência e garantir que não se fica ligado a um só fornecedor por vários anos.

Outro fator de resistência é a dificuldade em estabelecer uma relação de confiança entre o cliente e o fornecedor, podendo existir dificuldade por parte do cliente em encontrar o fornecedor adequado [CPW 2005; Gonzalez et al. 2013].

Para além destas, existem ainda outras condicionantes na adoção do outsourcing, como a partilha de informação confidencial, barreiras culturais, políticas, organizacionais e financeiras [IDC 2002].

A resistência dos colaboradores é outra das barreiras apontadas, dado que podem sentir dificuldade em adaptar-se a novas formas e novos ambientes de trabalho, dado que alguns podem passar a trabalhar com a empresa contratada [Aubert et al. 2005].

Encontrar o fornecedor com as competências adequadas às necessidades do cliente pode também ser uma das barreiras do outsourcing [Uçar e Bilgen 2013].

Os custos, apresentados anteriormente como um dos motivos para a adoção do outsourcing, podem, paradoxalmente, ser também uma barreira, devido ao custo elevado que podem atingir alguns contratos e projetos, sobretudo se não forem bem negociados e geridos [Albertin e Sanchez 2008].

2.5 Contratos

As entidades externas quando se responsabilizam por determinadas funções, fazem-no segundo um contrato que deve incluir a definição dos níveis de serviço [Fraga et al. 2012]. Os *Service Level Agreement* (SLA) são instrumentos que especificam as funções a cumprir, bem como os parâmetros para avaliação da qualidade de execução dessas funções [Carapuça 2003].

O sucesso dos contratos de outsourcing requer uma análise cautelosa das necessidades do negócio, uma definição clara de objetivos e uma cuidadosa avaliação e seleção do fornecedor [Uçar e Bilgen 2013].

No processo de seleção dos fornecedores são utilizados diversos critérios, que podem ser agrupados como: técnicos; estratégicos; e comerciais. Nos critérios técnicos são avaliadas questões como competência, experiência e metodologias. A avaliação dos critérios estratégicos é efetuada relativamente ao alinhamento com objetivos de negócio, posição estratégica no mercado, afinidade cultural e negócios futuros. Por último, os critérios comerciais têm em conta aspetos como preço, condições financeiras e modelo de contratação [Prescott 2008].

Aquando do processo de celebração de um contrato de outsourcing existem ainda outros aspetos a considerar, nomeadamente a definição de objetivos e a duração do contrato [Gonzalez et al. 2013]. Para Resina [2008], nesta fase do processo deve-se, entre outras coisas, identificar claramente os serviços e custos, definir os SLA (fixando os níveis e formas de medição) e clarificar as responsabilidades das partes envolvidas, para além de se estabelecerem mecanismos de governo e adaptação do contrato.

Normalmente os contratos de outsourcing são contratos de longo prazo. Segundo a Gartner, a duração média dos 20 maiores contratos de outsourcing estabelecidos em 2008 foi de seis anos. No entanto, verifica-se que a duração global tem vindo a cair, o que indica uma maior preferência por acordos de curta duração (inferior a cinco anos), eventualmente fruto da atual crise económica ou para evitar ficar preso por longo tempo a um determinado fornecedor [Fraga et al. 2012; Gonzalez et al. 2013].

Metade das organizações que assinam contratos de outsourcing ficam insatisfeitas com o acordo estabelecido ao fim do primeiro ano [CPW 2005]. Algumas das razões para tal insatisfação são

[CPW 2005]: o não cumprimento das expetativas; os custos não diminuírem; não se alcançarem os níveis de rendimento previstos; e a falta compreensão e informação sobre quais os objetivos dos serviço contratado .

As dificuldades verificadas na relação entre cliente e fornecedor são vistas como uma tendência incontornável, pois cada vez mais os clientes procuram um parceiro de negócio e não um mero fornecedor [SI 2009]. A capacidade do fornecedor para adicionar valor ao contrato, o aumento da produtividade, a diminuição da taxa de erros e o aumento da maturidade dos processos, assim como o facto de o fornecedor atingir conscientemente os níveis de serviços acordados, contribui para o desenvolvimento de uma boa relação cliente-fornecedor [Gonzalez et al. 2013].

Todos os pontos fundamentais devem ser descritos no processo de elaboração do contrato de outsourcing, evitando fazer alterações na fase seguinte à celebração do contrato, pois o cliente e o fornecedor podem não chegar a um consenso sobre eles [CPW 2005]. Adicionalmente, a elaboração de um contrato com deficiências pode provocar custos acrescidos de adendas contratuais [Uçar e Bilgen 2013].

2.6 Avaliação de projetos de outsourcing e satisfação relativamente ao outsourcing

A monitorização, acompanhamento e avaliação dos projetos de outsourcing é um aspeto primordial. Para tal, deve proceder-se, desde início, à identificação de quais as informações relevantes para a condução da avaliação, quando e a quem deverão ser divulgadas [Schmid 2007].

Um dos aspetos importantes na avaliação dos projetos é a comunicação entre as entidades envolvidas [Gonzalel et al. 2013]. Uma boa comunicação leva a que a captura, distribuição e armazenamento das informações do projeto sejam efetuadas de forma adequada. Uma das formas de transmitir informação ao longo das várias fases do projeto é através de reuniões periódicas internas, que envolvem pessoas da organização, e reuniões externas, que envolvem o cliente. Desta forma é assegurada a comunicação entre as partes interessadas, desde o início até ao fim de um projeto [Mehta 2002].

A avaliação dos projetos de outsourcing envolve a verificação da satisfação dos requisitos estabelecidos. O cumprimento dos termos contratuais e os bons resultados obtidos com a implementação dos serviços, traduzir-se-á no sucesso do outsourcing e conduzirá, consequentemente, à satisfação das organizações clientes [Uçar e Bilgen 2013]. Tipicamente as organizações mostram satisfação relativamente ao outsourcing se verificarem uma melhoria significativa nos processos, na qualidade dos resultados, nas competências e sofisticação do

fornecedor, acederem a tecnologia de ponta, bem como verificarem maior concentração na sua atividade principal [Varajão et al. 2008; Varajão et al. 2009b].

2.7 Outsourcing offshore

O outsourcing offshore consiste em contratar serviços a entidades externas, tendo como diferença relativamente ao outsourcing onshore o facto de esses serviços serem contratados fora do país [Varajão 2001; Dolgui e Proth 2013].

Um dos serviços mais adquiridos no processo de outsourcing offshore é o desenvolvimento de software [Xiong 2008]. Este processo é praticado em todo o mundo, embora muitas organizações procurem este tipo de serviço principalmente em países como a Índia, a Irlanda, o Brasil, a China, a Rússia, Singapura, entre outros [Green 2007]. O objetivo é obter vantagens competitivas em termos de custos, qualidade e diluição dos riscos [Prikladnicki e Audy 2004].

O outsourcing offshore possibilita o acesso a um leque maior de fornecedores, que apresentem capacidades mais diversificadas e capacidades de resposta mais elevadas e a um menor custo. Não obstante, esta prática pode também encerrar novos riscos, para além dos já referidos, como é o caso do idioma, os fusos horários e a adaptação a outras culturas [Varajão 2001; CPW 2008].

3 Método de Investigação

No sentido de cumprir o objetivo deste trabalho (caracterizar a realidade do outsourcing de serviços de SI na banca em Portugal), realizou-se um estudo que envolveu um inquérito baseado num questionário online, cujo convite para participação foi enviado via email e via CTT para Gestores de SI de 32 entidades bancárias presentes em Portugal.

O questionário foi estruturado em grupos de questões que abrangiam as diferentes temáticas em estudo, tais como, a identificação dos serviços sujeitos a outsourcing, as motivações e os riscos do outsourcing, as barreiras e os impactos organizacionais, os critérios para a escolha de fornecedores e os aspetos considerados aquando da celebração de um contrato. O questionário incluiu questões de vários tipos (escolha múltipla, texto livre, hierarquização), tendo sido baseado nos questionários utilizados em trabalhos prévios que se focaram no estudo do outsourcing em grandes empresas [Varajão et al. 2008; Varajão et al. 2009b].

O questionário foi disponibilizado por um período de cinco meses, verificando-se alguma dificuldade na obtenção das respostas. A taxa de respostas válidas obtida para o questionário foi de 15,6%. As Figuras 2 a 4 caracterizam as entidades bancárias participantes no estudo.

Como se constata pela Figura 2, relativamente ao número de utilizadores finais de TI, na altura da recolha de dados 40% das entidades bancárias respondentes apresentava um número superior a 5000, 40% apresentava um número que variava entre 1 e 50 e, finalmente, 20% apresentava entre 201 e 500 utilizadores finais de TI.

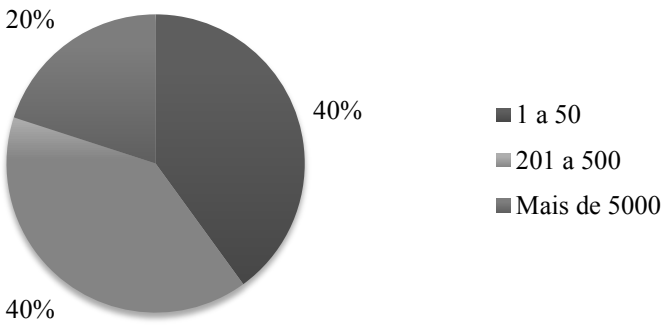


Figura 2 – Número de utilizadores finais de TI das entidades estudadas

Todas as entidades bancárias afirmaram ter um departamento de interno de informática, de SI ou similar.

Relativamente ao volume de negócios, apresentado na Figura 3, uma percentagem significativa (40%) de entidades bancárias respondentes apresentava um volume de negócios superior a 500.000.000 euros por ano.

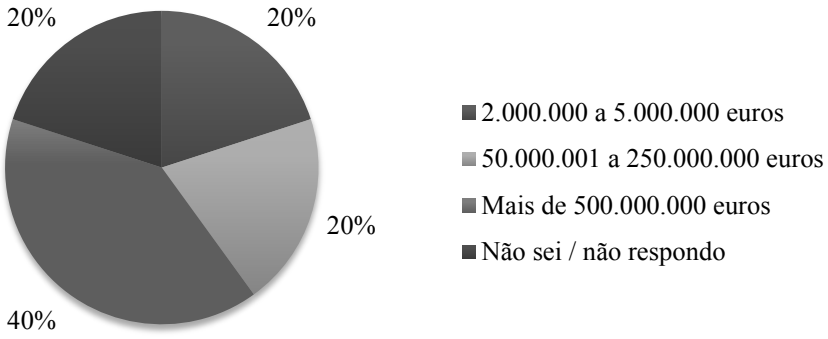


Figura 3 – Volume de negócios das entidades estudadas

No que concerne ao orçamento do departamento de informática (ou similar) existente, a realidade das entidades respondentes é a apresentada na Figura 4.

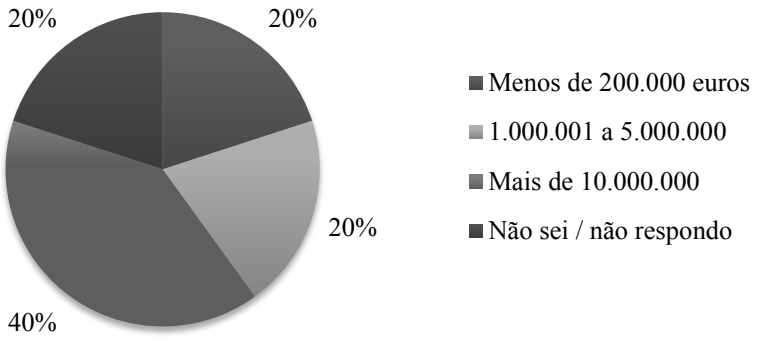


Figura 4 – Orçamento do departamento de informática (ou similar) das entidades estudadas

Como se verifica, em 40% das entidades participantes, o orçamento do departamento de informática (ou similar) é superior a 10.000.000 euros por ano.

4 Apresentação e Discussão dos Resultados

Caracterizadas as entidades bancárias participantes no estudo, analisam-se e discutem-se nesta secção os dados recolhidos no estudo.

4.1 Serviços de sistemas de informação

Nas entidades bancárias, à semelhança de organizações de outros setores de atividade, são vários os serviços de SI em que se verifica a contratação a entidades externas.

De acordo com as respostas obtidas, todas as entidades bancárias participantes recorrem ao outsourcing de serviços de SI, dedicando em média 42% do orçamento do departamento de informática (ou similar) para a contratação de serviços.

Na Figura 5 estão representados os serviços sujeitos a outsourcing nas entidades respondentes. Para cada um desses serviços pediu-se que fosse indicada a percentagem do serviço sujeita a outsourcing (um serviço totalmente fornecido por uma entidade externa, seria 100%).

Como se constata, há serviços, tais como a gestão de projetos e a formação de utilizadores, que não são alvo de outsourcing em 60% das entidades respondentes, sendo que mesmo nas entidades em que existe algum outsourcing deste tipo de serviços, isso não ultrapassa os 50% do serviço.

Figura 5 – Serviços sistemas de informação sujeitos a outsourcing

Há ainda um conjunto de serviços, nomeadamente "serviços web", "serviços profissionais de consultoria", "serviços de email e mensagens", "plataformas e serviços de comunicações", "manutenção de aplicações", "gestão de microinformática", "gestão de data center" e "desenvolvimento de aplicações", em que 80% das entidades recorrem ao outsourcing.

De modo a facilitar a comparação dos serviços em termos de outsourcing, foi criado um ranking. O gráfico da Figura 6 revela que os serviços mais sujeitos a outsourcing são "plataformas e serviços de comunicação" (60%), seguindo-se os "serviços Web" (50%) e a "gestão de redes" (56%).

Comparando os resultados obtidos com os resultados de outros estudos focados noutros tipos de empresas (ver, por exemplo, [Varajão et al. 2007b]), encontram-se diferenças significativas, por exemplo relativas aos serviços de desenvolvimento de aplicações.

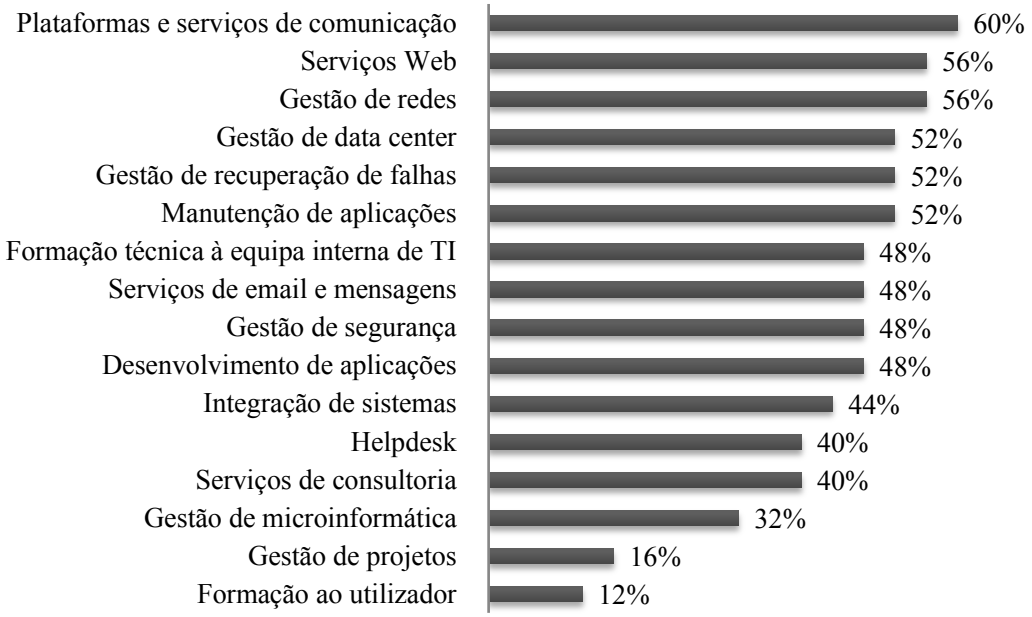


Figura 6 – Ranking dos serviços sistemas de informação sujeitos a outsourcing

Outro resultado pertinente do estudo foi que apenas 20% das entidades participantes afirmam ter procedimentos definidos para determinar se o serviço deve, ou não, ser sujeito a outsourcing.

4.2 Motivações e Riscos

O terceiro grupo de perguntas do questionário diz respeito às motivações para a adoção do outsourcing e aos riscos inerentes a essa adoção.

Na Figura 7 são apresentadas as motivações para proceder ao outsourcing.

Verifica-se que as principais motivações são o acesso a competências técnicas e a recursos humanos especializados (71%), a concentração no negócio (67%), o acesso a tecnologias da informação sofisticadas (63%) e o acesso a serviços especializados (62%).

Para as entidades respondentes, a redução de custos não é uma das principais motivações para a adoção do outsourcing (ficando na nona posição de uma lista de 19 motivações, com 38%). Este resultado acaba por ser surpreendente quando comparado com o resultado obtido num estudo de outsourcing em grandes empresas [Varajão et al. 2008; Varajão et al. 2009b], no qual a redução de custos surge classificada como a segunda principal motivação para o outsourcing.

Figura 7 – Motivações para outsourcing

Na Figura 8 são apresentados os riscos associados ao outsourcing que mais preocupam as entidades bancárias participantes.

Como se verifica, os principais riscos identificados pelas entidades bancárias no outsourcing são a exposição aos riscos de instabilidade do fornecedor (78%), a perda de capital intelectual (63%), os custos não previstos (57%), o ficar preso aos fornecedores num mercado dinâmico (57%) e a perda de competências e capacidades em sistemas de informação (55%).

Figura 8 – Riscos do outsourcing

4.3 Barreiras e Impactos

Outro aspeto estudado foram as barreiras que as entidades bancárias enfrentam ao adotar o outsourcing e os impactos organizacionais que surgem de um processo de outsourcing.

Curiosamente, as entidades bancárias não identificaram qualquer barreira para a adoção de outsourcing nos seus casos.

Relativamente aos impactos organizacionais, mediante uma lista de impactos apresentados, foi pedido aos participantes que atribuíssem uma classificação a cada impacto (onde 1 representava o menor impacto e 5 o impacto mais significativo). Na Figura 10 apresenta-se o ranking dos impactos verificados.

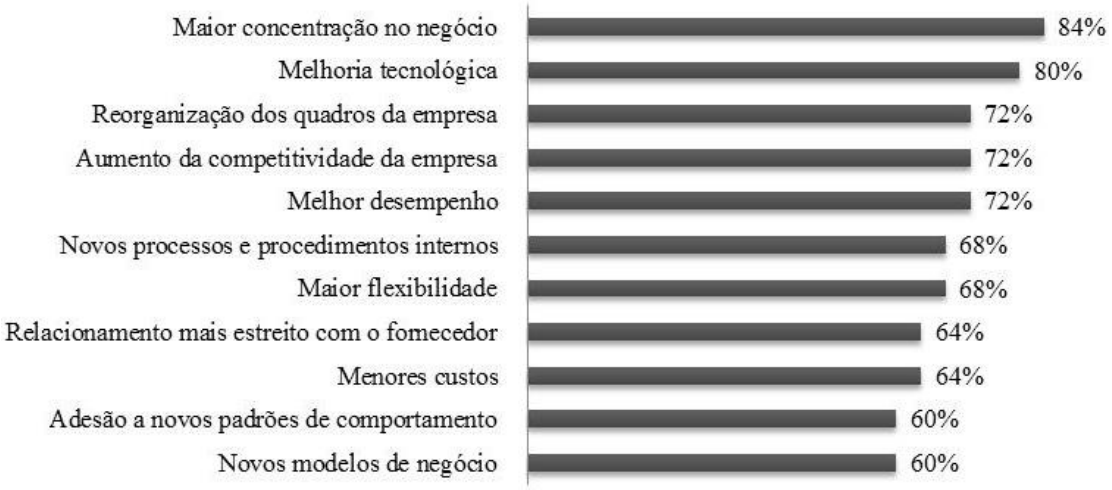


Figura 9 – Impactos organizacionais resultantes do outsourcing

Note-se que o maior impacto sentido pelas entidades bancárias é a capacidade de ter uma maior concentração no negócio (84%), seguindo-se a melhoria tecnológica (80%), reorganização dos quadros da empresa (72%), aumento da produtividade (72%) e melhor desempenho (72%).

4.4 Contratos de outsourcing e relação com os fornecedores

No processo de adoção de outsourcing, uma das fases que requer grande atenção é a de seleção dos fornecedores, onde é necessário utilizar um conjunto de critérios bem definidos.

Para as entidades bancárias, os critérios mais relevantes para a seleção dos fornecedores são a capacidade para fornecer o tipo de serviço pretendido (100%), capacidade técnica e de gestão (88%) e o conhecimento e experiência na área de negócio da empresa.

No que concerne à celebração de contratos, os aspetos considerados pelas entidades como os mais relevantes a este nível foram: a descrição precisa dos serviços a serem fornecidos (81,4%); a definição dos níveis de serviço (67,1%); e a definição de responsabilidades das partes (58,6%).

No que se refere à duração dos contratos, 40% das entidades bancárias afirmou que os contratos duram em média três anos.

As principais dificuldades verificadas na relação com os fornecedores são a falha no cumprimento de prazos (78,9%), desacordo relativamente ao âmbito/preço (76,8%) e falhas no cumprimento dos níveis de serviço acordados (75,8%).

Segundo as entidades bancárias, os aspetos que mais conflitos originam com o fornecedor são as falhas nos níveis de serviço (68,6%) e a falta de comprometimento e atitude de serviço por parte do fornecedor (60%).

Existindo a possibilidade de ocorrência de conflito entre cliente (entidade bancária) e fornecedor, as soluções apresentadas pelos respondentes para resolver este problema de conflito são a renegociação de contratos (80%) e a inclusão de cláusulas contratuais para a resolução de conflitos (60%).

De acordo com as entidades participantes, a avaliação/monitorização dos projetos de outsourcing é feita por comissão conjunta (cliente/fornecedor) (80%), elaboração de relatórios periódicos de desempenho (60%) e por equipa interna para avaliação/monitorização (60%).

Com o outsourcing de serviços de SI, as entidades participantes consideram que genericamente não haverá perda de controlo (60%).

No que refere à satisfação das entidades com os resultados do outsourcing, muitos são os fatores em destaque. A concentração no negócio (88%) é um dos fatores principais, assim como o relacionamento com os fornecedores (88%).

Finalmente, é de notar que nenhuma das entidades participantes referiu recorrer ao outsourcing offshore.

5 Conclusão

Este trabalho procurou caracterizar a realidade do outsourcing de serviços de SI na banca em Portugal em várias vertentes. Para tal, foi realizado um inquérito baseado em questionário. Através das respostas obtidas foi possível caracterizar e documentar aspetos como: motivações e riscos; barreiras e impactos organizacionais; aspetos a considerar na elaboração de contratos de outsourcing; entre outros.

Do conjunto de entidades bancárias respondentes, 40% tem mais de 5.000 utilizadores finais de TI, 40% têm de 1 a 50 utilizadores de TI, sendo que outros 20% têm de 201 a 500 utilizadores de TI. O volume de negócios das entidades bancárias (40%) enquadra-se em mais de 500 milhões de euros e o orçamento do departamento de informática em mais de 10 milhões de euros (40%).

Relativamente aos serviços SI sujeitos ao outsourcing, “plataformas e serviços de comunicação” (60%) é o serviço mais contratado a entidades externas, seguindo-se “serviços Web” e “gestão de redes” com 56%, no entanto, apenas 20% das entidades bancárias afirmam ter procedimentos definidos para determinar se um serviço deve, ou não, ser sujeito a um processo de outsourcing.

Várias são as razões que levam as entidades bancárias a optar pelo outsourcing, sendo as mais prevalecentes o acesso a competências técnicas e recursos humanos especializados (71%), concentração no negócio (67%) e acesso a TI de ponta (63%). As entidades bancárias também identificam diversos riscos no outsourcing, destacando-se a exposição aos riscos de instabilidade do fornecedor (78%), perda de capital intelectual (63%), custos não previstos (57%) e ficar preso aos fornecedores num mercado dinâmico (57%). No entanto, as entidades bancárias não identificam quaisquer barreiras na adoção de outsourcing e referem como principais impactos organizacionais resultantes do processo de outsourcing, a maior concentração no negócio (84%) e a melhoria tecnológica (80%).

Os critérios apontados como mais relevantes aquando da seleção de fornecedores são a capacidade para fornecer o tipo de serviço pretendido (100%), capacidade técnica e de gestão (88%) e conhecimento e experiência na área de negócio da empresa. Em relação ao processo de celebração do contrato de outsourcing, os mais destacados foram: a descrição precisa dos serviços a serem fornecidos (81,4%); a definição clara dos níveis de serviço (67,1%); e a definição de responsabilidades das partes (58,6%). No que diz respeito à duração dos contratos, 40% das entidades bancárias afirma que os contratos duram em média três anos.

As principais dificuldades verificadas na relação com os fornecedores são a falha no cumprimento de prazos (78,9%), desacordo relativamente ao âmbito/preço (76,8%) e falhas no cumprimento dos níveis de serviço acordados (75,8%). Segundo as entidades participantes, os aspetos que mais conflitos originam com os fornecedores são as falhas nos níveis de serviço (68,6%) e a falta de comprometimento e atitude de serviço por parte do fornecedor (60%). Existindo a possibilidade de ocorrência de conflito, a solução apresentada pelas entidades bancárias para resolver este problema é a renegociação de contratos (80%) e a inclusão de cláusulas contratuais para a resolução de conflitos (60%).

De acordo com as entidades bancárias participantes no estudo a avaliação/monitorização dos projetos de outsourcing é feita por comissão conjunta (cliente/fornecedor) (80%), elaboração de relatórios periódicos de desempenho (60%) e por equipa interna para avaliação/monitorização (60%). Com o outsourcing de serviços de SI, 60% das entidades participantes consideram não haver perda de controlo, 40% referem haver perda de controlo em processos vitais para o negócio e 40% referem a impossibilidade de adotar tecnologias do âmbito do contrato.

No que diz respeito à satisfação da entidade bancária relativamente aos resultados verificados com outsourcing, muitos são os fatores em destaque. A concentração no negócio (88%) é um dos fatores principais, assim como relacionamento com os fornecedores (88%). Nenhuma das entidades bancárias participantes no estudo referiu recorrer ao outsourcing offshore.

Verifica-se que o outsourcing tem hoje uma importância incontornável na vida das entidades bancárias. Este estudo ajuda a perceber os contornos atividade de outsourcing neste tipo de organizações. Embora um retrato pontual desta realidade nas entidades bancárias seja já um contributo enriquecedor, considera-se e propõem-se como trabalho futuro a reedição do mesmo estudo, de forma a ser possível efetuar uma análise longitudinal do fenómeno nas entidades bancárias e descobrir padrões evolutivos, úteis para as entidades que atuam neste setor de atividade.

6 Agradecimento

Este trabalho foi suportado pela FCT – Fundação para a Ciência e Tecnologia FCT – no âmbito do Projeto UID/CEC/00319/2013.

7 Referências

Albertin, A. e O. Sanchez, *Outsorcing TI : impactos, dilemas, discussões e casos reais*, FGV, Rio de Janeiro, 2008.

Amaral, L. e J. Varajão, *Planeamento de sistemas de informação*, 4.ª ed., FCA, 2007.

Aubert, B., M. Patry, e S. Rivard, "A framework for information technology outsourcing risk management", *ACM SIGMIS Database* 36, 4 (2005), 9-28.

Carapuça, R, "Outsourcing", *E_motion*, 12, Publicação Novabase, 2003.

Click, R. e T. Duening, *Bussines Process Outsourcing*. *The Competitive Advantage*, John Wiley and Sons, 2005.

CPW, "Os caminhos do Outsourcing de SI/TI", *Computerworld Portugal*, http://www.computerworld.com.pt/2005/10/06/os-caminhos-do-outsourcing-de-siti/ (5 janeiro 2015), 2005.

CPW, "Satisfação do cliente deve ser factor-chave no outsourcing", *Computerworld Portugal*, http://www.computerworld.com.pt/2008/05/15/satisfao-do-cliente-deve-ser-factor-chave-no-outsourcing/ (5 janeiro 2015), 2008.

Dibbern, J., W. Chin e A. Heinzl, "Systemic Determinants of the Information Systems Outsourcing Decision: A Comparative Study of German and united States Firms", *Journal of the Association for Information Systems*, 13, 6 (2012), 466-497.

Dolgui, A. e J. Proth, "Outsourcing; definitions and analysis", *International journal of Production Research*, 51, 23-24 (2013), 6769-6777.

Dubois, A., K. Hulthen, P. Jonsson e V. Sundquist, □"Supply Network Flexibility", *20th IMP Conference*, Denmark, 2004.

Faulhaber, T., *Outsourcing*, http://www.businessforum.com/outsourcing1.html, (12 maio 2012), 2005.

Fraga, M., J. Varajão, L. Amaral, J. Cruz, "Information systems outsourcing in major Portuguese companies - contracting services", *Journal of Research and Practice in Information Technology*, 44(1) (2012), 81-106.

Frost & Sullivan, *Fazendo o Outsourcing de TI dar certo*, http://rmaia.redes.googlepages.com/MA_UnisysWhitePaper.pdf (6 março 2015), 2005.

Gonzalez, R., J. Gasco, e J. Llopis, "Information systems outsourcing: a literature analysis", *Information & Management*, 43, 7 (2006), 821-834.

Gonzalez, R., J. Llopis e J. Gasco, "Information technology outsourcing in financial services", *The Service Industries Journal*, 33, 9-10 (2013), 909-924.

Gorla, N. e T. Somers, "The impact of IT outsourcing on information systems success", *Information & Management*, 51 (2014), 320-335.

Green, A., *Offshore Outsourcing*, 2007.

IDC, *Caderno n.º 51 – Outsourcing*, IDC Portugal, Lisboa, 2002.

Lacity, M. e R. Hirschheim, *Beyond the Information Systems Outsourcing*, Wiley, Bandwagon, 1995.

Leeney, M., J. Varajão, A. Ribeiro e R. Palacios, "Information systems outsourcing in large companies: evidences from 20 Ireland companies", *International Journal of Information Technology Project Management*, 2 (4) (2011), 44-58.

Liston, P., J. Byrne, P. Byrne, e C. Heavey, "Contract costing in outsourcing enterprises: Exploring the benefits of discrete-event simulation", *International Journal of Production Economics*, 110, 1-2 (2007), 97-114.

Mehta, A., *Communication in Project Management*, PMI, 2002.

Prescott, R., Especial outsourcing: as respostas da PricewaterhouseCoopers, http://www.itweb.com.br/noticias/index.asp?cod=51781, (3 abril 2014), 2008.

Prikladnicki, R. e J. Audy, *Um Modelo de Referência para Desenvolvimento Distribuído de Software*, Faculdade de Informática, Rio Grande do Sul, 2004.

Reid, W., *Outsourcing: The 20 Steps to Success*, http://www.wsrcg.com/Articles/OutSourcing.pdf, (10 outubro 2014), 1996.

Resina, F., *Outsourcing: o caminho para um projeto de sucesso*, Deloitte, 2008.

Sá-Soares, F., D. Soares e J. Arnaud, "A catalog of information systems outsourcing risks", *International Journal of Information Systems and Project Management*, 2(3) (2014), 23-43.

Schmid, B., *Motivation in Project Management: The Project Manager's Perspective*. The Florida State University, Florida, 2007.

SI, "Outsourcing da novo folego as empresas", *Semana Informatica*, 921, (2009).

Tho, I., *Managing the Risks of IT Outsourcing*, Burlington, USA, 2005.

Valença, M. e A. Barbosa, A Terceirização e seus Impactos: Um Estudo em Grandes Organizações de Minas Gerais, http://www.anpad.org.br/rac/vol_06/dwn/rac-v6-n1-mcv.pdf, (20 maio 2014), 2002.

Uçar, E. e S. Bilgen, "A Case-Based Model for Assessing the Effectiveness of Information Systems Outsourcing", *Journal of Information Technology Case and Application Research*, 15, 3 (2013), 27-60.

Varajão, J., *Outsourcing de Servicos de Sistemas de Informacão*, FCA - Editora de Informatica, Lisboa, 2001.

Varajão, J., *A Arquitectura da Gestão de Sistemas de Informacão*, FCA - Editora de Informatica, Lisboa, 2005.

Varajão, J., *Função de Sistemas de Informação: Contributos para a melhoria do sucesso da adoção de tecnologias de informação e desenvolvimento de sistemas de informação nas organizações*, Tese de Doutoramento, Universidade Minho, Guimarães, 2002.

Varajão, J., A. Trigo, N. Figueiredo, J. Barroso, "Information systems services outsourcing reality in large Portuguese organisations", *International Journal of Business Information Systems*, 4(1) (2009a), 125-142.

Varajão, J., N. Ferreira, M. Fraga e L. Amaral, "Outsourcing de sistemas de informacão nas empresas nacionais", *CXO - Tecnologias de sistemas de informação para executivos*, 30 (2008), 28-33.

Varajão, J., N. Ferreira, M. Fraga, L. Amaral, "96% das empresas portuguesas recorre ao outsourcing", *Computerworld*, 800 (2009b), 16-17.

Varajão, J., R. Goncalves, J. Barroso e L. Amaral, "Outsourcing de Serviços de Sistemas de Informação", *Dirigir*, 93/separata (2006), 1-13.

Varajão, J., A. Trigo, N. Figueiredo e J. Barroso, "TI nas empresas nacionais", *CXO-Tecnologias de sistemas de informação para executivos*, 13 (2007a), 19-23.

Varajão, J., A. Trigo, N. Figueiredo e J. Barroso, "TI nas empresas nacionais: Papel das TI e Outsourcing", *Computerworld*, 729 (2007b), 9.

Xiong, W., X-T. Wang e Z-X Wu, "Study of a customer satisfaction-oriented model for outsourcing software quality management using Quality Function Deployment (QFD)", Proceedings of the *4th WiCOM'08*, Dalian, 2008.

Pontos de verificação para contratação de serviços de *Outsourcing* em Segurança da Informação de TI

Teofilo Branco Junior

Universidade do Minho, Portugal

teofilotb@hotmail.com

Resumo

Este artigo relata um estudo realizado sobre os pontos de verificação considerados relevantes para o processo de contratação de serviços de *Outsourcing* em Segurança da Informação. O presente estudo teve por base uma revisão da literatura sobre os fatores críticos de sucesso relacionados com a prestação de serviços de Segurança da Informação (SSI) por empresas terceiras. Neste estudo foram investigadas as preocupações relevantes apontadas pelos autores pesquisados a respeito da relação necessária entre clientes e fornecedores envolvida neste tipo de serviço. Estas questões foram catalogadas e organizadas com o objetivo de apresentar as principais preocupações para verificação da conformidade na prestação do serviço da Segurança da Informação por terceiros. Para que estes pontos possam se constituir em ferramentas de auditoria, há a necessidade de que sejam atribuídas métricas viáveis de registro e relacioná-las com as cláusulas respectivas nos acordos de serviço (SLAs).

Palavras chave: Sistemas de Informações, Segurança da Informação, Outsourcing de TI, Serviços de SSI, Auditoria de TI.

1. Introdução

A contratação de serviços terceirizados na área de TI tem sido crescente nos últimos tempos. Diversas empresas tem ofertado variados serviços de TI oferecendo vantagens aos clientes ao desonerar os custos de infraestrutura dos clientes em troca de serviços especializados. A princípio, trata-se uma proposta interessante e vantajosa para ambos, pois os clientes delegam as tarefas de TI a especialistas, concentrando seus esforços na sua atividade fim e estes últimos tem condições de prestar melhores serviços ao contarem com uma carteira de clientes que lhes fornecem recursos para aprimorarem suas especialidades. Porém, a qualidade da prestação do serviço, a competência e a capacidade do prestador do serviço tornam-se fatores fundamentais nesta relação.

Diante disso, promovemos um estudo, a partir de uma revisão da literatura, sobre os riscos e fatores críticos de sucesso envolvidos na contratação de serviços terceirizados (*Outsourcing*) na área da Segurança da Informação (SSI), promovendo uma discussão específica para oferta e consumo deste tipo de serviço.

A contratação do serviço de Segurança da Informação sob a forma de *Outsourcing* levanta questões amplamente exploradas na área acadêmica sobre a gestão do relacionamento de terceirização. O assunto é bastante explorado na literatura, sendo importante considerar os fatores que influenciam a promoção e manutenção de um *Outsourcing* bem sucedido (Pereira, de Sá-Soares, & Caldeira, 2011).

2. A Revisão da Literatura

Uma boa revisão de literatura cria uma base sólida para o avanço do conhecimento. Deste modo, facilita o desenvolvimento da teoria, fecha as áreas onde existe uma infinidade de pesquisas e descobre áreas onde é necessária a investigação e proporciona uma contribuição importante no estabelecimento de orientações para o futuro da investigação que são fundamentais para o fortalecimento na área em estudo (Webster & Watson, 2002).

Primeiramente, procurou-se identificar as principais questões de estudo envolvidos na prestação de serviços de SSI sob a forma de *Outsourcing* e as questões notadamente relacionados a relação de confiança envolvidas nesta prestação de serviço. Uma pesquisa da literatura foi realizada no sentido de:

i. Caracterizar os trabalhos apresentados sobre a problemática tratada, identificando os conceitos envolvidos, os pontos relevantes, as questões polêmicas envolvidas e os pontos-de-vista apresentados;

ii. Fazer uma catalogação das questões envolvidas de forma crítica em relação as opiniões dos autores pesquisados, a fim de definir os pontos de interesse dos autores;

iii. Ressaltar as premissas e os requisitos necessários para a boa prestação do serviço de SSI nesta modalidade.

O objetivo da pesquisa da literatura foi a de encontrar trabalhos que tratassem das seguintes questões:

i. Quais os riscos envolvidos no serviço de *Outsourcing* na área de SSI?

ii. Quais as condições e requisitos a serem observados para verificação da conformidade no processo de auditagem de prestação do serviço de SSI em *Outsourcing*?

Foram utilizadas as seguintes palavras-chave e termos utilizados para a pesquisa por título, nos idiomas Inglês e Português:

i. “Outsourcing de Segurança de SI”,

ii. “Riscos em Outsourcing de TI e de Segurança de SI”,

iii. “Requisitos para Outsourcing de TI”,

iv. “Auditagem de Outsourcing de Segurança em SI”.

Foram selecionadas as seguintes fontes de acordo com o seguinte critério de prioridade:

i. Publicações de artigos em revistas científicas;

ii. Teses e dissertações;

iii. Anais de conferências internacionais.

Na pesquisa das revistas científicas, foi realizada consulta no SCImago Journal & Country Rank (http://www.scimagojr.com) para verificação da visibilidade das revistas (SJR). Foi atribuída uma maior atenção aos aspectos qualitativos considerando o número de citações dos artigos, de forma a selecionar aqueles que tivessem maior impacto. Os seguintes *Journals* foram selecionados com base no nome da revista, na descrição, área de estudo e categoria:

i. ACM Transactions on Information and System Security;

ii. Computers & Security;

iii. IEEE Security & Privacy;

iv. Information Management & Computer Security;

v. Information Security Journal: A Global Perspective;

vi. Infosecurity Today;

vii. International Journal of Information Security and Privacy;

viii. Journal of Information Privacy and Security;

ix. Journal of Information System Security.

Considerando o escopo da pesquisa, os artigos foram preferencialmente pesquisados nas seguintes bases de dados levando-se em conta o vasto conjunto de artigos publicados relacionados com o assunto em estudo:

i. ACM Transactions on Information and System Security;

ii. Computers & Security;

iii. IEEE Security & Privacy;

Foram também realizadas consultas na Web of Science (http://workinfo.com), Scopus (http://www.scopus.com) e J. Stor (http://www.jstor.org).

Os artigos foram selecionados a partir do título, resumo e palavras-chave. Destas fontes, foi dada prioridade a referências com menos de cinco anos. Foram selecionados os trabalhos levando-se em conta critérios como adequação ao tema do presente trabalho, relevância das publicações, produção acadêmica dos autores e citações de seus trabalhos.

A partir do material pesquisado, foi realizada uma leitura preliminar nos resumos dos artigos, no sentido de selecionar aqueles que se encontram dentro do objetivo da pesquisa. Assim, a revisão da literatura foi realizada numa primeira fase, em função do título, do resumo e nas palavras-chave, mas, em alguns casos, devido à inadequação do conteúdo do artigo com a temática em investigação, não foi integrada na bibliografia. Como resultado da pesquisa da temática em investigação, foi elaborada uma matriz de conceitos-chave por referências bibliográficas (Webster & Watson, 2002) , conforme relacionado na Tabela 1.

Referências	**Conceitos Chave**			
	Definições e conceitos Outsourcing	**Segurança e Riscos em Outsourcing**	**Requisitos de Outsourcing**	**Auditagem em Outsourcing**
Abbott, LJ et al. 2007.				●
Armbrust, M et al. 2009.	●			
Brandas, Claudia. 2010.				●
Chen, Y et al. 2010	●	●		
Hamlen et al. 2013	●	●		
Jensen, Meiko et al. 2009	●			
Karyda, Maria et al. 2006.	●	●	●	
Khajeh-Hosseini et al. 2010	●			
M Kanchana et al. 2013			●	●
Mansfield-Devine, Steve. 2008.		●		
Mouratidis et al. 2013.	●	●		
Pereira et al., 2011		●	●	
Randeree, Kishore, & Rao, 2005		●	●	
Tafti, Mohammed H.A. 2013		●	●	
Weinhardt, Christof et al. 2009	●			

Tabela 1 - Matriz de referência autor x conceitos chave

3. *Outsourcing* de SSI – Um serviço em expansão

O crescimento da Internet combinado com uma dependência tecnológica cada vez maior por parte das empresas aumentou a vulnerabilidade dos Sistemas de Informações (SI), razão pela qual se tornou essencial voltar o foco para a implementação da segurança dos SI. Os custos elevados com a infraestrutura de equipamentos de proteção, conjuntamente com a necessidade

de profissionais altamente especializado em Segurança das Informações, tem levado às organizações a busca de alternativas de terceirização deste serviço (Randeree et al., 2005).

Esta contratação passa a ser interessante para as empresas, devido a ser uma área de alta especialização devido as tarefas de segurança exigirem recursos significativos em termos de recursos humanos e infraestrutura. Assim, as empresas especializadas em segurança fornecem serviços de segurança para as organizações que não possuem esse conhecimento, ou não possuem recursos suficientes para executar essas funções internamente (Karyda, Mitrou, & Quirchmayr, 2006).

Um outro fato que justifica a busca pelas organizações pela terceirização da segurança é a dificuldade de atrair e reter funcionários qualificados e experientes nesta área, bem como os seus custos elevados. Além disso, a área da segurança sempre está evoluindo e se atualizando tecnologicamente, tornando-se necessário a aquisição contínua de novos equipamentos e softwares e elevado custo de treinamento e capacitação do pessoal especializado na área (Karyda et al., 2006).

Os serviços de SSI que as organizações são mais propensas a terceirizar incluem as seguintes atividades descritas na Tabela 2 e 3 (Karyda et al., 2006).

A utilização da Internet aliada a uma crescente terceirização das operações de suporte pelas organizações resulta em aumento de vulnerabilidades de segurança. Portanto, é desejável que a preocupação com a SSI seja direcionada para uma empresa especialista nesta área. Entretanto, é de fundamental importância que a empresa que preste esse serviço seja confiável o suficiente para presta-lo de forma segura e eficiente (Hamlen & Thuraisingham, 2013).

Antivírus de *e-mail* e filtragem de *Spam*	Gerenciamento de redes virtuais (*Vpn)*
Testes de invasão e detecção de intrusões	Gerenciamento de acesso ao usuário
Configuração de *Firewall*	Classificação de dados
Gerenciamento de servidores	Planos de contigências
Monitoramento de rede	Planos de continuidade
Treinamento de usuários em Segurança	Recuperação de desastres
Atualizações de Segurança	

Tabela 2 – Lista de serviços diretos mais ofertados em SSI

Escopo	Serviço
Análise de segurança	Verifica o nível de segurança atual da infraestrutura de tecnologia da informação do cliente, emitindo relatório detalhado das vulnerabilidades encontradas e quais ações devem ser tomadas para sua correção;
Definição e implementação de política de segurança	A política de segurança a ser adotada pela organização é definida com base no levantamento de suas necessidades de negócio, principais aplicações, infraestrutura de servidores e comunicação (local e distribuída, incluindo Internet, Intranet e Extranets), estrutura organizacional e análise de vulnerabilidades. Essa política envolve processos, pessoas e ferramentas num esforço coordenado para dotar a organização do nível de segurança adequado a seu negócio
Fornecimento de ferramentas de segurança	O fornecimento e configuração das principais ferramentas de segurança necessárias às organizações (Antivírus, firewall, ferramentas de análise de intrusos (IDS), inspeção e controle de conteúdo, análise de vulnerabilidades, redes privativas virtuais (VPN); Planos de contingência. Elaboração do plano de contingência contendo a definição de procedimentos para recuperação de informações e de continuidade operacional. Os planos podem incluir simulações de desastres para garantir a adequação do plano as aplicações instaladas.

Tabela 3 - Serviços de consultoria mais ofertados em SSI

4. A importancia da auditoria de *Outsourcing* de SSI

Os contratos entre fornecedores de serviços de TI e clientes normalmente é definido através de um documento denominado *Service Level Agreement* (SLA), ou traduzindo para o idioma português: “Acordo de Nível de Serviço”, uma nomeclatura de padrão internacional.

Um SLA especifica em geral e frequentemente em termos mensuráveis, quais são serviços que o fornecedor vai prestar e em que padrão de qualidade. SLAs são definidos no início de qualquer relação de contratação de serviços de TI e são usados para mensurar e monitorar o desempenho da prestação do serviço por um fornecedor (Armbrust, Fox, Griffith, & Joseph, 2009).

Para proceder a verificação da conformidade dos serviços de TI no ambiente computacional, são realizadas auditorias. A auditoria foca sua atividade prioritariamente na avaliação dos processos de governança, gestão de riscos e de forma complementar, na avaliação das principais atividades, processos e produtos da organização, especialmente aqueles considerados vitais para se atingir os objetivos estratégicos.

Com esse foco, a auditoria visa proporcionar relativa segurança às partes interessadas. A auditoria de sistemas, portanto, envolve a avaliação dos Sistemas de Informações e dos recursos tecnológicos que englobam o processo de geração, guarda e disponibilização da informação.

Neste sentido, uma das principais funções da auditoria de TI é a de promover a adequação, revisão, avaliação e recomendações para o aprimoramento dos controles internos dos Sistemas de Informações (SI) da empresa, bem como avaliar a utilização dos recursos humanos, materiais e tecnológicos envolvidos no processamento dos mesmos (M. Kanchana, Sk. Nazar Hussain, Kumar, & Praveen, 2013).

A análise dos riscos e as estratégias de mitigação dos riscos de forma eficiente e eficaz em uma empresa que terceirizou alguns ou todos os seus serviços de TI só pode ser realizado com o apoio de auditorias de Outsourcing (Brandas, 2010). Estas auditorias tornam-se necessárias e são fundamentais para se verificar:

i. O estabelecimento de controles adequados;

ii. A garantia do cumprimento das cláusulas contratuais e que os serviços são prestados em boas condições; A eficácia da gestão de segurança;A garantia dos processos de negócio da organização.

Desta forma, o foco da auditoria de Outsourcing é o de elaborar um diagnóstico independente do processo de terceirização de TI e relacioná-las com a realização dos objetivos da terceirização através da verificação do cumprimento dos acordos, da aferição, do faturamento e da correção das instabilidades identificadas. (Abbott, Parker, Peters, & Rama, 2007).

Os processos envolvidos na auditoria de Outsourcing desencadearam uma série de mudanças nas normas de auditoria que especificam orientações de auditoria para todas as partes envolvidas. Além disso, este tipo de auditoria pode ser estendido do fornecedor de serviços para o cliente, a fim de que possa atender a conformidade dos procedimentos do cliente em relação ao fornecedor. Independentemente de sua posição, o auditor deve se concentrar em avaliar os riscos específicos da área auditada e testar a eficácia e eficiência dos controles implementados. (Brandas, 2010)

Desta forma, a auditoria de *Outsourcing* se reveste de grande importância neste contexto e deve ser considerada como parte integrante do processo de terceirização dos serviços de TI. Ao contratar forncedores de serviços de TI, deve-se prever também os custos relacionados com a auditoria destes serviços. Nesse sentido, constituem-se como objetivos da auditoria de Outsourcing a verificação da terceirização dos serviços descritos na Tabela 4 em relação a decisão do cliente consumidor de TI (Brandas, 2010).

Quesito	Objetivo
Administrativo	Avaliar o processo de decisão de terceirização
Técnico	Avaliação do serviço prestado pelo fornecedor
Custo benefício	Analise do custo e previsão contrato de terceirização
Conformidade	Verificação dos Acordos de Nível de Serviço (SLAs)
Orçamento	A alocação e execução do orçamento de TI
Comunicação	A comunicação e relacionamento do cliente com o provedor de serviços
Confiança	A confiança na execução do serviço
Riscos	O gerenciamento de riscos
Monitoramento	O monitoramento da relação do prestador de serviços com o cliente;
Resultados	O impacto na estratégia de TI do cliente

Tabela 4- Objetivos da auditoria de Outsourcing de TI em relação ao Cliente

Por outro lado, torna-se necessária também verificar a prestação de serviços de TI relacionados na Tabela 5 por parte do fornecedor (Brandas, 2010).

Quesito	Objetivo
Analise do Cliente	A análise dos requisitos do cliente na fase pré-contratual
Contratual	A análise técnica e jurídica do contrato de terceirização e os SLA
Ambiental	A verificação das ferramentas de controle e de segurança do ambiente
Conformidade	A análise do atendimento dos requisitos exigidos do cliente consumidor
Feed-back	A entrega de serviços e o registro da satisfação do cliente
Seleção de Pessoal	Os critérios de recrutamento e de seleção de pessoal especializado na área de TI
Monitoramento	O monitoramento do treinamento e do desempenho
Avaliação	Avaliação do grau de dependência do cliente em relação ao provedor

Tabela 5 - Objetivos da auditoria de Outsourcing de TI em relação a prestação de serviços pelo fornecedor

5. Os riscos envolvidos na prestação de *Outsourcing* de SSI

Na análise da revisão da literatura, percebe-se que o tratamento do aspecto relacionado ao risco é bastante abordado no âmbito do Outsourcing da TI. Alguns riscos são reforçados pelos pesquisadores e devem ser considerados relevantes para a prestação de serviço de SSI.

Com base em resultados obtidos através de estudos de caso, (Khajeh-Hosseini, Greenwood, & Sommerville, 2010) concluíram que a migração de Outsourcing de TI traz benefícios em relação ao desenvolvimento dos negócios da empresa e que não houveram mudanças significativas em relação a gestão de projetos. Entretanto, estes autores ressaltam que a maior fonte de riscos está relacionada a deterioração do atendimento ao cliente e da qualidade do serviço prestado pelo provedor, além da dependência que se cria em relação a este.

Um dos componentes mais importantes de risco em relação a esta prestação de serviço é o da formulação de um contrato mal elaborado (SLA). Embora seja crucial incluir SLAs em um contrato de terceirização, defini-los pode ser uma proposição complicada. Detalhes como suposições incorretas sobre o desempenho de um fornecedor, tempo médio de resposta para a resolução de bilhetes de help desk, por exemplo, podem não se mostrar efetivos na prática. Um baixo nível de desempenho pode ser facilmente camuflado por um SLA que se baseia exclusivamente em médias. (Tafti, 2013)

Os riscos associados com as questões de privacidade e segurança podem vir a ser grandes se não for tratada no início de decisões em matéria de terceirização de TI. Entre eles, incidentes já ocorridos demonstram os riscos alarmantes envolvidos quando as empresas envolvidas em Outsourcing de TI não tomam medidas proativas para proteger a privacidade individual e os ativos de informação de seus clientes. (Tafti, 2013)

Os riscos associados com a replicação da terceirização envolvendo a contratação de terceiros pelas prestadoras de serviços em regime de Outsourcing é grande. Neste caso, a questão da responsabilidade passa a envolver mais de um agente. Há consequências graves para a terceirização de atividades que deveriam ser realizadas internamente no âmbito do prestador de serviço de *Outsourcing*. Um exemplo disto pode ser a rotatividade de integrantes da equipe de especialistas em Segurança de TI, levando a casos em que a prestadora tenha de recorrer a profissionais externos ou a outras empresas para suprir esta demanda. (Tafti, 2013)

Outra consideração relevante é a perda de competências do pessoal de TI a nível interno da organização, o que pode resultar em dificuldades no caso se ter que fazer um *insourcing* (trazer o serviço de volta), caso o provedor dos serviços não prestar serviços inadequados ou não atender as necessidades do cliente. (Khajeh-Hosseini et al., 2010).

6. Os pontos críticos para verificação da conformidade do serviço de Outsourcing de SSI

Com base na revisão da literatura, um fator de consenso entre os investigaores, é considerar que encontrar o fornecedor certo e firmar acordos viáveis de SLAs são de suma importância.

Reconhecendo que o sucesso não depende apenas de aspectos contratuais, mas também na relação estabelecida entre as partes, foram citados também outros fatores importantes na relação dos fornecedores de serviço de SSI, tais como a participação, a comunicação, a compreensão mútua, compartilhamento de informações, apoio à gestão e coordenação (Pereira et al., 2011).

Para estabelecer requisitos contratuais adequados, torna-se necessário adotar tecnologias capazes de coletar os dados necessários para informar as decisões de risco, como o acesso de uso, controles de segurança, localização e outros referenciais relativos a utilização do serviço de *Outsource* de TI. Além disso, é desejável que os consumidores disponham de métricas e controles para auxiliar a gestão do uso da nuvem em cumprimento aos SLAs acordados entre as partes (Weinhardt et al., 2009).

Pelo fato dos auditores ainda não estarem familiarizados com a o nível de complexidade do outsourcing de TI, inclusive com a popularização da Computação em Nuvem e sobre o que auditarem, é necessário levar em consideração os seguintes aspectos (Jensen, Schwenk, Gruschka, Lo Iacono, & Ieee, 2009):

i. A aplicabilidade regulatória para uso em contratos de outsourcing de TI;

ii. A divisão das responsabilidades entre o provedor do serviço e o cliente acordado nos SLAs.

No decorrer do processo de migração para Outsourcing de TI, é importante que o provedor de serviços forneça os detalhes operacionais de como os dados do cliente será tratado, quais são os requisitos de segurança são aplicados e os planos de contingências previstos para o caso do sistema ficar comprometido (Mouratidis, Islam, Kalloniatis, & Gritzalis, 2013).

Em uma pesquisa realizada acerca dos fatores preponderantes na relação provedor-cliente na prestação de serviços terceiros de SSI, Pereira et al (2011) apontam a necessidade da verificação desta relação a partir dos seguintes critérios:

Em relação a escolha do provedor de SSI:

i. A compatibilidade entre as culturas organizacionais dos envolvidos na terceirização;

ii. A relação custo-benefício do serviço de SSI a partir do ponto de vista do cliente;

Em relação a credibilidade do prestador de serviços:

i. Prova da capacidade do prestador de serviços ISS para realizar o serviço em questão;

ii. Existência de um relacionamento de sucesso anterior na parceria entre o cliente e o prestador de serviços;

iii. Experiência do prestador de SSI na prestação deste tipo de serviço;

iv. Conhecimento por parte do prestador de serviços ISS do mercado em que o cliente opera;

v. Diversidade na oferta de produtos e serviços por parte do prestador de serviços;

vi. Apoio no cumprimento dos requisitos regulamentares ou legais por parte do prestador de serviço SSI.

Envolvendo a estruturação do relacionamento entre as partes:

i. Definição clara das responsabilidades do cliente e do prestador de serviços;

ii. Garantia de não-repúdio das tarefas realizadas pelo prestador de serviços ISS;

iii. Salvaguarda de confidencialidade das informações do cliente pelo prestador de serviços de SSI.

Na construção do relacionamento mútuo fornecedor-cliente:

i. Compromisso pelo prestador de serviços para melhorar continuamente o serviço prestado;

ii. Manutenção dos canais de comunicação entre o cliente e o prestador;

iii. Existência de um plano de continuidade de negócios;

iv. Confiança no relacionamento entre as partes envolvidas;

v. Gestão compartilhada da Segurança entre o cliente e o prestador de serviços de SSI

Da gestão do serviço prestado:

i. Capacidade para avaliar o nível de serviço realizado, bem como a avaliação da execução e gestão deste;

ii. Competência dos recursos humanos do prestador de serviços;

iii. Identificação inequívoca das tarefas que estão sendo executadas dentro do âmbito do serviço de SSI;

Os clientes não devem confiar cegamente nas alegações do provedor sobre a robustez de suas ferramentas e de sua infraestrutura operacional, devendo serem esclarecidos todos os detalhes das políticas adotadas referentes as tecnologias empregadas. O provedor deve ser investigado pela organização, sendo a ele solicitado todos os pontos importantes e relevantes a serem considerados para prestação dos seus serviços.(Mouratidis et al., 2013).

Fator importante é a realização de testes do ambiente do provedor. Simulações e testes de consistência devem ser exaustivamente realizados para testar a confiabilidade dos serviços regularmente. De preferência, estes testes devem ser realizados por empresas terceiras (Mansfield-Devine, 2008).

7. Conclusões

Este trabalho objetiva dar ciência a partir de uma revisão da literatura, sobre os riscos e fatores críticos de sucesso envolvidos na contratação de serviços terceirizados (*Outsource*) na área da Segurança da Informação (SSI), promovendo uma reflexão de como pode ser viabilizada a auditoria específica para este tipo de serviço.

Na revisão da literatura, verificou-se que embora diversos pesquisadores tratem de assuntos relacionados a problemática da confiabilidade envolvida na relação entre a as partes envolvidas onde se prioriza os riscos envolvidos neste tipo de serviço, nenhum material relacionado a ferramentas de auditagem, como métricas estabelecidas foi encontrado para a área específica de Outsourcing de SSI. Percebe-se que as questões envolvendo a relação de confiança exigida nesta modalidade de serviço foi condiserada como sendo a mais relevante.

O aprofundamento deste estudo na perspectiva de trabalhos futuros, abre possibilidades para novas investigações, como por exemplo, o do desenvolvimento de um framework para auditoria da prestação do serviço de Segurança da Informação (SSI) sob a forma de *Outsourcing*, a partir do estabelecimento de padrões de conduta e de ferramental (métricas e formas de aferição) adequado a este tipo de serviço.

8. Referências

Abbott, L., Parker, S., Peters, G., & Rama, D. (2007). Corporate governance, audit quality, and the Sarbanes-Oxley Act: Evidence from internal audit outsourcing. *The Accounting Review*. Retrieved from http://www.aaajournals.org/doi/abs/10.2308/accr.2007.82.4.803

Armbrust, M., Fox, A., Griffith, R., & Joseph, A. (2009). Above the clouds: A Berkeley view of cloud computing. *University of California, Berkeley, Tech. Rep. UCB*, 07–013. Retrieved from http://scholar.google.com/scholar?q=intitle:Above+the+clouds:+A+Berkeley+view+of+cloud+computing#0

Brandas, C. (2010). Risks and Audit Objectives for IT Outsourcing. *Informatica Economica*, *14*(1), 113–119.

Hamlen, K. W., & Thuraisingham, B. (2013). Data security services, solutions and standards for outsourcing. *Computer Standards and Interfaces*, *35*(1), 1–5.

Jensen, M., Schwenk, J., Gruschka, N., Lo Iacono, L., & Ieee. (2009). *On Technical Security Issues in Cloud Computing*. *Cloud: 2009 Ieee International Conference on Cloud Computing*. Retrieved from <Go to ISI>://WOS:000275314400015

Karyda, M., Mitrou, E., & Quirchmayr, G. (2006). A framework for outsourcing IS/IT security services. *Information Management & Computer Security*.

Khajeh-Hosseini, A., Greenwood, D., & Sommerville, I. (2010). Cloud Migration: A Case Study of Migrating an Enterprise IT System to IaaS. In *2010 IEEE 3rd International Conference on Cloud Computing* (pp. 450–457). IEEE. doi:10.1109/CLOUD.2010.37

M. Kanchana, Sk. Nazar Hussain, Kumar, M. K., & Praveen, C. (2013). Preserving Audit of Secure Data Storage Services in Cloud Computing. *International Journal of Advanced Research in Computer Science*, *4*(5), 70–73. Retrieved from http://www.ijarcs.info/?wicket:interface=:3::::

Mansfield-Devine, S. (2008). Danger in the clouds. *Network Security*, *2008*(12), 9–11. doi:10.1016/S1353-4858(08)70140-5

Mouratidis, H., Islam, S., Kalloniatis, C., & Gritzalis, S. (2013). A framework to support selection of cloud providers based on security and privacy requirements. *Journal of Systems and Software*, *86*(9), 2276–2293. doi:10.1016/j.jss.2013.03.011

Pereira, L., de Sá-Soares, F., & Caldeira, M. (2011). Information systems security outsourcing key issues: a service providers' perspective. Retrieved from http://www.repository.utl.pt/handle/10400.5/3080

Randeree, E., Kishore, R., & Rao, H. R. (2005). Managed Security Service Providers: Issues in Outsourcing Security. In *Proceedings of the Eleventh Americas Conference on Information Systems* (pp. 2731–2734).

Tafti, M. H. A. (2013). Risks factors associated with offshore IT outsourcing. Retrieved from http://www.emeraldinsight.com/doi/full/10.1108/02635570510599940

Webster, J., & Watson, R. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. Retrieved January 23, 2015, from https://www.google.com.br/#q=writing+a+literature+review+webster+and+watson+2002

Weinhardt, C., Anandasivam, A., Blau, B., Borissov, N., Meinl, T., Michalk, W., & Stößer, J. (2009). Cloud Computing – A Classification, Business Models, and Research Directions. *Business & Information Systems Engineering*, *1*(5), 391–399. doi:10.1007/s12599-009-0071-2

Portuguese Public Procurement Governance Model

Isabel da Rosa[1], José Tribolet[2], Miguel Mira da Silva[3]

1) Universidade de Lisboa, Portugal

Darosa.isabel@gmail.com

2) Universidade de Lisboa, Portugal
Jose.tribolet@inesc.pt

2) Universidade de Lisboa, Portugal

Mms@tecnico.ulisboa.pt

Abstract

Public Procurement in Europe represents 19% of GDP. The 2014 European Directives on Public Procurement highlight the need for governance measures in Member States.

Portugal has been considered a pioneer in electronic Public Procurement and, following Europe's last guidelines, is dedicating attention to a formal governance model set up.

A governance model is fundamental to enhancing integrity in Public Procurement, supporting sustainable and inclusive growth, and reinforcing citizens' trust.

This study aims to analyse the evolution of Public Procurement governance in the Portuguese state and design a model within the Enterprise Governance context.

Our proposal builds on the Enterprise Ontology principles and the design uses DEMO methodology to represent the model. It mainly aims to address the lack of articulation among key public entities and to increase alignment between government strategic objectives and public entities action plans.

The result of our study observes the recent OECD recommendations on Public Procurement.

Keywords: Enterprise Ontology, Enterprise Governance, DEMO, Public Procurement

1. Introduction

Within the scope of the European Union Single Market principle, public procurement has been considered as a determinant tool [EC Grow Directorate 2015].

Portugal has been influenced by European Union (EU) policies and public procurement has changed dramatically over the last 15 years with the creation of public entities and the reorganization of services and responsibilities of the National Public Purchasing System

(NPPS). In Figure 1 it is illustrated the main milestones regarding the creation of public entities that introduced significant changes in the NPPS.

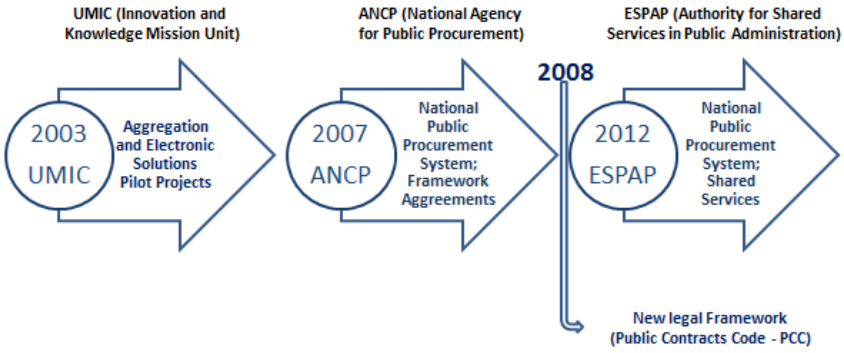


Figure 1 - Public entities related with National Public Procurement System

Some of these changes represented a profound revolution in public procedures with the onset of aggregation of purchases for several public entities, the conduct of electronic auctions and the conclusion of framework agreements for categories common to the whole public administration.

All these changes were achieved through the individual effort of key public bodies and the only articulation between entities derived from the enforcement of some measures. With the growing complexity of the system, the lack of any explicit governance solution for public procurement all over this period is surprising. However, and because of that, it is not surprising that organizations stood for different interests and different positions and the absence of orchestration and integrity led to difficulties in daily life of public procurers and poor results.

Without a governance model, public procurement lacks unity and integration, with no alignment among organizations and among people. Consequently, together with the growing complexity of the Information Technology (IT), it will be very difficult to implement any changes in the existing system.

1.1. Research Focus

In recent years, in most international and high level organisations, such as OECD (Organisation for Economic Cooperation and Development), UN (United Nations) and European Commission (EC), and the academia, governance has gained growing attention and particular relevance in development and sustainability of countries and enterprises.

According to the last EU directives on public procurement, from 2014, Member States (MS) will have to transpose them into national laws until April 2016. This is the first time that MSs have the need for a set of obligations under the governance scope [Commission 2014].

Portugal has not implemented, so far, a governance scheme to ensure the holistic and integrative orchestration of all key entities. In this case, public entities will have to handle EU obligations on an individual basis, which in many cases can be very difficult, especially if the result depends on more than one organisation. Usually several problems occur, such as: late responses; deviations from the expected results; lack of information integrity; dilution of duties and failure to assume responsibilities.

This study presents a brief summary of the evolution of key public entities in the field of Public Procurement over the last 12 years in Portugal and proposes a governance model based on the Enterprise Ontology and Enterprise Governance background to address the abovementioned problems.

1.2. Research Methodology

Our research methodology is based on Design Science Research (DSR), which comprises two complementary phases: build and evaluate. Unlike behaviour research, design-oriented research builds a "to-be" conception and then seeks to build the system according to the defined model taking into account restrictions and limitations [Österle, et al. 2011]. Building and evaluation of artefacts designed to meet identified business needs are the foundation of research in design science [Hevner, et al. 2004].

The building process covers two stages, whereas the evaluation process comprises only one. Figure 2 describes the governance model construction steps in each phase. This research approach has already been used in other research papers [Pereira & Silva 2012], [De Haes & Grembergen 2008] and [Vicente & Silva 2011].

BUILD		EVALUATE
Construct Definition	**Model Construction**	**Evaluation**
Domain definition	Analysing the relation between actors and roles	Pratical evaluation - Experts interviews
Model objective definition	Integrate constructs	
Definition of the model components	DEMO representation	

Figure 2 - Design Research Methodology stages

We also followed the four principles a design-oriented IS research must observe [Österle et al. 2011]:

- **Abstraction**. This paper proposes a governance model for public procurement in the Portuguese Public Administration. This model can be used in any MS, because the level of abstraction is guaranteed under the three dimensions of the model: type of structure,

functions and organisational framework. In this study we also describe its application to the present situation in Portugal.

- **Originality**. The proposed artefact is not present in the body of knowledge of the domain.
- **Justification**. The method proposed to evaluate the artefact is the first step to understand the viability of our proposal.
- **Benefit**. It is necessary to implement formal solutions to deal with public procurement governance in the near future. This necessity is stated by many national and international organisations, and is imposed by the 2014 European directives. Growing complexity requires the development and implementation of strategic, evidence-based and innovative policies in order to strengthen public governance respond effectively to diverse and disruptive economic, social and environmental challenges and deliver on government's commitments to citizens.

A design artefact is complete and effective when it meets the requirements and constraints of the problem it is meant to solve. In this paper we evaluated our artefacts through interviews with experienced specialists from public procurement entities, at national and international level.

1.3. Structure of this Paper

This paper is divided into seven sections identified by a number, as follows: 1. Introduction; 2. Problem Definition; 3. Related Work; 4. Proposal Presentation; 5. Proposal Demonstration; 6. Proposal Evaluation; and 7. Conclusions.

2. Problem Definition

The European Directives define procurement as the acquisition, by means of a public contract, of works, supplies or services, by one or more contracting authorities from economic operators chosen by those contracting authorities, whether or not the works, supplies or services are intended for a public purpose [Commission 2014].

Public Procurement in Portugal is expected to undergo a profound change, following the publication of three legal instruments that will change the implemented processes, the supporting systems and the strategic objectives defined until now.

These instruments are: (1) the law governing the operation of electronic platforms in Portugal, (2) the updated code that regulates public procurement, by transposing European Directives 2014/24/EU, 2014/25/EU (repealing Directives 2004/17/EC and 2004/18/EC) and 2014/23/EU

on the award of concession contracts, adopted by the European Parliament and the Council of the European Union, and (3) the national system definition for electronic invoice in public procurement following the transposition of Directive 2014/55/EU.

Particularly in the cases described in (1) and (3), these instruments will bring about significant changes in the information system architecture that had been implemented since November 2009, when electronic public procurement became mandatory.

In both cases, and considering the architectures under design, a set of public entities will have a critical role in the implementation of new services and the development of new IT solutions. The overall success will depend not only on all the individual organisations, actions and performances but also, and above all, on the integration of those individual efforts and results.

References to the need for governance in public procurement in Directive 2014/24/EU only concern a reduced set of principles, and they derive from the perspective of the European Union, as a central structure that requires coordination in the creation of the Single Market.

Joining, both perspectives, national transformation processes and European Union requirements, a national governance scheme should be put in place to accommodate these three questions:

Q1: How to ensure data consistency and quality in the formal reports to international organisations regarding national public procurement performance?

Q2: How to ensure unity and integration of all key stakeholders in public procurement solutions?

Q3: How to ensure that action plans of public entities are in line with government strategic objectives for public procurement?

The evolution of Public Procurement over the last years shows that the system has managed to survive without any governance model. Although this might be considered an amazing fact, it is also true that this was possible due to the evolution of public procurement, which recently gained some projection even within the government concerns. Indeed, there have been significant changes in the nature of the procedures and in the way they were conducted, but also in the systems that supported them all. Their complexity has increased exponentially since 2009, when the national electronic framework was set up.

As illustrated in Figure 3, despite the growing relevance of public procurement, there has been no unity orchestration of all parties involved.

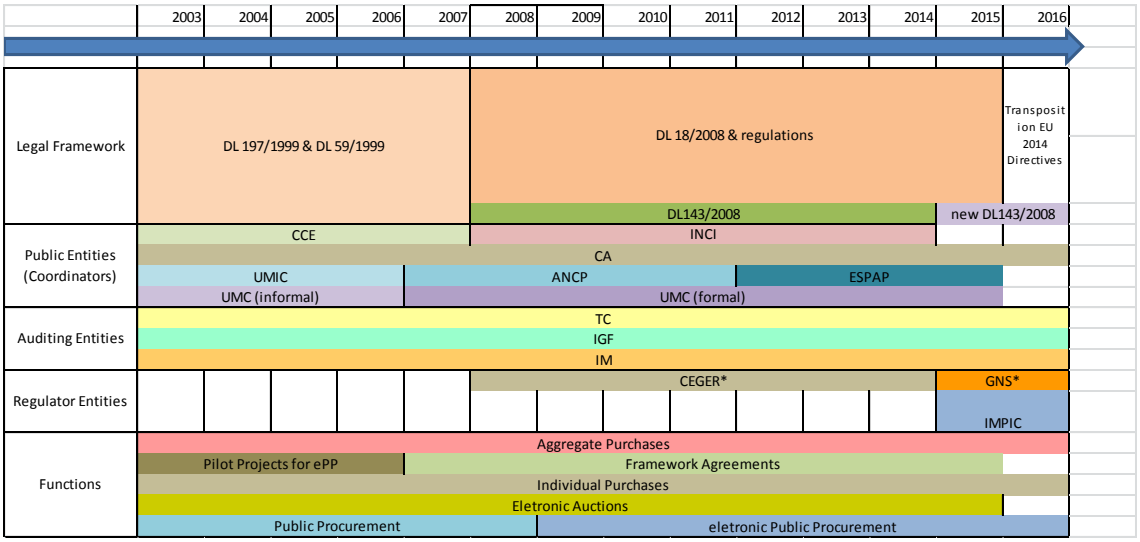


Figure 3 - Public procurement responsibilities evolution

Changes in existing public bodies have been a constant source of instability.

Recent years have frequently seen the merger and termination of several bodies and the creation of new ones, and this is not merely an aspect of public procurement management, it also reflects the need to reorganize the state and decrease its weight in public expenditure.

The next challenges, and most particularly in the specific case of interoperability in the pre-award phase, will require highly complex technical solutions. Whereas change was once a contained transactional event (and easier to manage), it is now more open-ended, radical, complex, personal, and continuous. “Transformation” has emerged as the new type of change, and it is by far the most prevalent and complex type occurring in organizations today. In general, leaders do not understand transformational change or how to lead it, which is causing virtually all the change-related problems they are now facing [Anderson & Anderson 2010].

In Public Administration this reality is exacerbated by the fact that their universe is not restricted to a single organization, but rather comprises a wide range of organizations that compete with each other and sometimes have overlapping responsibilities and diverging interests.

We are witnessing a time when leaders and managers are becoming more aware of the need for a structure that leads institutions throughout the transformation process, failing which the objectives pursued will probably not be achieved, with the risk of causing tremendous resistance and burnout in people, loss of employee morale, and turmoil in the culture of the organizations [Anderson & Anderson 2010].

A governance solution is therefore needed in order to answer those questions and promote a single, integrated view of public procurement in Portugal.

3. Related Work

This research was based on a context outlined by three references: Ontological Model, Enterprise Governance and Business Transformation Management.

The reason for choosing these references is the fact that some concepts referred to in this study have different definitions and different perspectives in literature and in several sciences. This is mainly the case for governance, our major subject.

Governance

For this study, we based our approach on the governance concept advocated by Hoogervorst because it is clearly linked to transformation and makes a distinction between administration and governance, which in many concepts is not easy to understand.

As referred by Hoogervorst, the term 'governance' stems from the Latin word *gubernáre* (in turn borrowed from the Greek language), which means to control or steer, in the original meaning, the steering of a ship. Governance can thus be associated with guiding and giving direction. It is important to distinguish governance from management. The latter term has its origin in the Latin word *manus* (hand). Both terms are relevant within the enterprise context. To differentiate management from governance, we will consider 'management' from an operational, implementing point of view, and will use the term 'governance' in the context of enterprise change. In other words, governance guides developments that lead to a new (or partly new) enterprise [Hoogervorst 2012].

Governance provides the guardrails and steering wheel to keep the transformation on the right track [Westerman, Bonnet & McAfee 2014].

In fact, for the last three decades, governance has been a prominent subject in Public Administration and we can find different definitions depending on the context [Frederickson 2007]. Besides that, there are many different types of governance and we need to make a brief review in order to understand each one of them and decide which governance we will focus on:

- Corporate Governance – is the system by which organizations are directed and controlled [De Haes & Grembergen 2008]; it is the responsibility delegated by stakeholders and the public, defined by legislators and regulators, and shared by boards, in some measure, with managers [Webb et al. 2006]. Briefly stated, corporate governance regards directions for internal (managerial) control that aim to arrange

financial transparency and prudent financial conduct in order to safeguard the interests of shareholders [Coley et al. 2005].

- Enterprise Governance – is a set of responsibilities and practices exercised by the board and executive managers, with the goal of providing strategic direction, ensuring that plans and objectives are achieved, assessing that risks are proactively managed, and assuring that the enterprise's resources are used responsibly [De Haes & Grembergen 2008]. Enterprise governance is the overarching type of governance that encompasses corporate and IT governance [Hoogervorst 2012].

- IT Governance (ITG) – literature has demonstrated a lack of a clear shared understanding of the term ITG [Pereira & Silva 2012]. Pereira identified several ITG definitions in literature, with minor differences from where we select two: ITG is the responsibility of the board of directors and executive management. It is an integral part of enterprise governance and consists in the leadership and organizational structures and processes which ensure that organization's IT sustains and extends the organization's strategies and objectives [Information Technology Governance Institute 2004]; the strategic alignment of IT with business such that maximum business value is achieved through the development and maintenance of effective IT control and accountability, performance management, and risk management [Webb et al. 2006].

Ontological Model

Ontological theories deal with the nature of things. Within the discipline of enterprise engineering, there is a particular concern with cause-effect relationships in systems.

Ontological theories are valuated by their soundness and appropriateness.

The Enterprise Ontology theory produces conceptual models that are [Dietz 2006]:

- Coherent – the distinguished aspect models constitute a logical and truly integral whole;
- Comprehensive – all relevant issues are covered, the whole is complete;
- Consistent – the aspect models are free from contradictions or irregularities;
- Concise – no superfluous matters are contained in it, the whole is compact and succinct;
- Essential – the conceptual model shows only the essence of the enterprise (its deep structure) and abstracts from all realization and implementation issues.

DEMO (Design & Engineering Methodology for Organizations) is a methodology for modelling, (re)designing and (re)engineering organizations and networks of organizations [Dietz 2006]. DEMO allows modelling the Enterprise Ontology (EO) of an organization. Enterprise Ontology is defined as the fully implementation independent understanding of the essence of an organization. We have chosen EO because it can help us expand the expressiveness of the model description and, consequently, allow for a better alignment between expectations and perceptions. EO is a viable basis for practically dealing with organizational changes of all kinds [Institute 2015].

DEMO consists of four aspect models, represented by particular diagrams, tables and lists [Mendes 2013]:

- The Construction Model (CM) – specifies the identified transaction types and the associated actor roles, as well as the information links between the actor roles and the information bank. The CM specifies the composition, environment, and structure of an organization:
    - The Interaction Model (IAM) – shows the active influences between actor roles i.e. the execution of transactions;
    - The Interstriction Model (ISM) – shows the passive influences between actor roles, i.e., the taking into account of existing facts by an actor role when being active.
- The Process Model (PM) – contains, for every transaction type in the CM, the specific transaction pattern of the transaction type. The PM also contains the causal and conditional relationships between transactions. These relationships determine, in addition to the transaction patterns, the possible trajectories in the Coordination-world. In other words, the PM specifies the state space as well as the transition space of the Coordination-world;
- The Action Model (AM) – specifies the action rules that serve as guidelines for the actors in dealing with their agenda. For every coordination step, there is an action rule that guides how the performing actor role should respond to the reached status. At the ontological level of abstraction there is nothing below the AM;
- The State Model (SM) – specifies the state space of the P-world: the object classes and fact types, the result types, and the ontological coexistence rules. SM is the ideal starting point for developing and maintaining the data dictionary of an enterprise and it

simplifies the identification of business components (software components), based on the chunks of fact types around categories.

These models constitute the complete ontological model of Business-organization and subsequently represent the ontological model of the corresponding enterprise.

The Interaction Model (IAM) is expressed by the Actor Transaction Diagram (ATD) and the Transaction Result Table (TRT). These two high level artefacts are the ones we use in this paper to illustrate the essence of the proposed governance model.

Business Transformation Management

Within the scope of Business Transformation Management we paid particular attention to a recent methodology named Business Transformation Management Methodology (BTM^2) that aims to be different from other methodologies by assuring integration and holistic transformation management through a high level layer called Meta Management.

Meta Management is business-driven and value-oriented, and integrates three pillars, namely: management disciplines, transformation lifecycle and leadership.

In the case of leadership it is defended that the business transformation process is only successful if leaders are aware of their roles and communication and have established a solid culture.

In fact, in change management theory, an often neglected aspect of the process is the need to build a coalition [Jick & Peiperl 2002]. Though this appears implicitly in a number of change models, the dimensions of the team needed and their roles has been largely missing from major change conceptualizations [Uhl & Gollenia 2012].

BTM^2 aims to handle this by providing an organizational structure for the management roles (formal and informal) that will accompany the life cycle of the transformation process.

The essence of roles in an organizational structure is the reference to the model we are proposing.

4. Proposal

We propose the creation of a governance model for the Portuguese Public Procurement as a solution to deal with the problems experienced in the past (identified in section 2), as a reference to promote the continuous development of the system that has been implemented, and

as a solution to face the significant challenges implicit in the new strategic objectives of the electronic system.

The structure of the presentation of our proposal is in line with the structure of the methodology described in section 1. 2. Thus, we describe the construction of our model in two steps: the construct definition and the model construction of the building phase.

The construct definition step refers to the activities:

- Domain definition
- Model objective definition
- Definition of the model components

On the domain definition, the main concern is to delimit the scope, which in our study is the conception of a governance model for public procurement in Portugal, and therefore, is limited to public institutions that play a key role in guiding Public Procurement as a result of the Portuguese government policy and in accordance with the legal framework in force.

At the model objective definition level, we established the correlation with the problem identification in section 2. Our proposal seeks to provide an answer to the three questions in the problem set-up, which means:

Answer to Q1: To provide one single and accountable voice for national public procurement information that is able to collect and integrate data from all the relevant components of the national system;

Answer to Q2: To provide an orchestration scheme for the key public bodies related to public procurement;

Answer to Q3: To provide direction and enforce the articulation for organisations activities through alignment guidance.

Considering governance as the vehicle through which the enterprise stakeholders can participate for guiding and achieving an integrated construction model [Henriques, Tribolet, & Hoogervorst 2010], and considering Public Administration as a set of public entities with different objectives, interests, autonomy levels and development degrees we propose a model, in Figure 4, based on three components: structure, attributions and empowerment positioning.

The last component it is very critical and fundamental because the positioning inside organic structure of Public Administration will determine the authority level in the relation with other public entities.

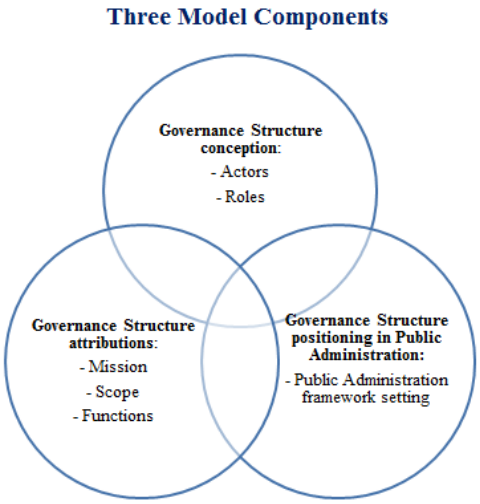


Figure 4 - Public procurement model components

Within the phase model construction we analysed the relation between actors and roles, we integrate all the components of the model and, finally, we use DEMO for representation purposes.

The structure we propose encloses some of the principles defined in BTM[2] for roles definition within the Meta Management layer.

BTM[2] advocates that roles should be defined at 3 levels: strategy, program and project, which correspond respectively to the concern levels of "why", "what" and "how" in the process to be implemented [Uhl & Gollenia 2012].

By analogy, the proposed structure is multi-level as shown in Figure 5.

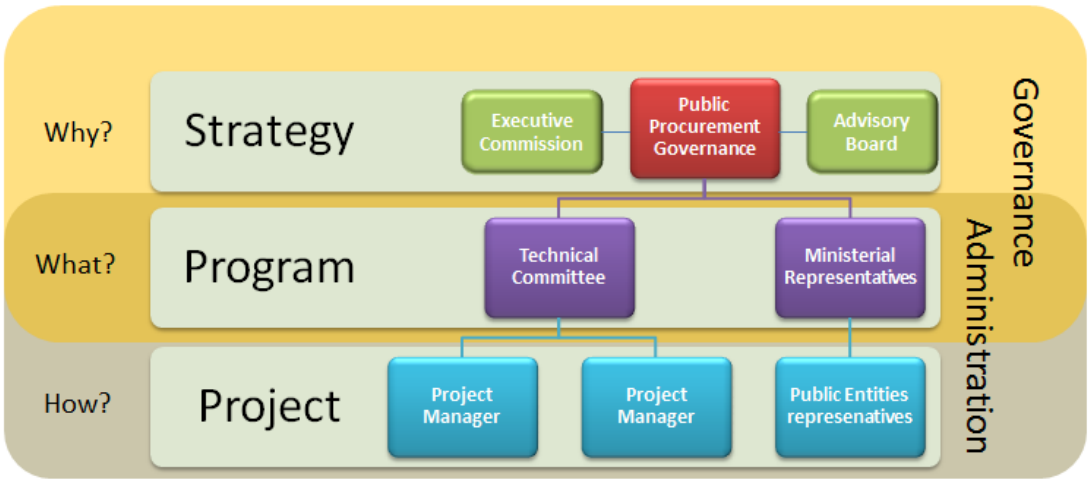


Figure 5 – Governance structure

The governance structure has four actors: the executive commission, the advisory board, the technical committee and the ministerial representatives group. The executive commission is the top level of the structure and the one that articulates with the government.

The responsibilities of the governance structure derive from the objective that is inherent to its creation and is driving the transformation processes. We organize those responsibilities in four dimensions as represented in Figure 6.

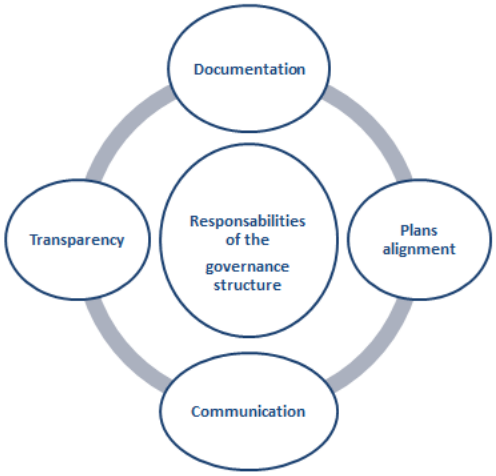


Figure 6 – Responsibilities of the governance structure

In the first case, the documentation dimension, the responsibility of the governance structure is to create a repository of documents on the transformation process, which should include plans, reports, diagrams, stakeholders, KPIs, etc., and to establish the rules for managing that repository. This is a fundamental responsibility because the effectiveness of the transformation process will depend on the current level of knowledge (representation of the situation AS-IS) and the degree of knowledge of what is to be achieved (representing the TO-BE situation) and the effective coordination of the steps that mediate both states.

With respect to the second dimension, plans alignment aims to ensure that the activity of the various public bodies involved in the transformation process is properly aligned with the strategic objectives set by the government and underlying the legal framework.

Since the activity of public entities is embodied in their activity plans and projects, it is critical to ensure the alignment between strategy and operation.

Regarding the communication dimension, the goal is to ensure the structure and the rules that allow the dissemination of information (top-down perspective) and to collect feedback from all the stakeholders involved (bottom-up perspective).

Finally, and very important for the sustainability of the governance function, we have transparency and accountability. Both, the results of the governance structure activity and the progress of the transformation process must be published.

This last dimension is particularly relevant in Public Administration where decisions and activities might affect organizations, citizens and the economy in general.

In the public procurement field there is a direct impact on the socio-economic system of the country.

In order to formalize the proposal, we present a DEMO white-box model for the public procurement governance. A white-box model is a direct conceptualization of the ontological system definition and captures the construction and the operation of a system, while abstracting from implementation details [Dietz 2006].

The first model is the Construction Model and is described by the Transaction Result Table (TRT), the Result Structure Chart and the Actor Transaction Diagram (ATD). In order to achieve this model we applied the guidelines defined in [Dietz 2006] to a text that describes the function of the governance structure. Within this phase, we made a distinction between the Ontological, Infological and Datalogical actions described, as referred to in the Enterprise Ontology Distinction axiom.

The next step is the identification of C-acts/facts, P-acts/facts and actor roles, using the Performa (Ontological) items identified in the previous step.

This step is called Coordination-Actors-Production Analysis but we reduced its complexity when compared with other methodologies, because from now on we will only consider the Ontological actions identified in this step.

After these analyses, we identified the transactions by clustering the identified C-acts/facts and P-acts/facts in the so called Transaction Pattern Synthesis. The Transaction axiom can be helpful in this step, because it guarantees that each P-act/fact or C-act/fact previously identified corresponds to a complete transaction. Then, for each identified transaction type, the result type (i.e., the Production fact created) was formulated. The result is presented in Figure 7 below, called Transaction Result Table.

Transaction	Result	Executed by
T01 – TPM Start	R01 – TPM started	Executive Commission
T02 – TPM Impacts Analysis	R02 – TPM impact analyzed	Advisory Board
T03 – TPM and SO Alignment Analysis	R03 – TPM and SO alignment analyzed	Executive Commission
T04 – SO Start	R04 – SO started	Government
T05 – SO End	R05 – SO ended	Government
T06 – PGM Start	R06 – PGM started	Technical Committee
T07 – PGM End	R07 – PGM ended	Technical Committee
T08 – Program KPIs Definition	R08 – Program PG KPIs defined	Executive Commission
T09 – PM Repository Creation	R09 – Repository of PM created	Technical Committee
T10 – Program Viability Evaluation	R10 – Program PG viability evaluated	Ministerial Representatives
T11 – PJM Definition	R11 – PJM PJ details defined	Project Manager
T12 – PJM Start	R12 – PJM PJ started	Project Manager
T13 – PJM End	R13 – PJM PJ ended	Project Manager
T14 – Project Monitoring	R14 – Projects for period P monitored	Technical Committee
T15 – PGM Monitoring	R15 – Programs for period X monitored	Executive Commission
T16 – PGM Results Publish	R16 – PGM results published	Executive Commission
T17 – Communication Plan Definition	R17 – Communication Plan Objectives	Executive Commission

Transformation Process Management – TPM
Strategic Objectives – SO
Program Management – PGM
Project Management – PJM

Figure 7 – Transaction Result Table for Public Procurement Governance Model

By performing production acts, subjects contribute to bringing about the goods and/or services that are delivered to the environment of the enterprise. By performing coordination acts, subjects enter into, and comply with, commitments towards each other regarding the performance of production acts [Mendes 2013].

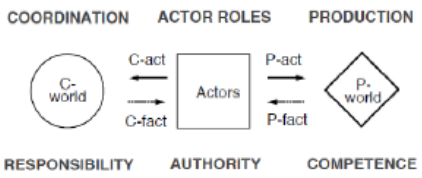


Figure 8 – Graphical representation of the operation axiom and the notions of responsibility, authority, and competence [Dietz 2006].

After identifying the Ontological transaction types, its results and dependencies, we modelled the Actor Transaction Diagram (ATD), using the notation proposed by Dietz and represented in Figure 9.

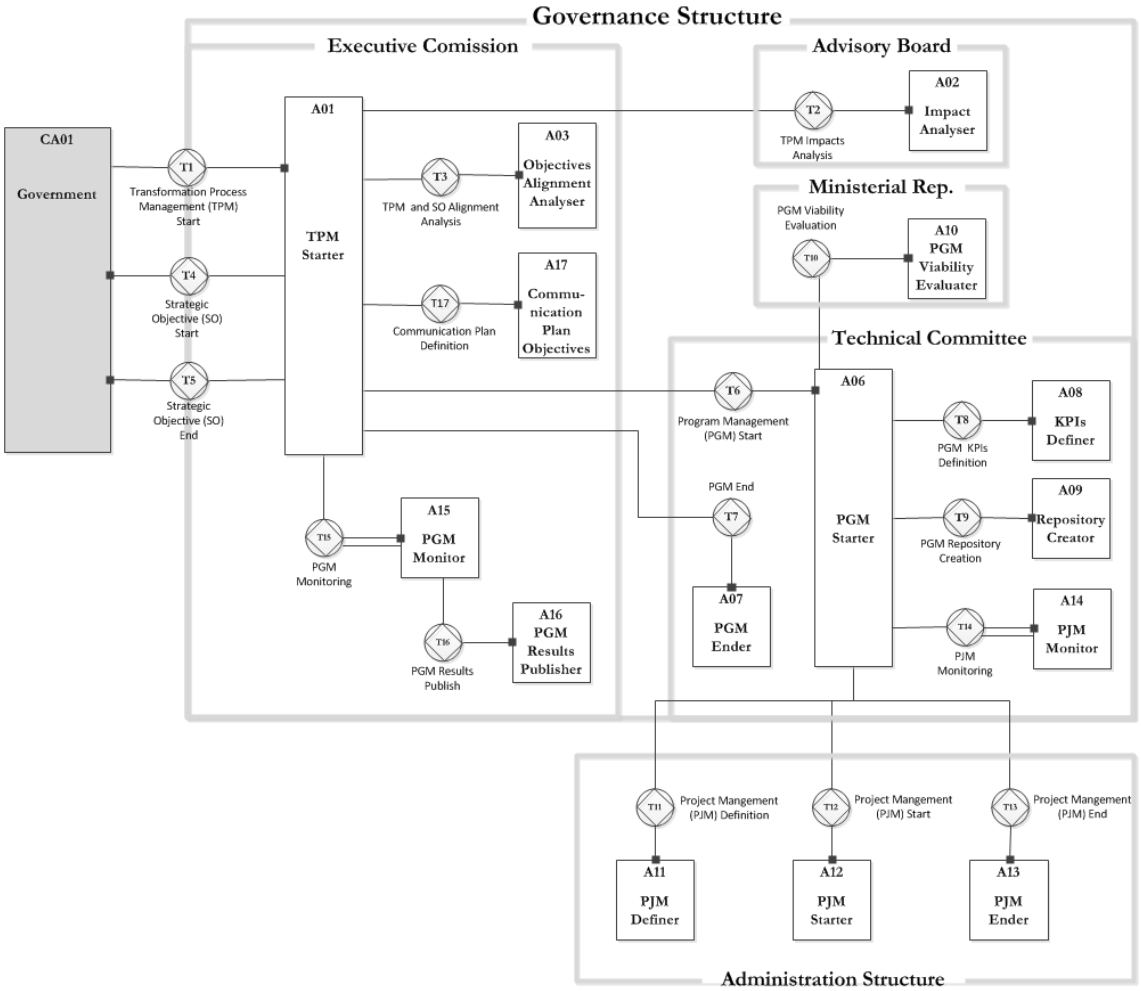


Figure 9 – Actor Transaction Diagram for Public Procurement Governance Model

With this model construction we identified and represented in a schematic way the essentials of our governance proposal for public procurement.

Besides the actors and their relations that are explicit in Figure 5, we can summarize the major model attributes as follows:

- Provides strategic and operational alignment;
- Provides a structure to enhance key entities orchestration;
- Provides transparency;
- Reinforces accountability;
- Provides information integration;

- Provides unity and articulation for program management;
- Provides unity and articulation for project management;
- Provides viability analysis;
- Provides standards on KPIs for public procurement;
- Enforces a communication plan.

All these attributes are related to the objectives defined in the construct definition phase (under the model objective definition step) and aligned with OECD recent recommendations to governments for promoting development and sustainability that were launched in February 2015.

5. Proposal Demonstration

In this section we describe the characterization of the proposed Governance Model for Public Procurement in the context of the current Portuguese Public Administration set-up.

This means, defining a concrete type and composition of the proposed structure for the Portuguese situation and also the setting within the State organic.

The attributions and responsabilities are the ones defined under the proposal section..

The first and most important question in the Portuguese economic conditions concerned the type of structure. For the governance structure there were two basic options: create a new dedicated entity or create a flexible structure.

Both options present several advantages and disadvantages but, for the time being, the choice we consider the most adequate is the flexible structure type because it will not require as much time, resources and capital as the creation of a new entity.

Based on this point, the next question is what should be its positioning in the organic of the Public Administration. This is equally a sensitive point because there are direct implications for the public contracts policy and, of course, for the expenditure management at a national level.

Therefore, we firmly advise that such structure should report to the Prime Minister's Cabinet, which means a framework setting as shown in Figure 10.

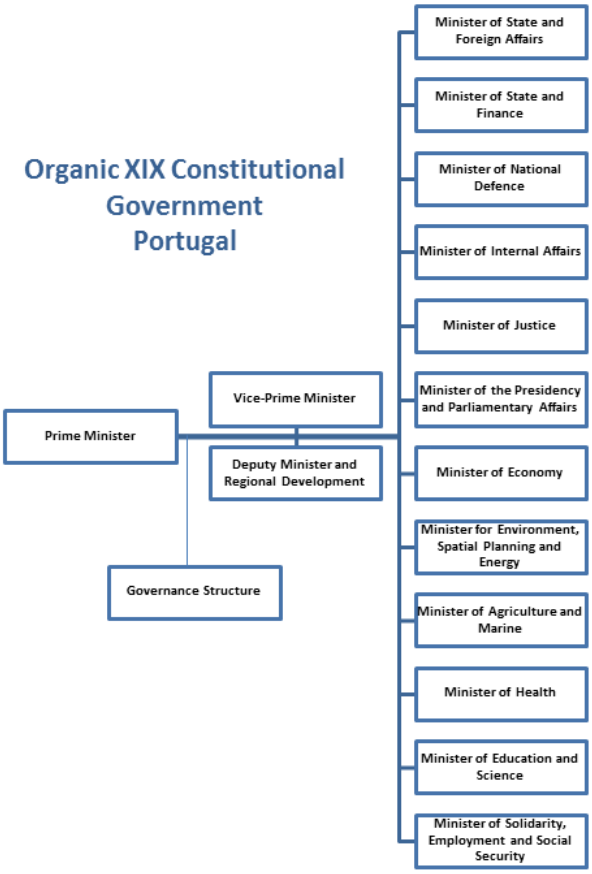


Figure 10 – Governance structure framework

Only this hierarchy can guarantee a strong sponsorship throughout the transformation processes.

The structure was defined under section 4 (proposal) and illustrated in Figure 5. The names of some actors are self-evident, as in the case of Ministerial Representatives, but in other cases, namely the Executive Committee, the Advisory Board and the Technical Committee, a small description becomes necessary because the composition of these actors will be crucial to the achievement of the objectives of the governance structure.

For these three actors in particular, we describe their composition in Figure 11 and can say they include the key stakeholders in Public Procurement as defined in the legal framework and in the government guidelines.

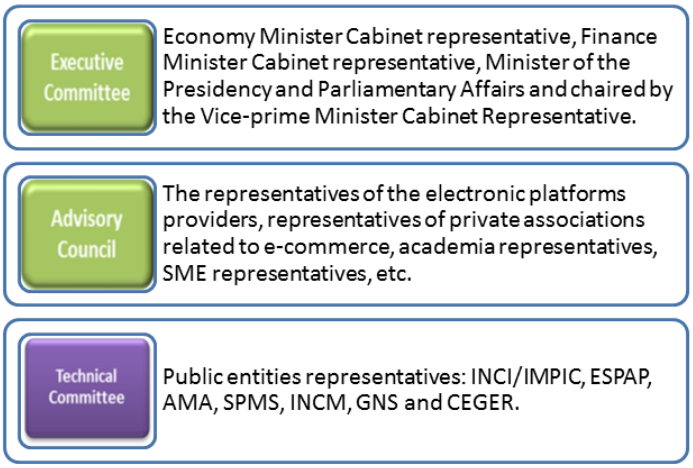


Figure 11 – Actors identification in the Portuguese governance structure

During the design process, an important concern was whether the governance structure should report on policy and functional aspects. In order to address this concern, the composition of the executive committee basically aims to ensure an alignment with the policy set, and the technical committee seeks to ensure the management of a set of measures that will guarantee that alignment.

As referred by Westerman, the only effective way we've seen to drive transformation is top-down, through strong senior executive direction coupled with methods that engage workers in making the change happen [Westerman, Bonnet & McAfee 2014].

The type of structure, the functions and the empowerment that results from the model application to the Portuguese context it will enable the government to ensure a solution for the governance European guidelines and to improve the articulation among all the key stakeholders.

6. Proposal Evaluation

In order to validate that our solution meets the objectives [Osterle, et al. 2011] and get some feedback on our conclusions, we decided to use a "User Opinion study" [Pries-Heje, Baskervile, & Venable 2004]. We planned interviews with public bodies that are key stakeholders in Public Procurement in Portugal and that in our model should be part of the Technical Committee. We also had the privilege of interviewing a senior expert from OECD who is involved in the development of recommendations on Public Procurement. The interviews aimed to present them the results of the research and obtain their overall evaluation and feedback. In total we conducted three interviews.

The questions were classified into two types

- Open questions – answers provided with no format condition
- Agreement questions – answer provided by 1 to 5 (1 – nothing, 2 – little, 3 – enough, 4 – quite and 5 – very much).

All the interviewees are experienced managers, two of them are board members and the other one holds a head of unit post.

An analysis of the results of the interviews shows in general a highly positive feedback to the proposed model, although it was felt that in many cases there was a clear distinction between Governance and Administration functions and very often there was a clear temptation to add more responsibilities to the governance structure that in our proposal belongs to the administration field.

Questions	1	2	3	4	5
A.Do you consider the proposal model adequate?					
B. Do you consider it necessary?					
C. Do you consider the proposal model will contribute for the success of next transformations processes?					
D. Had it in your power would you support the implementation of a governance model?					
E. Had it in your power would you support the implementation of the proposed model?					

Figure 12 - Interview Agreement Questions

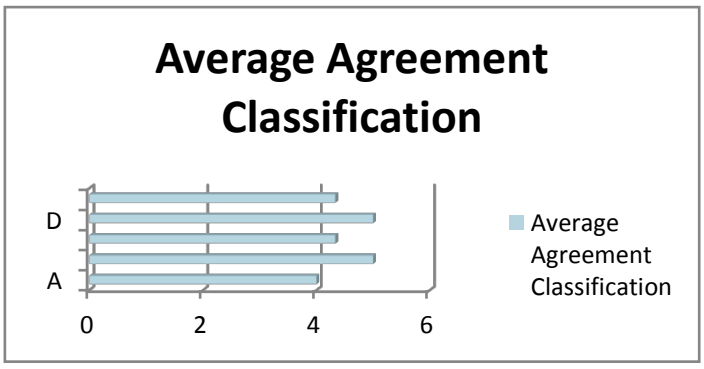


Figure 13 - Average Agreement Answers

Agreement questions (Figure 12) were intended to assess, among the respondents, the correlation with the designed model.

The overall evaluation (Figure 13) is very high, as all answers range between 4 and 5, leading us to conclude that our proposed model is welcome in practice.

Questions
F. Do you consider necessary to have a governance model for procurement in Public Administration? Why?
G. Do you consider that public entities are able to settle an articulation schema for a common goal without a top orchestration? Why?
H. Did you have ever felt the need of an explicit governance model in public services? Why?
I. List the advantages of implementing a governance model for public procurement
J. List the disadvantages of implementing a governance model for public procurement
L. Which kind of governance structure you consider more adequate: fixed one (creating another public entity or set this competence in a existing one)or flexible one (as proposed)? Why?
M. With growing complexity on the services that will have to implement in near future if there is no governance structure which alternatives do you consider will solve the same problems ?
N. If you could propose changes to the model which ones would be? Why

Figure 14 - Interview Open Questions

With the open questions we wanted to find the arguments on which the interviewees based their opinion, namely the advantages and disadvantages of a generic governance model for public procurement.

Some of the answers should be highlighted:

- On the disadvantages of a generic governance model: “a significant risk of simply creating more bureaucracy, more control loops”.
- On the type of structure for the governance model: “variable geometry to be adapted as time goes by, and rather than a strict hierarchical structure must really correspond to a collaborative network of entities”.
- On the advantages of a generic governance model: “Coordination, better implementation, integration, shared vision and objectives, better visibility into the process”.
- On the changes to the model topic: “Consider the possibility of having permanent staff to monitor the implementation and the performance of the model and to guarantee adequate transmission of decisions”.
- On the question about the need for an explicit governance model in public services: “Yes. In several public procurement reforms that I implemented and experienced in different countries there was always the need to develop a governance model. No reform can succeed without a solid governance structure”.

It is possible to say that all the interviewees can list more advantages than disadvantages of implementing governance models and they all underline the importance of this type of solutions for a successful implementation of public policies.

7. Conclusion

Public procurement is a crucial pillar of strategic governance and services delivery for governments. Because of the sheer volume of spending it represents, well governed public procurement can and must play a major role in fostering public sector efficiency and establishing citizens' trust. Well-designed public procurement systems also contribute to achieving pressing policy goals such as environmental protection, innovation, job creation and the development of small and medium enterprises [OECD 2015].

With our contribution, we aim at a paradigm shift in current public procurement leading. We also aim that governance becomes more than an interesting topic in literature and it becomes a concrete reality in Public Administration.

The 2014 European Directives stress out the governance topic in the public procurement legal framework that will be transposed to national laws in all EU Member States, the latest, by April 2016.

In Portugal there has been no integrated vision of public procurement so far and all the achievements depended on the performance of individual public entities.

To address all these questions a governance scheme is necessary and we designed a model based on enterprise ontology, namely based on a methodology (DEMO) with a strong theoretical background, thus using the existing knowledge in an innovative way (one of the design-science research objectives).

Using DEMO for the representation of our proposal surely brings a certain "purist" stance in the sense that we can have pure "objective" and "essential" models of reality. However, one cannot talk about purely ontological and purely objective observations of reality. Humans are always constrained by available resources and the goals they have in mind. So models produced will always be biased by a certain motive and dependent on implementation issues, including the resources available for the modelling process itself [Pombinho, Tribolet, & Aveiro 2014].

This proposal was evaluated by three senior experts from entities highly involved in public procurement. The overall feedback is very positive but we cannot forget the fact that it was not tested in practice and, for the time being, it can only be considered as a theoretical reference for any solution to be implemented in the future.

From our research point of view the most positive aspects of the model are:

- It represents an additional layer to the existing reality;
- It will not increase the costs with the creation of a new public body;
- It addresses transformation as a distinct dimension when compared to the common operational daily tasks (administration), as advocated in the governance definition.
- It clearly distinguishes policy competences from technical competences;
- It clearly defines who should be at each level;
- Although for demonstration purposes it specifically addresses the Portuguese Public Administration case, its principles can be adapted to any Public Administration, particularly in those cases where responsibilities are distributed among several public entities.;
- It demands transparency and accountability as one of the pillars for sustainability development;
- It requires a repository of all types of documentation on the conduction and monitoring of the transformation process;
- It has an inherent obligation to ensure a communication structure as well as gather feedback from all stakeholders.

The less positive points are:

- A flexible structure implies that people in charge accumulate these with other functions, a situation that may create a conflict as regards the time to be allocated to the necessary involvement.
- Flexible structures do not have their own resources, which may reduce capabilities if the existing administration structures cannot incorporate additional features.
- The model, per se, does not ensure that leadership in its essence will depend on the characteristics of the individuals appointed to perform key roles.

The convergence of the designed governance model with the problem definition does not indicate that the proposal is fully complete; it only indicates that the proposal is compatible with the realities of the given case. However, the positive feedback obtained with the interviews gives a qualitative sense of confidence about the usefulness of the solution proposed.

8. References

Anderson, D., & Anderson, L. (2010). *Beyond Change Management*. San Francisco. U.S: Pfeiffer.

Coley, J. L., Doyle, J. L., Logan, G. W., & Stettinius, W. (2005). *What is Corporate Governance?* New York,: McGraw-Hill.

Commission, E. (2014, 02). *http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2014.094.01.0065.01.ENG*. Retrieved from Euro Law: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0024&from=EN

Commission, E. (2014, 07). *Single Market Scoreboard*. Retrieved from internal_market/scoreboard/performance_per_policy_area/public_procurement: http://ec.europa.eu/internal_market/scoreboard/performance_per_policy_area/public_procurement/index_en.htm

De Haes, S., & Grembergen, W. (2008). Analysing the Relationship between IT Governance and Business/IT Alignment Maturity. *Proceedings of the 41st Annual Hawaii International Conference on System Science*,, (p. 428). Waikoloa.

Dietz, J. (2006). *Enterprise Ontology: Theory and Methodology*. Springer.

EC Grow Directorate . (2015, 02 22). *Public Procurement*. Retrieved from Grow Internal Market, Intrustry, Entrepeneurship and SMEs: http://ec.europa.eu/growth/single-market/public-procurement/index_en.htm

Frederickson, H. G. (2007). Whatever Happened to Public Administration?: Governance, Governance Everywhere. In L. E. Ewan Ferlie, *The Oxford Handbook for Public Managment* (p. 282). Oxford University Press.

Henriques, M., Tribolet, J., & Hoogervorst, J. (2010). Enterprise Governance and DEMO: a reference method to guide enterprise design and operation with DEMO. *Conferência da Associação Portuguesa de Sistemas de Informação*. Viana do Castelo, Portugal.

Hevner, A., March, S., Park, J., & Ram, S. (2004, March). Design Science in Information Systems Research. *MIS Quarterly*, pp. 75-105.

Hoogervorst, J. (2012). *A Paradigm Shift Needed: Enterprise Governance and Enterprise Engineering as essential concepts*. www.antwerpmanagementschool.be.

Information Technology Governance Institute (ITGI). (2004). *Board Briefing on IT Governance*. USA: Information Technology Governance Institute.

Institute, E. E. (2015, 06 15). *DEMO Homepage*. Retrieved from DEMO: http://www.demo.nl/

Jick, T. D., & Peiperl, M. A. (2002). *Managing Change: Cases and Concepts*. London: McGraw-Hill.

Mendes, C. (2013). *PhD Thesis: A Method Based on DEMO for Managing Service Quality*. Lisboa, Portugal: Information Systems and Computer Engineering, Instituto Superior Técnico.

OECD. (2015, 05 30). *Recommendation on public procuremen*. Retrieved from OECD website: http://www.oecd.org/gov/ethics/recommendation-on-public-procurement.htm

Osterle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcma, H., . . . Sinz, E. (2011). Memorandum on Design-Oriented Information Systems Research. *European Journal of Information Systems,*, 7-10.

Pereira, R., & Silva, M. (2012). Designing a New Integrated IT Governance and IT Management Framework Based on Both Scientific and Practitioner Viewpoint. *International Journal of Enterprise Information Systems*, 6-8.

Pombinho, J., Tribolet, J., & Aveiro, D. (2014). Linking Value Chains – Combining e3Value and DEMO for specifying Value Networks. *4th Enterprise Engineering Working Conference, EEWC 2014* (pp. 105-109). Funchal, Madeira: Springer.

Pries-Heje, J., Baskervile, R., & Venable, J. (2004). Strategies for Design Science Research Evaluation. *European Conference on Information Systems*. Turku, Filand.

Uhl, A., & Gollenia, L. (2012). *Business Transformation Management Methodology*. Gower Publishing Limited.

Vicente, P., & Silva, M. (2011). A Conceptual Model for Integrated Governance, Risk and Compliance. *Proceedings of the 23rd International Conference on Advanced Information Systems Engineering*,, (pp. 199-213). London.

Webb, P., Pollard, C., & Ridley, G. (2006). Attempting to Define IT Governance: Wisdom or Folly? *39th Annual Hawaii International Conference on System Sciences*, (p. 194a). Hawaii, USA.

Westerman, G., Bonnet, D., & McAfee, A., (2014). *Leading Digital, Turning Technology into Business Transformation*. Harvard Business Review Press.

Sistema de gestão e auditoria fiscal na nuvem CAPSI/2015

Luís Rolo [1], Alexandre Fonte [2], Eurico Lopes [3]

1) Instituto Politécnico de Castelo Branco, Portugal
luisrolo.tec@gmail.com

2) Instituto Politécnico de Castelo Branco, Portugal
adf@ipcb.pt

3) Instituto Politécnico de Castelo Branco, Portugal
eurico@ipcb.pt

Resumo

A evolução da computação na nuvem tem atraído bastantes adeptos, de tal forma que existem cada vez mais aplicações e serviços disponibilizados por este tipo de infraestruturas, a custos razoáveis, consoante a necessidade e quantidade, devido aos seus recursos elásticos sob solicitação. A grande capacidade de armazenamento, processamento e execução de grandes volumes de informação, tem potenciado a que arquiteturas como *Data Warehouses* sejam hospedadas na nuvem, provocando o surgimento de novas aplicações analíticas. A necessidade de envio de informação sobre a atividade fiscal dos contribuintes ou empresas para os portais do governo de forma automática e integrada, bem como, a ausência de um local centralizado para a sua consulta tem-se revelado um problema. Daí que, este artigo apresente uma proposta de arquitetura para um sistema de gestão e auditoria na nuvem de documentação fiscal, possibilitando o envio automático da informação e a sua consulta.

Palavras chave: computação, *Data Warehouse*, fiscal, nuvem, sistema

1. Introdução

A computação na nuvem tem vindo a ganhar posição, na Internet, com a oferta de vários serviços e recursos de computação, para indivíduos e empresas, colmatando uma lacuna existente nos sistemas de informação ao nível do aumento e adição de capacidades e recursos de computação sob solicitação sem a necessidade de investimento numa nova infraestrutura [Eric Knorr, 2008].

A potencial diversidade de oferta de serviços é uma outra característica intrínseca à computação na nuvem, a qual tem permitido o surgimento de novas aplicações no mundo da computação na nuvem, originando novos modelos de negócios. Daí que faça sentido criar e desenvolver novos

projetos na nuvem como sistemas analíticos que possibilitem tirar partido da grande capacidade de recursos elásticos destas infraestruturas a custos reduzidos.

Além do mais, a computação na nuvem é um sistema bastante automatizado, o que permite solucionar problemas de comunicação entre aplicações. Por exemplo, pode permitir estabelecer uma ponte de ligação entre aplicações de *software* de gestão e portais *web* das instituições governamentais competentes, pois atualmente o processo de submissão periódico da documentação fiscal gerada por estes programas para estes portais é realizada de forma manual, e não integrada. Soma-se a impossibilidade de se poder efetuar o rastreamento da informação fiscal, porque após a sua submissão não se tem conhecimento ao certo para que efeito esta servirá no futuro, bem como, por quanto tempo é mantida nestas organizações no final da sua utilização.

Nesse contexto, o objetivo deste artigo prende-se na apresentação de uma proposta arquitetónica que visa criar um mecanismo automático de submissão de documentação fiscal para as entidades correspondentes, bem como, guardar essa informação num repositório central de dados, tornando-a mais útil por forma a auxiliar na tomada de ilações ou decisões mais inteligentes.

O artigo está organizado da seguinte forma: na seção 2 é apresentado o conceito de computação na nuvem, evidenciando as suas principais características, bem como as suas principais vantagens e desvantagens. Na seção 3 é descrito um problema como caso de aplicação desta tecnologia e apresentada uma solução arquitetónica para o mesmo. Na seção 4, é definido e concretizada uma possível modelação para um sistema analítico para *E-Gov*, resultado do problema identificado e enunciado na seção anterior, dando igualmente ênfase às suas mais valias em termos de centralização e consulta da informação. Depois, na seção 5 são descritos os principais resultados esperados na concretização dum projeto desta natureza. Por fim, na seção 6, é realizada uma breve conclusão e a descrição de trabalho futuro.

2. Computação na nuvem

Embora não sendo um conceito novo, e tendo aparecido pela primeira vez nos anos cinquenta, com a introdução da computação de mainframe [Maximiliano Neto, 2014], a computação na nuvem tem ganho particular destaque, no inicio do século XXI com o crescimento acelerado da Internet.

Na generalidade, a computação na nuvem pode ser vista como um tipo de computação que reside essencialmente na partilha de recursos de hardware e software, onde diferentes serviços,

como servidores, armazenamento de informação ou aplicações são disponibilizados através da Internet [Beal Vangie, 2015].

A computação na nuvem é conhecida por ser uma arquitectura abstracta, uma vez que encapsula o seu sistema de implementação. As aplicações são executadas em sistemas fisicos não especificados, sendo a informação armazenada em locais desconhecidos, e a sua administração terceirizada por administradores de sistemas. O acesso à nuvem é realizado pelos utilizadores de forma ubiqua [Sosinsky, 2011]. Além disso, os sistemas na nuvem são virtualizados através de agrupamento e partilha de recursos, sendo fornecidos e escalados consoante a necessidade [Sosinsky, 2011].

Com diferentes modelos de implementação que indicam a natureza para esta foi construída, a computação na nuvem pode ser pública, privada, hibrida ou de comunidade.

A nuvem pública é propriedade e é operada por várias empresas que oferecem acesso rápido e recursos de computação a outras organizações ou individuos [IBM, 2015], automaticamente e em demanda, tais como, aplicações, armazenamento e outros serviços. É baseada no modelo pagamento consoante o uso, isto é, o pagamento dos serviços disponibilizados por este modelo são cobrados consoante a sua utilização e o seu custo aos seus utilizadores [Parsi et al, 2013].

A nuvem privada é propriedade e é operada por uma única empresa que controla os recursos virtualizados e outros serviços automáticos que são personalizados e utilizados por vários grupos empresariais [IBM, 2015]. Pode ser interna *(on-premise),* alojada no próprio centro de dados das organizações que requerem este género de serviços, garantindo uma maior protecção e segurança ou pode ser externa *(off-premise)* alojada num fornecedor do serviço, assegurando uma maior privacidade em relacão às nuvens públicas devido aos riscos associados de partilha de recursos fisicos [Parsi et al, 2013].

A nuvem comunidade é uma nuvem com o objetivo de atender um determinado propósito ou função [Sosinsky, 2011]. É partilhada entre organizações que detêm interesses e requisitos semelhantes, assim como os seus custos. Pode ser alojada e gerida interna, externamente, ou mesmo por fornecedores *third-party* [Parsi et al, 2013], e tem a vantagem de tirar partido de politicas de segurança comuns e de escalabilidade superior em relação a nuvens privadas [Ken, 2012].

Por último, a nuvem hibrida pode combinar múltiplas nuvens, privada, comunidade, ou pública mantendo as suas identidades originais. No entanto, estão unidas como uma unidade. [Sosinsky, 2011], originando um único serviço, podendo oferecer a flexibilidade, controlo e segurança combinada dos vários modelos de implementação.

Os fornecedores deste tipo de computação podem oferecer três principais modelos de serviço, que podem ser organizados sob a forma de uma arquitetura de camadas. O primeiro modelo encontra-se na camada superior, designado *Software as a Service (SaaS),* onde os utilizadores fazem uso de navegadores web para aceder a aplicações desenvolvidas por outros, mantendo a oferta como um serviço na Internet [Armbrust et al., 2009].

O modelo intermédio deste serviço é denominado de *Platform as a Service (PaaS)*. Trata-se da camada onde as aplicações são desenvolvidas usando um conjunto de linguagens de programação e ferramentas que são suportadas e fornecidas pelo fornecedor deste tipo de serviço. Os programadores obtêm um elevado nível de abstração que permite que o seu foco seja exclusivamente no desenvolvimento de aplicações, sem a necessidade de definir ou manter a infraestrutura por elas suportada [Boniface, et al., 2009].

Por último, a camada mais baixa onde os utilizadores usam os recursos de computação como bases de dados, poder de processamento, memória, e armazenamento, ou recursos para implementar e executar as suas aplicações é designado por *Infrastructure as a Service (IaaS)*. Ao contrário do modelo PaaS, o IaaS permite aos utilizadores dar acesso à infraestrutura subjacente através do uso de máquinas virtuais que automaticamente são escaladas para cima ou para baixo. O IaaS é mais flexível do que o PaaS, pois permite que o utilizador possa implementar qualquer tipo de software no sistema operativo [Abraham et al, 2009].

Em termos de vantagens na adoção em computação na nuvem por parte das empresas ou a nível individual, os motivos tornam-se inúmeros e bastante atrativos, pois a sua confiabilidade em planos complexos para recuperação de desastres, deixam de ser preocupação para as empresas que usam este tipo serviço [Salesforce, 2015].

Os seus fornecedores disponibilizam constantemente atualizações automáticas de software e os recursos podem rapidamente ser fornecidos, a qualquer altura e em qualquer quantidade [Sosinsky, 2011], de forma agrupada e num sistema *multi-tenant*, sendo dinamicamente alocados ou realocados consoante necessidade. O sistema de medições permite que estes sejam medidos, auditados, e reportados para o cliente [Sosinsky, 2011].

Outra grande vantagem é o custo reduzido das redes na nuvem, pois estas operam com maior eficiência, o que diminui bastante os custos, para além de fornecerem escalabilidade, balanceamento de carga e *failover*, tornado-a altamente confiável [Sosinsky, 2011], diminuindo drasticamente a necessidade de investimento em infraestruturas e recursos como armazenamento e processamento.

Por outro lado, existem também algumas desvantagens da computação na nuvem. Por exemplo, e um pouco contraditório, são os custos elevados de computação associados à sua utilização a médio/longo prazo, sobretudo se estes não forem aprovisionados eficientemente. Especialistas dizem para ter sempre o controlo dos custos em mente, porque os erros mais comuns que costumam ocorrer são pedidos de poder de computação excessivos e desnecessários, ou a não desativação de recursos quando estes não estão a ser usados, ou a não utilização de ferramentas de monitorização dos ciclos de computação desperdiçados, ou ainda fazer pensar aos programadores que esses ciclos são gratuitos [Boulton, 2015].

Outro problema é o facto de as aplicações disponibilizadas na nuvem sofrerem com a latência inerente das conetividades *WAN*, e se houver necessidade de transferir grande quantidade de dados, a computação na nuvem pode não ser a melhor opção [Sosinsky, 2011].

Além disso, a computação na nuvem é um sistema sem estado, como a Internet, e solicitações do género HTTP: PUTs, GETs, podem perder-se se houver uma desconexão arquitetónica entre as mensagens de pedido-resposta [Sosinsky, 2011], porque a computação na nuvem é baseada no paradigma cliente-servidor, tirando partido das comunicações pedido-resposta entre clientes e servidores sem estado. Isto é, não necessita de um cliente para estabelecer a ligação, ao invés disso, verifica o seu pedido como sendo uma transacção independente, respondendo a este. Assim, o cliente não é afetado se o servidor for abaixo e voltar ao seu estado operacional, nem tem que se preocupar com o estado do sistema, o que faz com que este seja simples, mais robusto e escalável [Marinescu, 2013].

Contudo, a maior preocupação é sem dúvida a privacidade e a segurança da informação, pois os dados navegam e são armazenados em sistemas fora do controlo do utilizador, o que aumenta o risco de interceção e prevaricação por parte de terceiros [Sosinsky, 2011]. Isto embora, seja usado como contra-argumento, o fato da gestão dos serviços e da infraestrutura na nuvem seja realizada por equipas especializadas, bem treinadas, e sejam adotados rígidos protocolos de segurança nos centros de dados.

Goste – se ou não, a verdade é que a computação na nuvem tem vindo a crescer nos últimos tempos e a ganhar novos adeptos, quer para criar novos modelos de negócio, quer para gerar novo tipo de aplicações e com isso atingir rapidamente o *time-to-market*, reduzindo drasticamente o custo de desenvolvimento dos projetos e os custos inerentes com as infraestruturas que os suportam, o que indica claramente que este é e será o futuro das *TIC* nos próximos anos.

Um dos serviços que poderá tirar partido dos recursos da computação na nuvem é sem dúvida o *Data Warehouse*, um repositório central de dados, utilizado para armazenar informações

relativas às atividades de uma determinada organização de forma histórica e consolidada. É visto como um sistema de dados orientado por assuntos, integrados, não voláteis e variantes no tempo [Bill Inmon, 2009]. Tem por objetivo integrar dados de diferentes fontes e formatos, não sendo construído para suportar o processo funcional ou operacional, de uma empresa, mas sim fornecer um meio para facilitar o uso da informação [Joachim *et al*, 2011; Ralph Kimball et al, 2002].

Na computação na nuvem, o processo de *Data Warehousing* poderá usar as potencialidades e recursos do *Big Data* para o processo de extração, transformação e carga, *(ETL do Inglês Extract, Transform, Load)* pois trata-se de uma tecnologia veloz com a capacidade de suportar e manter grandes quantidades de informação, onde é possível armazenar dados em dispositivos de baixo custo de forma estruturada e não estruturada [Bill Inmon, 2014].

3. Descrição do Problema

Atualmente existem diversas aplicações de faturação e contabilidade, como Primavera, Filosoft, F3M, etc. que geram determinada informação em formato digital, de forma padronizada, em ficheiros *XML* para diversas entidades governamentais, como a segurança social, ministério das finanças e instituto nacional de estatística, de forma automatizada. Um exemplo de formato padrão para o ficheiro *XML*, e obrigatório para uso por este tipo de aplicações, conforme estabelecido na Portaria n.º 321-A/2007, é o SAFT-PT (*Standard Audit File for Tax Purposes-Versão Portuguesa*), o qual permite a recolha dos dados fiscais relevantes referentes aos períodos das declarações fiscais e/ou para análise, em qualquer altura, pela Autoridade Tributária e Aduaneira (AT) [Portaria 321-A/2007, 2017]. A estrutura atual do ficheiro *SAFT-PT* está definida em [SAFT-PT, 2013].

É importante referir ainda que nem todos os formatos exigidos por estas entidades são em formato *XML*, e todos estes obedecem a um conjunto de caracteristicas muito especificas consoante o seu fim. Por exemplo, para o envio de eletrónico de dados referentes à declaração mensal de remunerações, quer para a Segurança Social, quer para a Autoridade Tributária, o formato exigido é em código ASCII. Daí que as finanças disponibilizem no seu portal informação e modelos de ficheiros com uma estrutura especifica para entrega de obrigações declarativas em formato eletrónico via Internet [Finanças, 2015b] de dados referentes a contribuintes e empresas, para a entidades governamentais como as Finanças, Segurança Social e Instituto Nacional de Estatística.

Contudo, e apesar da geração destes ficheiros ser efetuado de forma automática, o processo de submissão desta informação é realizado de forma manual e não integrada para cada um dos

portais *web* destes organismos institucionais, o que obriga a que a pessoa responsável ou com privilégios para tal tenha que realizar um processo algo moroso e bastante inflexível.

Além disso, o controlo da informação gerada é perdido ao longo do tempo, pois apenas estes organismos a mantêm nas suas bases de dados, impossibilitando a consulta e navegação sobre dados como facturação ou impostos que são contabilizados e declarados ao longo dos anos fiscais por parte dos contribuintes, ou empresas o que a acontecer permitiria ter uma visão sobre a informação fiscal destas entidades ou indivíduos ao longo dos anos.

Nesse sentido, e perante este cenário uma proposta de solução para este problema é a criação de um sistema de gestão e auditoria fiscal, na nuvem, e responsivo com auxilio a um *Data Warehouse* que permitirá persistir informação proveniente de variadas fontes de informação estruturadas e não estruturadas, podendo realizar e fornecer relatórios dinâmicos e personalizados, para além de poder enviar os ficheiros ou dados requeridos por estas entidades de forma automatizada por remessa direta nos portais das autoridades ou de preferência através dos seus *web services*.

A grande vantagem de um sistema desta natureza será a possibilidade de aceder e navegar sobre dados em qualquer ponto do globo, em qualquer dispositivo móvel, e em qualquer altura, pois esta estará alojada na nuvem, para além de possibilitar o envio da informação para os diferentes organismos governamentais de forma automática [Joachim *et al*, 2011; Diarmuid & Donal, 2015].

Os autores identificam situações adicionais onde a arquitetura proposta permitirá um sistema de armazenamento e controle cibernético (DW), o qual permitirá complementar a extração de conhecimento e alinhar a empresa com as estratégias que esta vier a definir em termos de gestão [Gates & Germain, 2015].

4. Arquitetura de Sistema Analítico para E-Gov

Mediante a solução apresentada na seção anterior, no sentido de tornar o processo de submissão de ficheiros fiscais de forma automática, bem como, a possibilidade de manter a informação num repositório central como um DW para uma posterior consulta, um possível modelo de arquitetura para o sistema de gestão e auditoria fiscal *(E-Gov)*, pode ser observado abaixo na Figura 1.

Figura 1 - Proposta de arquitetura de sistema de gestão e auditoria fiscal

De acordo com a Figura 1, numa primeira fase a informação é lida das diversas fontes de informação, quer esta seja estruturada ou não estruturada. De seguida, os dados são colocados num ambiente de estágio, isto é, o processo *ETL*, passando por um longo processo de transformação e limpeza, no sentido de poder carregar a informação para os modelos multidimensionais definidos, designados *datamarts*, que em conjunto formarão o repositório central de dados, ou seja, o *Data Warehouse.*

Após o carregamento do *DW*, a informação pode ser passada para o cubo *OLAP (do Inglês On-Line Analytical Processing)*, ilustrado na Figura 2.

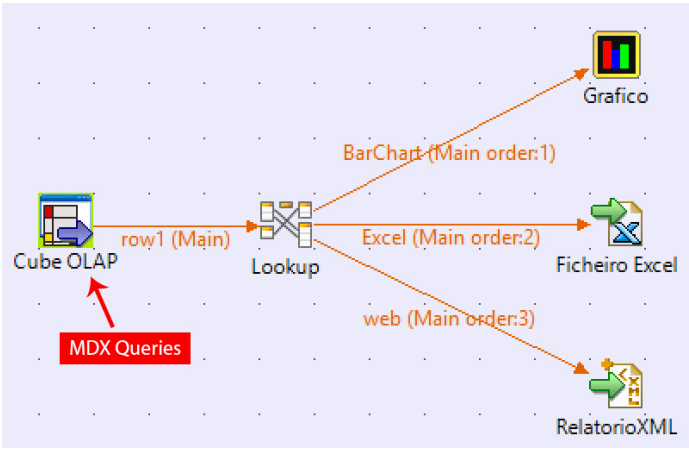


Figura 2 - Processo de criação e visualização do cubo OLAP

O objetivo do cubo *OLAP* é disponibilizar a informação de forma tridimensional com o recurso a *MDX queries*, permitindo navegar sobre a informação e disponibilizá-la no final em aplicações analíticas em relatórios *ad-hoc* ou gráficos, dando origem ao processo convencional de *Data Warehousing*.

No entanto, a grande diferença desta arquitetura em relação ao tradicional processo de *Data Warehousing* reside no fato da informação poder ser novamente alvo de transformação, usando ferramentas de extração e transformação de dados, (e.g., exemplo de software de integração em [Talend, 2015]) pois assim que esta exista no DW, poderão ser agendados novos processos *ETL*, podendo a informação transformada no final ser disponibilizada nas mais diversas formas, como em aplicações web analíticas na nuvem, em forma de relatórios ou gráficos, indicado na Figura 3, suportadas em infraestruturas de fornecedores deste género de serviços como a *Google* [Google, 2015] ou *Open Shift* [Open Shift, 2015], usando as suas *APIs*.

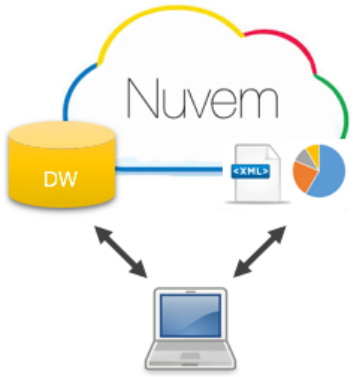


Figura 3 - Serviços integrados na nuvem

Na prática, e conforme ilustra a Figura 3, a informação seria inicialmente extraída das suas fontes, e transformada localmente com uma ferramenta de integração de dados, sendo de seguida persistida no DW na nuvem, no sentido de tirar partido do seu poder de computação e capacidade de armazenamento. Depois, esses dados poderiam ser usados diretamente para alimentar as aplicações web desenvolvidas para a nuvem que necessitassem de usar o DW como fonte de informação, para efeitos de análise e navegação sobre dados, dispondo assim de um repositório de informação centralizado. Seguidamente, e usando a mesma ferramenta, poderia ser agendada e efetuada uma nova transformação à informação anterior, mas desta vez com o DW como fonte de informação, dando origem aos ficheiros padronizados requeridos pelas instituições fiscais com o formato por elas requerido, o que tornaria esta arquitetura bastante flexível e versátil.

Um componente intrínseco ao sistema e com um papel importante na automatização do processo de submissão dos ficheiros ou dados padronizados, é, portanto, o módulo de comunicação com os sistemas de informação destas entidades. Este módulo adotará, consoante o sistema alvo, as formas de submissão disponíveis ou por remessa direta nos portais ou através de *web services*, observando os aspetos legais e técnicos requeridos.

A título de exemplo, a Autoridade Tributária e Aduaneira (AT), disponibiliza vários *web services SOAP* para comunicação da informação fiscal, e.g., declaração *IRC (Imposto Rendimento Coletável)*, faturas ou documentos de transporte [Finanças, 2015a]. Neste caso, é requerida a construção dos pedidos SOAP *(Simple Object Access Protocol)* de acordo com a especificação *WSDL* do serviço disponibilizado (e.g., comunicação de faturas em [Finanças, 2015c]), e a adoção de mecanismos de segurança. Mais especificamente, o cabeçalho dos pedidos – *SOAP:HEADER* - têm que incluir campos de autenticação do utilizador responsável com cifragem das senhas utilizando chave pública da AT, e a criação de um canal de transporte seguro *HTTPs* utilizando certificado *SSL* previamente submetido e assinado pela AT.

No final, e integrados todos esses serviços, poderia ser desenvolvida uma aplicação web que disponibilizaria relatórios analíticos para os seus utilizadores finais, podendo a informação ser acedida através de computadores convencionais ou dispositivos móveis e expondo serviços *web*, no sentido de permitir que a informação requerida pelas instituições governamentais possa ser consumida por estes o que traria benefícios em termos processuais, pois simplificaria e centralizaria os serviços e informação num único lugar.

5. Resultados esperados

Com o desenvolvimento deste sistema espera-se conseguir facilitar o processo de armazenamento e obtenção da informação fiscal relevante, como por exemplo, faturação diária, declarações mensais de remuneração, relatórios únicos, *IVA* ou *IRS* de contribuintes e empresas, no sentido de possibilitar a navegabilidade sobre a informação, isto é, tornar os dados gerados em informação mais útil, e com isso poder tomar ilações ou mesmo tomadas de decisão mais inteligentes.

Por outro lado, espera-se igualmente que este sistema possa auxiliar e solucionar o problema processual de submissão da documentação fiscal para as diversas entidades governamentais, como o ministério das finanças e a segurança social, através automatização do processo, utilizando as formas disponíveis de submissão por remessa nos portais destas entidades ou atravéso de *web services* disponibilizados.

6. Conclusão

A diversa oferta de serviços, recursos elásticos e modelos de implementação de computação da nuvem tem vindo a crescer e a ser uma forte aposta por parte dos fornecedores deste tipo de computação.

As inúmeras vantagens como o custo reduzido, alta confiabilidade e *failover* da computação da nuvem permitem que surjam cada vez mais novos modelos de negócio e aplicações baseadas neste tipo de paradigma.

Contudo, é preciso ter em conta que o risco de segurança é ainda um fator que impede que a adoção da computação na nuvem seja maior e mais acelerada.

Um dos exemplos de grande utilidade da computação na nuvem, seria o desenvolvimento de um sistema de gestão fiscal e auditoria, com dados das empresas ou dos contribuintes, o que permitiria resolver problemas de navegabilidade sobre a informação, bem como, automatizaria o processo de submissão dos dados para as diferentes entidades governamentais de forma fácil, simples e rápida.

A implementação deste modelo servirá como trabalho futuro, no sentido de validar a arquitetura proposta e apresentar resultados concretos e satisfatórios.

7. Referências

Abraham, M. et al., *Autonomic Clouds on the Grid. Journal of Grid Computing*, pp. 1-18, 2009.

Armbrust, M. et al., *Above the clouds: A berkeley view of cloud computing*, EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-28., 2009.

Boulton, Clint, Wall Street Journal – The Hidden Waste and Expense of Cloud Computing, http://www.wsj.com/articles/the-hidden-waste-and-expense-of-cloud-computing-1424139032, (07 de Agosto de 2015), 2015.

Boniface, M. et al., *Platform-as-a-Service Architecture for Real-time - Quality of Service Management in Clouds*, 2009.

Diarmuid Ó Coileáin & Donal O'mahony. 2015. Accounting and Accountability in Content Distribution Architectures: A Survey. ACM Computing Survey 47, 4, Article 59 (May 2015), 35 pages. DOI=10.1145/2723701.

Finanças, Portal das, Apoio ao Contribuinte, http://info.portaldasfinancas.gov.pt/pt/apoio_contribuinte/ (15 de junho de 2015), 2015.

Finanças, Portal das, Especificação de *Webservice* para comunicação dos dados das faturas (WSDL), http://info.portaldasfinancas.gov.pt/NR/rdonlyres/02357996-29FC-4F11-9F1D-6EA2B9210D60/0/factemiws.wsdl (15 de junho de 2015), 2015.

Finanças, Portal das, Ajuda – Suporte Informático – Formato de ficheiros, https://www.portaldasfinancas.gov.pt/de/ajuda/DGCI/FAQSI.htm (07 de Agosto de 2015), 2015.

Gates, Stephen & Germain, Christophe (2015) *Designing Complementary Budgeting and Hybrid Measurement Systems that Align with Strategy, Management Accounting Quarterly*. Winter2015, Vol. 16 Issue 2, p1-8. 8p.

Google, *Google Cloud Platform*, https://cloud.google.com/, (07 de Agosto de 2015), 2015.

IBM, *IBM Cloud – What is cloud? Computing as a service over the Internet.* http://www.ibm.com/cloud-computing/us/en/what-is-cloud-computing.html (6 de Janeiro de 2015), 2015.

Inmon, B, *Beye Network. Big Data Implementation vs Data Warehousing*, http://www.b-eye-network.com/view/17017 (23 de Novembro de 2014), 2014.

Inmon, B., *My DBA World – Blog for Database Concepts,* http://mydbaworld.wordpress.com/2009/07/23/bill-inmon-vs-ralph-kimball/, (21 de Novembro de 2014), 2009.

Joachim Götze, Tino Fleuren, Bernd Reuther, and Paul Müller. 2011. *Extensible and scalable usage accounting in service-oriented infrastructures based on a generic usage record format*. *In Proceedings of the 6th International Workshop on Enhanced Web Service Technologies* (WEWST '11), Walter Binder and Heiko Schuldt (Eds.). ACM, New York, NY, USA, 16-24. DOI=10.1145/2031325.2031328.

Portaria 321-A/2007, Portaria 321-A/2007 de 26 de Março, Diário da República, 1.ª série — N.º 60 — 26 de Março de 2007.

Ken, Cloud 101: The Four Deployment Models, https://kensvirtualreality.wordpress.com/2012/10/15/cloud-101-the-four-deployment-models/, (07 de Agosto de 2015), 2015.

Kimball, R. et al, *The Data Warehouse Toolkit. Second Edition*. John Wiley & Sons, Inc, New York, EUA, 2002.

Knorr, E., *InfoWorld: What cloud computing really means,* http://www.infoworld.com/article/2683784/cloud-computing/what-cloud-computing-really-means.html, (23 de Janeiro de 2015), 2008.

Marinescu, Dan C., *Cloud Computing – Theory and Practice*, Morgan Kaufmann, Waltham, USA, 2013, 978-0-12404-627-6.

Neto, M., *Thoughts On Cloud – A brief history of cloud computing*, http://thoughtsoncloud.com/2014/03/a-brief-history-of-cloud-computing/, (15 de Dezembro de 2014), 2014.

Open Shift, Develop, *Host, and Scale your Apps in the cloud. The hybrid cloud application platform by Red Hat*, https://www.openshift.com/, (07 de Agosto de 2015), 2015.

Parsi, Kalpana et al., *International Journal of Advanced Research in Computer Science and Software Engineering – A Comparative Study of Different Deployment Models in a Cloud*, http://www.ijarcsse.com/docs/papers/Volume_3/5_May2013/V3I5-0229.pdf, (7 de Agosto de 2015), 2013.

SAFT-PT. Estrutura de dados, versão 1.03_01, associada à Portaria n.º 274/2013, de 21 de agosto de 2013, http://info.portaldasfinancas.gov.pt/apps/saft-pt03/SAFTPT1.03_01.xsd, (15 junho de 2015), 2015.

Salesforce, *Salesforce.com. Why Move to the Cloud? 10 Benefits of Cloud Computing*, http://www.salesforce.com/uk/socialsuccess/cloud-computing/why-move-to-cloud-10-benefits-cloud-computing.jsp, (6 de Janeiro de 2015), 2015.

Sosinsky, B., *Cloud Computing Bible*, Wiley Publishing, Inc, Indianapolis, Indiana, 2011, 978-0-470-90356-8.

Talend, Open Studio, Free, open source integration software, https://www.talend.com/products/talend-open-studio, (07 de Agosto de 2015), 2015.

Vangie, B. *Webopedia - Cloud Computing (the cloud)*, http://www.webopedia.com/TERM/C/cloud_computing.html, (02 de Janeiro de 2015), 2015.

Sistema Inteligente de Recolha, Armazenamento e Visualização de Informação proveniente do Twitter

Gaspar Bogueira[1], Fernando Batista[1] e Joao Paulo Carvalho[2]

1) ISCTE-IUL – Instituto Universitário de Lisboa & INESC-ID Lisboa, Portugal
gmrba@iscte.pt

1) ISCTE-IUL – Instituto Universitário de Lisboa & INESC-ID Lisboa, Portugal
fernando.batista@inesc-id.pt

2) Instituto Superior Técnico - Universidade de Lisboa & INESC-ID Lisboa, Portugal
joao.carvalho@inesc-id.pt

Resumo

As redes sociais têm ganho bastante popularidade para a partilha de informação sobre os mais diversos tópicos desde a política, desporto ou mesmo aspetos do quotidiano. As mensagens partilhadas no Twitter[1] (tweets) são essencialmente públicas constituindo uma fonte de informação que por ser difundida em tempo real pode revelar-se útil em domínios como o turismo, marketing, saúde ou segurança. Este artigo descreve o desenvolvimento de um Sistema de Informação envolvendo a criação de um repositório de tweets (*corpus*) escritos em Português Europeu e publicados em Portugal. O sistema envolve também uma REST API que permite o acesso à informação armazenada e um Dashboard Web para análise e visualização de indicadores sobre os dados.

Palavras-chave: Sistema de Informação, Twitter, MongoDB, Geolocalização, Redes Sociais

1. Introdução

Uma grande parte da sociedade atual vive hiperconectada. Apesar da distância, a todo o momento estamos próximos de amigos ou de outras pessoas cujas vidas queremos acompanhar através de fotografias colocadas no Instagram, conteúdos que são partilhados no Twitter ou no Facebook, vídeos que se colocam no Youtube, etc.. A facilidade com que se pode acompanhar a vida de alguém, seja pela sua localização geográfica difundida no Foursquare, ou pelas fotografias das férias colocadas no Flickr ou ainda pelas comunidades em que participa no Orkut, faz com que os gostos interesses e demais características sejam expostas ao mundo através das redes sociais. A rede social tal como é definida por Boyl *et. al.*[2007] é um serviço baseado na *web* que permite construir um perfil público ou semipúblico num sistema delimitado e gerir uma lista de outros utilizadores com os quais se mantem uma conexão e se acede quer à sua lista de conexões quer à lista de conexões de outros utilizadores, visualizando as suas

[1] https://twitter.com/

atualizações de estado. As redes sociais tendem a transformar-se não só num ambiente para seguirmos a vida de amigos, mas também numa plataforma de interação com empresas, marcas, aplicações ou serviços.

A análise dos conteúdos e informações partilhadas nas redes sociais tem-se revelado bastante útil em diversas áreas, incluindo a política, o turismo, a saúde pública ou a segurança. No caso particular do Twitter, a rede social em que incidirá o estudo apresentado neste trabalho, são publicados em média cerca de 500 milhões de novos tweets por dia [Twitter 2015]. O Twitter disponibiliza o livre acesso a alguma da informação produzida pelos seus utilizadores através de APIs (*Application Programming Interface*) públicas. A popularidade do Twitter como fonte de informação tem conduzido ao desenvolvimento de inúmeras aplicações e investigações em diversos domínios. Com cerca de um milhão e meio de tweets relacionados com a saúde, onde constavam menções a diversas doenças incluindo alergias, obesidade e insónias, Paul *et al.* [2011] desenvolveu um método para rastreamento de doenças medindo fatores de risco ao nível comportamental, localizando doenças por regiões geográficas ao analisar os sintomas e medicação aplicada. Igualmente, com informação obtida do Twitter, Santos *et al.* [2013, 2014] utilizou um conjunto de aproximadamente 2700 tweets produzidos em Portugal para prever a taxa de incidência e disseminação do vírus *influenza* na população Portuguesa. Widener *et al.* [2014] utilizou uma *framework* de Data Mining para, através da aplicação de métodos de extração de informação e análise de sentimentos, tentar compreender como tweets geolocalizados podem ser utilizados na pesquisa da prevalência de alimentos saudáveis e não saudáveis em regiões contíguas dos Estados Unidos. Outros estudos relacionados com a saúde pública foram também desenvolvidos por Culotta [2010] e Scanfeld *et al.* [2010].

O Twitter foi também usado como fonte de informação para ajudar a identificar ou localizar a ocorrência de sismos [Sakaki *et al.* 2010] dado que "ao ocorrer um sismo as pessoas produzem muitos *posts* no Twitter relacionados com o acontecimento o que permite a identificação de sismos simplesmente pela observação do aumento do volume de tweets". No estudo desenvolvido por Kumar *et al.* [2013b] é proposta uma abordagem para identificar um subconjunto de utilizadores e sua localização que justifique serem seguidos em situações de catástrofe, de modo a obter um acesso rápido a informação útil sobre o acontecimento. Durante uma situação de crise a localização de determinado utilizador é um fator importante para determinar se é provável que publique informações relevantes sobre o estado da crise. Por exemplo, no caso de um sismo, os tweets produzidos num local próximo ao sismo são provavelmente mais pertinentes para avaliação do estado da situação do que tweets produzidos num local mais distante. Outros estudos foram produzidos tendo por base tópicos semelhantes [Mendoza *et al.* 2010, Qu *et al.* 2011, Lachlan *et al.* 2014].

Gerber [2014] refere que o Twitter é uma fonte de dados ideal para problemas de suporte à decisão. Utilizando tweets marcados no espaço e no tempo tentou prever a atividade criminal na maior cidade dos Estados Unidos. Dado que os tweets são informação pública disponibilizados oficialmente por serviços do Twitter, o desenvolvimento de modelos de análise linguística que permitam a identificação automática de tópicos relacionados com a prática de crime pode ter bastante relevância não só na prevenção do mesmo, como na tomada de decisão em fase de julgamento em tribunal.

O desenho de uma arquitetura de software para a captura e extração de informação contida em tweets configura-se como um grande desafio considerando as limitações no acesso à informação impostas pela Twitter API [Oussalah *et al.* 2013]. Perera *et al.* [2010] descreve uma arquitetura de software baseada na Twitter API que, recorrendo a Python e MySQL, recolhe os tweets enviados para utilizadores específicos. Na obtenção de dados espaciais (localização, nome, descrição...) dos autores dos tweets foi utilizado a Twython API. O processo de recolha é executado em intervalos de 5 minutos sendo recolhidos os tweets enviados para determinado *id* de utilizador do Twitter, nomeadamente o Presidente Barack Obama. Anderson *et al.* [2011] apresenta uma arquitetura com elevado grau de concorrência de modo a permitir a recolha de um grande volume de dados com um máximo teórico de 500 milhões de tweets por dia. O código desenvolvido é *multithreaded* e está adaptado para executar em máquinas com diversos processadores. Foram utilizadas as *frameworks* Spring, MVC, Hibernate e JPA e os componentes de infraestrutura Tomcat, MySQL e Lucene. Por outro lado, Marcus *et al.* [2011] desenvolveu o TwitInfo, uma plataforma para recolha e processamento de tweets em tempo real para efeitos de Análise de Sentimento. A arquitetura proposta por Oussalah *et al.* [2013] tem como objetivo a análise semântica e espacial de dados recolhidos do Twitter. Na recolha dos *tweets* foi utilizada a Streaming API por permitir o acesso aos tweets em tempo real e de forma contínua. Os tweets recolhidos são restringidos a uma região retangular delimitada por coordenadas geográficas (longitude e latitude). A implementação do software é baseada em Django (*framework* para desenvolvimento de aplicações web em Python), Apache Lucene (para indexação dos tweets) e MySQL para armazenamento dos dados recolhidos. Esta arquitetura permite a pesquisa de tweets por texto, nome de utilizadores ou localização.

Independentemente do grau de conhecimento e utilização das redes sociais é inegável a sua importância na sociedade contemporânea. Publicitar um evento, comentar ou divulgar uma ideia são práticas comuns nas redes sociais, tornando-as num meio propício à expressão da opinião individual e consequentemente um dos principais formadores da perceção da sociedade e do mundo que nos rodeia. Hoje em dia os eventos importantes são muitas vezes comentados nas redes sociais, mesmo antes de se tornarem "notícia pública" e até mesmo as agências noticiosas

tiveram de se adaptar e começar a usar as redes sociais como fonte de informação. Apesar da sua importância, muitas questões sobre o efeito das redes sociais na sociedade estão ainda por analisar devidamente. Para que possam ser analisadas questões como quais os temas, tópicos ou eventos importantes partilhados nas redes sociais, quais os atores principais dentro dos temas, qual a origem e propagação dos temas (*timeline*), o que faz determinado evento tornar-se importante nas redes sociais, quanto tempo é necessário para que um evento tenha impacto nas redes sociais e na sociedade, qual o papel dos "principais atores" na propagação dos eventos, é necessária a recolha, processamento e análise de informações partilhadas nas redes sociais.

O objetivo principal deste trabalho está relacionado com o desenvolvimento de uma arquitetura para recolha, processamento e armazenamento da informação partilhada no Twitter em Portugal e escrita em Português Europeu, resultando num Sistema de Informação que agrega quatro módulos para recolha, armazenamento, acesso e visualização dos dados. Este artigo encontra-se estruturado nas seguintes secções: a Secção 2 introduz de forma resumida o Sistema de Informação resultante deste trabalho; nas Secções 3 e 4 são descritos os processos de recolha, armazenamento e processamento de informação extraída do Twitter; as Secções 5 e 6 apresentam as ferramentas desenvolvidas para acesso e visualização dos dados e por fim, na Secção 7, são mencionadas as principais conclusões e as propostas para continuação do trabalho desenvolvido.

2. Sistema de Informação

A metodologia para recolha, processamento, armazenamento, acesso e visualização de dados do Twitter apresentada neste artigo culminou no desenvolvimento de um Sistema de Informação que poderá permitir a realização de diversos tipos de estudos e aplicações baseadas nos dados recolhidos. A Figura 1 apresenta o fluxo da informação entre os quatro módulos que constituem o Sistema de Informação: no módulo *Recolha* são capturados tweets geolocalizados em Portugal e destes são descartados os tweets cujo texto não se encontra escrito em Português Europeu; no módulo *Expansão* é recolhida a timeline de cada um dos utilizadores portugueses identificados no módulo *Recolha*, conduzindo à expansão da informação capturada sobre cada utilizar; o módulo *Acesso* possibilita o acesso ao *corpus* através de uma REST API que abstrai o acesso à base de dados; por último, o módulo *Visualização* apresenta na forma de um Dashboard Web algumas métricas e indicadores sobre os dados recolhidos. Nas Secções seguintes descrevem detalhadamente cada um dos módulos.

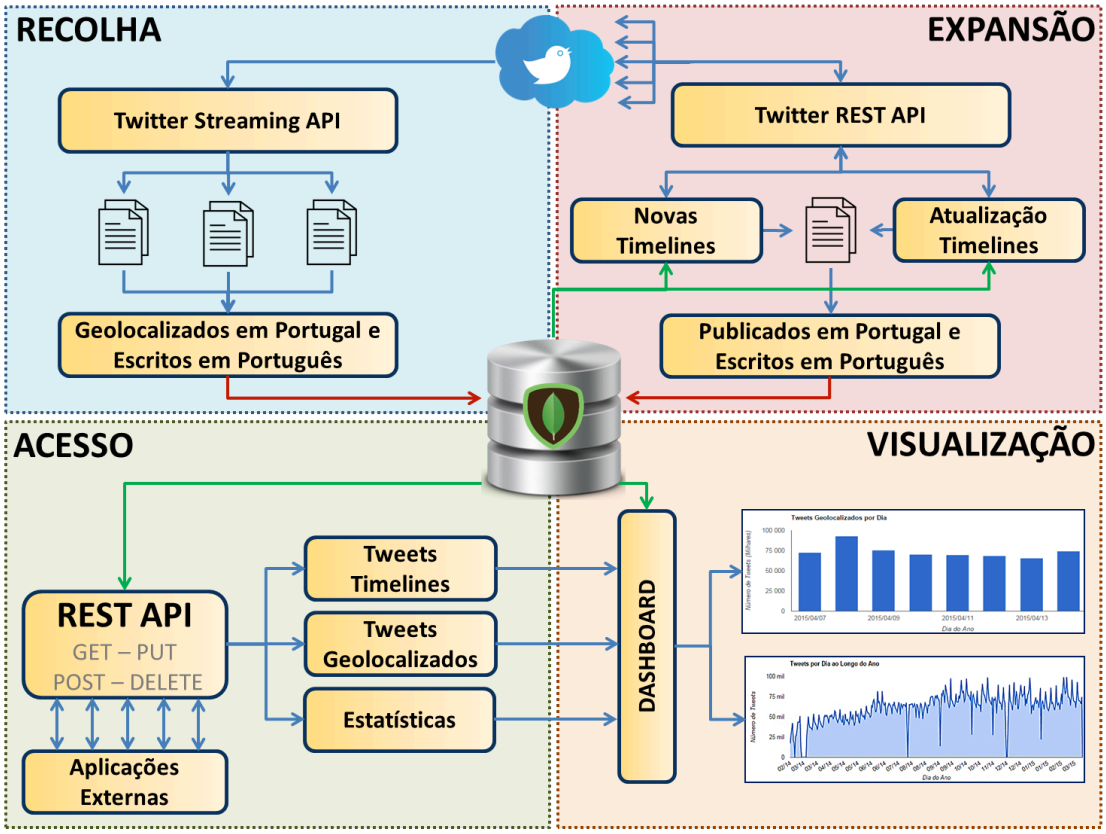


Figura 1: Sistema de informação para recolha e visualização de dados do Twitter.

3. Recolha e Armazenamento de Dados Geolocalizados

A localização geográfica de um tweet pode ser obtida seguindo uma de duas abordagens:

A) Geolocalização: os utilizadores podem optar por tornar pública a informação da sua localização no momento da publicação de um novo tweet. Esta informação pode ser bastante precisa se o tweet for publicado com recurso, por exemplo, a um *smartphone* com GPS.

B) Perfil do Utilizador: a localização de determinado utilizador pode ser extraída ou inferida pela informação contida no campo *location* do perfil de utilizador. Esta informação é igualmente disponibilizada na API do Twitter.

Tendo como objetivo a recolha de tweets geolocalizados em Portugal foi explorada a primeira abordagem, recorrendo à informação disponibilizada pela Twitter API, nomeadamente a Streaming API *statuses/filter*[2]. O acesso aos dados disponibilizados pela Twitter API só é

[2] https://dev.twitter.com/streaming/reference/post/statuses/filter

permitido após autenticação pelo protocolo aberto OAuth[3]. Após autenticação e envio do primeiro pedido HTTP para a Streaming API *statuses/filter* é disponibilizado o acesso ao fluxo de tweets produzidos em tempo real, permanecendo ativo o acesso ao fluxo sem necessidade de posterior interação por parte da aplicação cliente que iniciou a conexão à referida API.

O armazenamento dos tweets provenientes da *statuses/filter* API é efetuado continuamente em ficheiros comprimidos em disco. A estrutura de ficheiros utilizada para armazenar os tweets é criada dinamicamente sendo os ficheiros organizados em diretórios cuja nomenclatura é constituída pela referência ao ano, mês e ao dia em que os dados são recolhidos. A cada hora é criado um novo ficheiro guardado no diretório correspondente ao ano, mês e dia em que os tweets são capturados.

O fluxo de tweets retornado pela *statuses/filter* API está de acordo com um ou mais filtros. Esta API permite a pesquisa por palavras-chave, *hashtags*, *id* de utilizador ou regiões delimitadas geograficamente [Kumar *et al.* 2013a]. Porém, genericamente, o Twitter impõe limitações à utilização da sua API tanto no número de parâmetros como na quantidade de resultados devolvidos. Atualmente, cada pedido à *statuses/filter* API admite como parâmetros no máximo 400 palavras-chave, 25 delimitações de regiões geográficas ou 5000 *id* de utilizador. Como resultado são retornados todos os tweets que satisfaçam as restrições do pedido até um certo limite, que no caso da *statuses/filter* API é de 1% do volume total de tweets produzidos no Twitter em determinado instante [Kumar *et al.* 2013a]. Considerando que diariamente são produzidos a nível mundial aproximadamente 500 milhões de tweets [Twitter 2015] poderão teoricamente ser recolhidos da s*tatuses/filter* API cerca de 5 milhões de tweets por dia.

Em resultado de estudos prévios a este artigo concluiu-se que em Portugal são produzidos diariamente em média cerca de 60 mil tweets geolocalizados e escritos em Português Europeu [Brogueira *et al.* 2015], pelo que o volume de tweets que diariamente é possível recolher da s*tatuses/filter* API é bastante superior à quantidade de tweets geolocalizados que efetivamente são publicados em Portugal.

Na arquitetura proposta, o fluxo de tweets geolocalizados após ser guardado na estrutura de ficheiros comprimidos em disco referida anteriormente, é processado de modo a filtrar os tweets que não estejam escritos em Português Europeu ou que não tenham sido produzidos em Portugal. Esta filtragem é necessária visto que embora na invocação da *statuses/filter* API sejam indicadas as delimitações geográficas de Portugal, a dificuldade de representação exata dos limites geográficos de Portugal tem como consequência a recolha de alguns tweets de regiões

[3] http://oauth.net/

geograficamente próximas de Portugal, tal como Espanha ou Marrocos. Estes tweets são filtrados pela validação do campo *lang*[4] que deverá conter o valor igual a "pt" e o campo *place.country*[5] que deverá conter o valor "Portugal".

A verificação dos campos *lang* e *place.country* permite também detetar os casos em que os algoritmos de deteção automática de linguagem utilizados pelo Twitter [Twitter 2013] não identificam de forma correta a língua em que a mensagem de determinado tweet foi escrita. Por vezes verifica-se a ocorrência de tweets cujo campo *lang* tem erradamente o valor "pt", visto que o campo *text* não se encontra escrito em Português embora o valor do campo *place.country* esteja aparentemente de acordo com a língua em que o tweet foi efetivamente escrito. A Figura 2 apresenta um excerto de um tweet cujo valor do campo *lang* não corresponde à língua em que o texto do tweet foi escrito.

```
{
   "text": "a mi o me aportas o te apartas.",
   "lang": "pt",
   "created_at": "2014-02-20 16:23:59",
   "user":{
      "geo_enabled": true
   },
   "place":{
      "full_name": "Lucena del Puerto, Huelva",
      "country": "España",
      "country_code": "ES"
   }
}
```

Figura 2: Excerto de um tweet cujo campo *lang* foi incorretamente classificado pelo Twitter.

Como resultado do processo de filtragem são considerados "válidos" todos os tweets cujo valor do campo *lang* é igual a "pt" e o campo *place.country* é igual a "Portugal". Estes tweets serão armazenados numa base de dados MongoDB que conterá apenas os tweets geolocalizados em Portugal e escritos em Português Europeu capturados da *statuses/filter* API.

Aos tweets validados no contexto deste trabalho antes de serem inseridos no MongoDB é-lhes adicionado um novo campo que se convencionou designar por *created_at_object* que conterá o mesmo valor que o campo *created_at* (que indica a data de publicação do tweet), mas

[4] *Lang* – Campo que identifica a língua em que o texto do tweet foi escrito, através de algoritmos de reconhecimento de língua do Twitter.

[5] *Place.Country* – Campo que identifica o país em que o tweet foi publicado.

convertido para um objeto do tipo *Date* que facilita a posterior pesquisa e manipulação dos tweets em base de dados. A Figura 3 apresenta um exemplo da alteração efetuada a cada tweet.

- Tweet original recebido do Twitter:

```
{
   “_id”: “___”,
   “created_at": “Sun Jan 04 00:00:14 +0000 2015”,
   (...)
}
```

- Adição do campo *created_at_object* antes de inserir no MongoDB:

```
{
   “_id”: “___”,
   “created_at”: “Sun Jan 04 00:00:14 +0000 2015”,
   “created_at_object”: ISODate(“2015-01-04T00:00:14Z”),
   (...)
}
```

Figura 3: Documento JSON que guarda a informação sobre um utilizador.

Paralelamente ao armazenamento dos tweets geolocalizados em MongoDB é registado numa outra coleção MongoDB alguma informação relativa aos autores dos tweets. Para cada autor é construído um objeto JSON com a estrutura apresentada na Figura 4.

```
{
   ”_id”: Number,
   ”timeline_status”: String,
   ”last_timeline_status_date”: ISODate,
   ”user_date”: ISODate
}
```

Figura 4: Documento JSON que guarda a informação sobre um utilizador.

O campo *_id* corresponde ao identificador único do utilizador que se encontra no campo *user.id* de cada tweet, associando o tweet ao respetivo autor. Os campos *timeline_status* e *last_timeline_status_date* permitem controlar o estado do utilizador no âmbito da arquitetura apresentada neste artigo cujo significado será detalhado na Secção 4. No campo *user_date* é guardada a data em que o utilizador foi identificado pela primeira vez como tendo publicado um tweet geolocalizado capturado pelo procedimento apresentado neste artigo.

4. Expansão da Base de Dados

O conjunto de utilizadores cuja informação é armazenada pelo método apresentado na Secção 3 irá permitir a expansão da base de dados de tweets pela recolha da timeline de cada um dos utilizadores, isto é, serão lidos os tweets do histórico da atividade recente de todos os

utilizadores. A timeline está acessível através da REST API *statuses/user_timeline*[6] que dado um *id* de utilizador retorna os últimos 3200 tweets produzidos pelo respetivo utilizador. O procedimento de recolha da timeline de todos os utilizadores não é um processo trivial, considerando as restrições impostas pelo Twitter na utilização da API *statuses/user_timeline*. De uma forma geral o número de acessos permitidos à API do Twitter é contabilizado em períodos de 15 minutos sendo autorizados 180 pedidos à API em cada período utilizando autenticação do nível de utilizador ou 300 pedidos no caso de autenticação com nível de aplicação [Kumar *et al.* 2013a]. No presente trabalho foram utilizadas somente contas de acesso com autenticação do nível de utilizador. Em cada invocação à API *statuses/user_timeline* são retornados no máximo 200 tweets pelo que a leitura dos 3200 tweets da timeline necessita de 16 pedidos à API. Considerando que com uma conta de autenticação de nível de utilizador é possível efetuar 180 invocações à API *statuses/user_timeline* em cada período de 15 minutos e assumindo o caso limite em que todos os utilizadores têm na sua timeline 3200 tweets, é recolhida a timeline completa de cerca de 11 utilizadores a cada intervalo de 15 minutos. Neste cenário serão recolhidas 45 timelines por hora e 1080 timelines completas por dia. Com esta limitação é imperativo que o método de leitura das timelines seja otimizado de modo a minimizar o desperdício de tempo entre invocações, evitando não só a repetição de pedidos como a falha de invocações devido ao excesso de solicitações num dado período. O método de expansão do *corpus* apresentado nesta Secção é executado continuamente procedendo à recolha, processamento e armazenamento da timeline completa de cada novo utilizador integrado na base de dados e à pesquisa dos novos tweets publicados desde a última atualização das restantes timelines.

De acordo com o referido na Secção 6, são incluídos na base de dados diariamente em média 232 novos utilizadores. Considerando que o universo de utilizadores apresenta um crescimento constante e contando já com mais de 100 mil utilizadores, à data da escrita deste artigo, atualizar continuamente todas as *timelines* utilizando apenas uma conta para acesso à Twitter API, facilmente se depreende que o intervalo de tempo necessário para atualizar todas as timelines é bastante elevado e tendencialmente crescente. Com o intuito de minimizar o período entre duas iterações completas para atualização de todas as timelines, foram criadas e registadas diversas contas para acesso à Twitter API. Sobre cada conta é guardada a informação apresentada na Figura 5 onde os campos *ConsumerKey* e *ComsumerSecret* identificam univocamente a aplicação junto do Twitter e os campos *AccessToken* e *AccessTokenSecret* contêm a chave de autenticação no Twitter.

[6] https://dev.twitter.com/rest/reference/get/statuses/user_timeline

```
{
    “_id”: String,
    “user”: Integer,
    “client”: Integer,
    “timestamp”: ISODate,
    “timeline_status”: String,
    “APIKey”: String,
    “APISecret”: String,
    “AccessToken”: String,
    “AccessTokenSecret”: String,
}
```

Figura 5: Documento JSON com as credenciais de acesso à API do Twitter.

O campo *timeline_status* pode assumir diversos valores estando diretamente relacionado com o campo *timeline_status* da informação associada a cada utilizador (*cf.* Figura 4), uma vez que cada conta recolhe a timeline dos utilizadores cujo valor do campo *timeline_status* é igual ao valor do campo com o mesmo nome na informação do utilizador. Deste modo de entre o conjunto de contas para processamento das timelines existem contas dedicadas à recolha da timeline dos novos utilizadores e contas dedicadas à atualização das timelines dos utilizadores já incluídos na base de dados. Dado que o número de novos utilizadores integrados por dia é teoricamente inferior ao máximo de timelines que é possível processar com apenas um cliente, é portanto suficiente que apenas uma conta seja dedicada à recolha da timeline dos novos utilizadores, ficando os restantes clientes alocados à tarefa de atualização das timelines.

O procedimento de leitura das timelines tem em consideração o estado indicado no campo *timeline_status*, associado à informação de cada utilizador permitindo categorizar cada utilizador relativamente ao estado da recolha da respetiva timeline. Este campo pode assumir os seguintes estados: “Leitura Integral”, “Atualizar”, “Em atualização”, “Bloqueado”, “Casual” ou “Erro” cujo significado é apresentado na Tabela 1.

Tabela 1: Estados da recolha da *timeline* de um utilizador.

Estado	Descrição
Leitura Integral	Recolher todos os tweets da *timeline*
Atualizar	Recolher os novos tweets da *timeline*
Em atualização	“Bloqueia” o utilizador durante a atualização
Bloqueado	O utilizador não permite o acesso à *timeline*
Casual	Os novos tweets não foram publicados em Portugal
Erro	O processo de recolha terminou em erro

Quando um novo utilizador é adicionado à base de dados de utilizadores o campo *timeline_status* é colocado por defeito no estado “Leitura Integral”, indicando que o utilizador

foi introduzido no sistema e que se deverá proceder à recolha da totalidade da sua timeline. Concluída a recolha ou atualização da timeline, o campo *last_timeline_status_date* é atualizado com a data correspondente ao dia em que o processamento foi realizado. Este facto é bastante relevante uma vez que o processo de atualização das timelines dá prioridade aos utilizadores cujo campo *last_timeline_status_date* contém a data mais antiga.

O diagrama de transição de estados do processo de recolha da timeline é apresentado na Figura 6. Quando determinado utilizador é selecionado para que seja recolhida ou atualizada a respetiva timeline o seu estado é alterado para "Em Atualização". Deste modo evita-se que um mesmo utilizador possa estar a ser processado por mais do que um cliente em simultâneo. No caso do processamento da timeline terminar como esperado o campo *timeline_status* é colocado no estado "Atualizar" fazendo com que na próxima iteração do procedimento, sejam apenas pesquisados os novos tweets publicados por determinado utilizador desde o seu tweet mais recente, guardado na base de dados. O campo *last_timeline_status_date* é igualmente atualizado com a data do dia do processamento. No entanto, o estado "Em Atualização" pode originar a passagem a outros estados, nomeadamente, ao estado de "Erro" caso ocorra alguma situação inesperada durante a execução do processo; ao estado "Bloqueado" quando o utilizador não permite o acesso público à timeline ou ao estado "Casual" para o caso de utilizadores que já publicaram tweets em Portugal e escritos em português mas cujas posteriores mensagens não foram publicadas em Portugal ou escritas numa outra língua que não o português. A passagem a um destes estados não implica a atualização do campo *last_timeline_status_date.* Também para os estados de erro existe uma conta dedicada que periodicamente testa a possibilidade de "recuperação" dos utilizadores nestes estados. Quando o campo *user_date* permanece igual ao campo *last_timeline_status_date* significa que o utilizador nunca passou ao estado "Atualizar" pelo que a tentativa de "recuperação" pressupõe a passagem para o estado "Leitura Integral". Por outro lado, quando o campo *user_date* de determinado utilizador difere do campo *last_timeline_status_date* significa que em dado momento o processamento integral da sua timeline foi efetuado com sucesso pelo que a tentativa de "recuperação" irá recolocar o utilizador no estado "Atualizar".

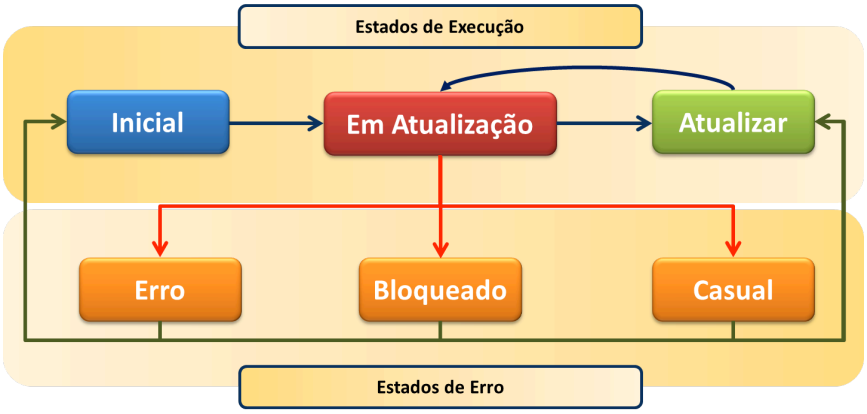


Figura 6: Diagrama de transição de estados na recolha da *timeline*.

À imagem do procedimento de recolha dos tweets geolocalizados os tweets recolhidos das timelines são armazenados em ficheiros comprimidos em disco. Neste caso os ficheiros são guardados numa estrutura organizada de forma diferente, onde o nome de cada diretório é composto pela sequência de caracteres 000, 001, 002 … 999. Por cada invocação da API *statuses/user_timeline* é gerado um ficheiro com os tweets devolvidos na resposta da API, sendo este guardado no diretório correspondente aos três últimos dígitos do *id* de utilizador processado. Esta nomenclatura permite uma distribuição uniforme dos ficheiros por cada diretório, visto que com a elevada quantidade de invocações à API e consequente número de ficheiros criados, rapidamente se atingiria o limite de ficheiros que é possível armazenar num só diretório caso todos eles fossem guardados em apenas um diretório.

Na Figura 7 é apresentado o fluxograma que resume a arquitetura proposta para recolha e expansão de uma base de dados de tweets produzidos em Portugal e escritos em português europeu. A parte esquerda representa a recolha dos tweets geolocalizados e a parte direita corresponde ao procedimento de leitura das timelines dos utilizadores portugueses que publicam tweets com recurso a dispositivos que permitem partilhar a sua geolocalização.

5. Acesso aos Dados

Os tweets recolhidos pelo mecanismo apresentado neste artigo não constituirão uma mais-valia considerando a informação que deles se pode retirar, se não forem processados e analisados. Por vezes este tipo de análises é efetuado por especialistas de áreas cujo conhecimento de conceitos como a representação dos tweets em objetos JSON ou a sintaxe da linguagem para pesquisa de informação numa base de dados MongoDB não é muito acentuado, constituindo uma dificuldade acrescida. Nesse sentido foi implementada uma REST API que pode ser vista como uma forma de abstrair todos estes conceitos, facilitando o acesso aos dados.

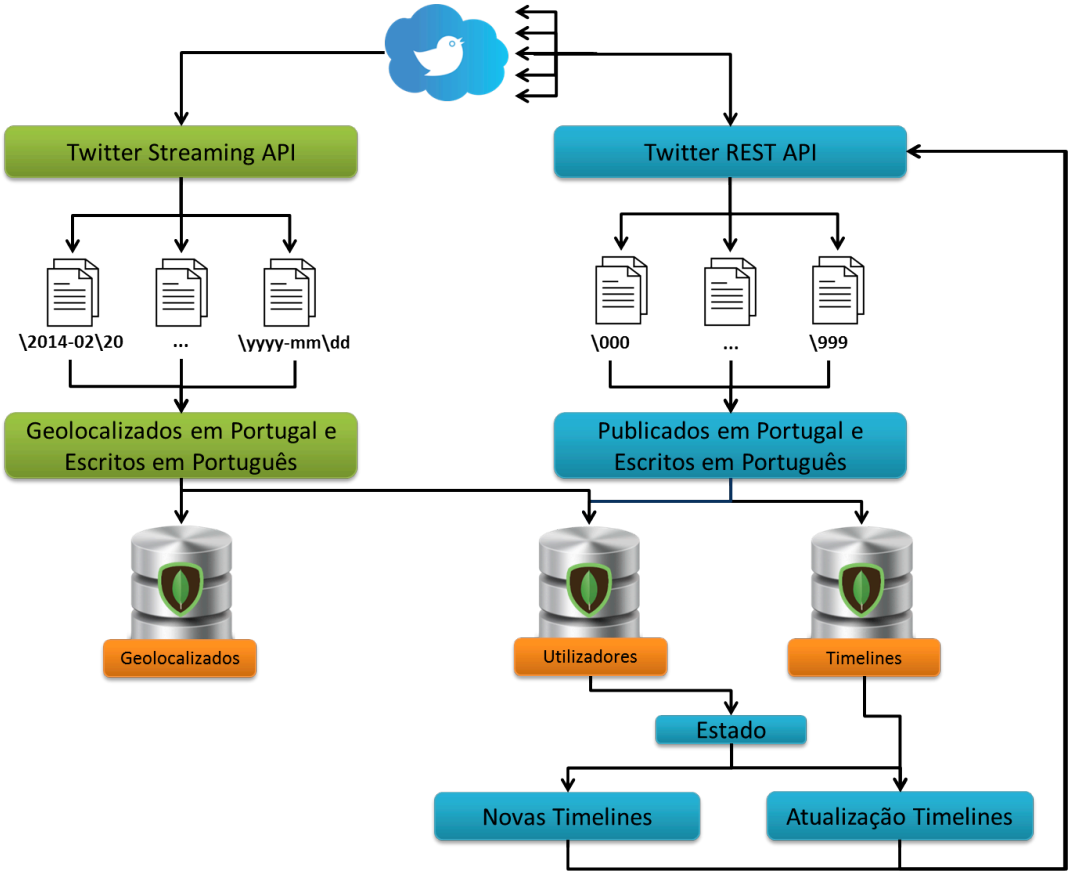


Figura 7: Arquitetura de recolha e expansão da base de dados de Tweets.

O conceito de REST (*Representational State Transfer*) apresentado por Fielding [2000] refere-se a um conjunto de princípios de arquitetura de interfaces web. De um modo geral uma API expõe um conjunto de dados e funções para facilitar a interação e a troca de informações entre aplicações e *web services*. Uma web API que é desenhada em conformidade com os princípios de arquitetura REST é designada por REST API.

No contexto deste Sistema de Informação foi implementada uma REST API que disponibiliza um conjunto de serviços (*endpoints*) através dos quais é permitido o acesso à informação guardada na base de dados MongoDB. Os serviços implementados podem ser vistos como uma camada de abstração em relação à base de dados MongoDB permitindo o desenvolvimento de aplicações baseadas na informação nela armazenada, sem a necessidade de conhecimento do modelo de dados ou da sintaxe para pesquisa e extração de dados. A REST API foi implementada com recurso à microframework Flask[7] que permite o desenvolvimento de aplicações web em Python. Os serviços disponibilizados com a REST API permitem apenas a consulta e leitura de dados, podendo ser divididos em três grupos: acesso à coleção de tweets,

[7] http://flask.pocoo.org/

acesso a informações sobre os utilizadores e acesso a estatísticas e indicadores gerais sobre os dados. Para acesso à coleção de tweets os serviços mais relevantes são indicados na Tabela 2.

Tabela 2: Serviços disponíveis na REST API para acesso aos tweets.

Nome Serviço	Parâmetros	Dados retornados pelo serviço
/api/<collection>/tweet/	<int: tweet_id>	Tweet correspondente ao *id* indicando em "tweet_id"
/api/<collection>/page/	<int: page_id>	Conjunto de 1000 tweets, ordenados decrescentemente pela ordem de publicação
/api/<collection>/day/	<int: day>	Tweets produzidos no dia indicado em "day"
/api/<collection>/user/	<int: screen_name>	Todos os tweets produzidos pelo utilizador indicado em "screen_name"
/api/<collection>/query/	<String: query>	Conjunto de tweets de acordo com o filtro especificado pelo parâmetro "query"

Dado que os tweets geolocalizados e os tweets lidos das timelines são armazenados em coleções de dados distintas, a definição do nome dos serviços da Tabela 2 é composto pelo nome da coleção (<collection>) que poderá assumir os valores "geolocated" ou "timeline", respetivamente. O acesso a informações relativas aos utilizadores é disponibilizado pelos serviços da Tabela 3.

Tabela 3: Serviços da REST API para acesso à informação de cada utilizador.

Nome Serviço	Parâmetros	Dados retornados pelo serviço
/api/user/state/	<int: user_id>	Estado em que o utilizador se encontra, no contexto deste trabalho. Ou seja, se a sua timeline já foi recolhida, se está em atualização ou se se encontra bloqueada pelo próprio utilizador
/api/user/first_profile/	<String: screen_name>	Perfil do utilizador @screen_name relativamente ao seu primeiro tweet incluído na base de dados
/api/user/last_profile/	<String: screen_name>	Perfil do utilizador @screen_name relativamente ao seu último tweet incluído na base de dados
/api/user/ageAndGender/	<String: screen_name>	Campos do perfil do utilizador que permitem inferir a idade e o género do utilizador indicado em @screen_name

Devido ao elevado volume de dados armazenados determinadas pesquisas podem tornar-se bastante demoradas. Para tal, o resultado de pesquisas mais complexas é pré-processado e guardado também em MongoDB permitindo a visualização de estatísticas sobre os dados de forma consideravelmente mais rápida. Algumas das pesquisas pré-processadas são disponibilizadas pelos serviços da Tabela 4.

Tabela 4: Serviços da REST API para acesso a estatísticas pré-processadas.

Nome Serviço	Dados retornados pelo serviço
/api/stats_geolocated_day/	Total por dia de tweets geolocalizados recolhidos (desde 20 de fevereiro de 2014)
/api/stats_geolocated_hour_day/	Soma dos tweets geolocalizados recolhidos por hora
/api/stats_geolocated_week_day/	Soma dos tweets geolocalizados recolhidos por dia da semana
/api/stats_timeline_day/	Soma dos tweets da timeline em cada dia (desde 20 de fevereiro de 2014)
/api/stats_users_day/	Número de novos utilizadores identificados por dia

Na Figura 8 é esquematizada a forma como através da REST API se realiza a ligação e extração de dados do MongoDB. A REST API permite a comunicação bidirecional com web browsers (Internet Explorer, Chrome, Firefox), com aplicações da linha de comandos (curl, http) ou com aplicações móveis (Android, iOS). As invocações à REST API são efetuadas mediante a realização de pedidos HTTP a cada *endpoint* sendo devolvida a informação solicitada, tal como é demonstrado na Figura 9. Na resposta às invocações é utilizada a biblioteca Python Jinja2[8] que auxilia na criação de templates HTML, disponibilizando os dados num formato adequado à visualização e interpretação dos mesmos.

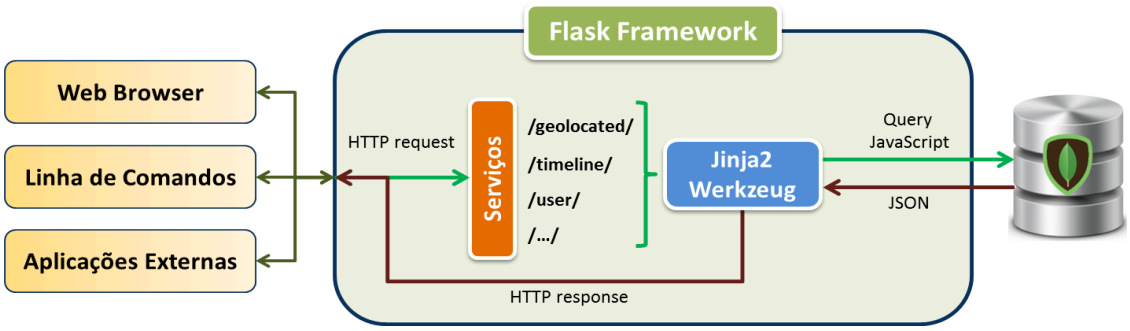


Figura 8: Interação entre aplicações externas e o MongoDB via REST API.

O exemplo da Figura 9 apresenta o resultado da invocação do *endpoint* api/geolocated/tweet/ passando como parâmetro o *id* de tweet 551528493021147136, obtendo-se como resposta o objeto JSON correspondente ao tweet.

[8] http://jinja.pocoo.org/docs/dev/

```
> curl -i http://localhost:27016/api/geolocated/tweet/551528493021147136
200 OK
Content-Type: text/html; charset = utf-8
Content-Length: 102
Server: Werkzeug/0.9.6 Python/2.7.8
Date: Sun, 26 Jul 2015 17:00:36 GMT
{
    "_id": "551528493021147136",
    "text": "Boa noite!",
    "created_at_object": "2015-01-04 00:00:14",
    (...)
}
```

Figura 9: Exemplo de utilização da REST API.

6. Visualização e Métricas sobre os Dados

A execução contínua do procedimento de recolha de tweets descrito neste artigo resulta na captação de um grande volume de informação, conduzindo à necessidade de agregar parte dessa informação num formato que permita uma visão geral de alguns indicadores de desempenho e performance de todo o Sistema de Informação. Justamente por esse motivo foi desenvolvido um Dashboard Web que resume parte da informação recolhida de uma forma clara, concisa e acima de tudo, visual. Recorrendo à API do Google para criação de gráficos Google Charts[9] e com a invocação dos serviços da REST API, descritos na Secção 5, relativos às estatísticas sobre os dados armazenados ou consultando diretamente a base de dados MongoDB foi desenvolvido um Dashboard. Um dos cinco quadros do Dashboard é resumidamente apresentado na Figura 10, onde se podem observar algumas estatísticas relativamente aos tweets geolocalizados. No gráfico 1 são indicados os tweets recolhidos por dia, no gráfico 2 é indicado o número de tweets recolhidos por hora e no gráfico 3 é indicado o número de tweets recolhidos para cada um dos dias da semana. No gráfico 4 observa-se o número de tweets recolhidos por dia ao longo do último mês de recolha de dados, que neste caso concreto corresponde a maio de 2015. O gráfico 5 apresenta a soma de tweets geolocalizados recolhidos por mês ao longo do período durante o qual foi efetuada a recolha de dados.

[9] https://developers.google.com/chart/

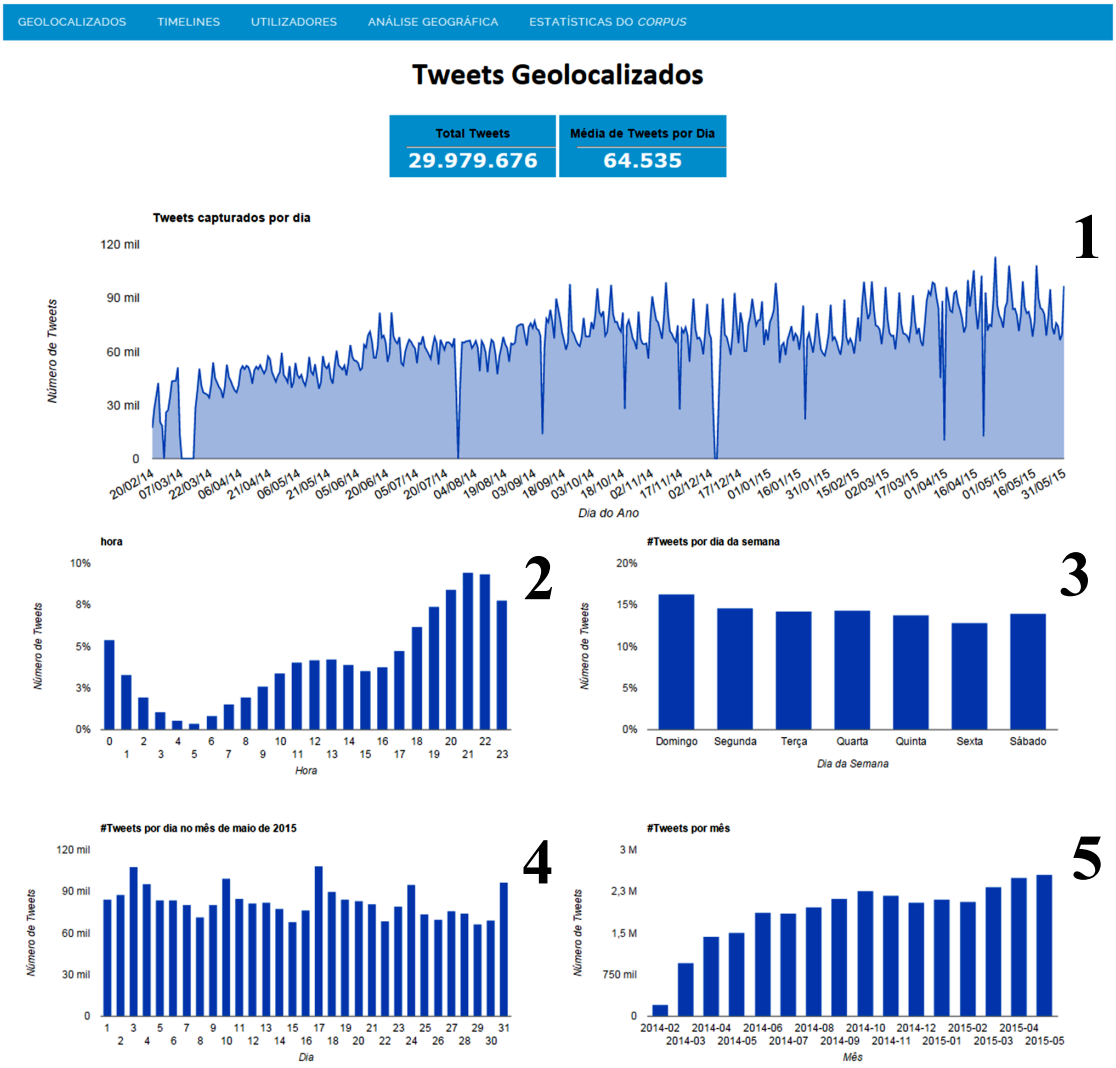


Figura 10: Dashboard para visualização de indicadores relativos aos dados recolhidos.

O processo de recolha de tweets foi iniciado a 20 de fevereiro de 2014 e terminou a 31 de maio de 2015. Na Figura 11 é apresentado o número de tweets geolocalizados recolhido por dia no período de fevereiro de 2014 a maio de 2015 observando-se um aumento do número de tweets capturados ao longo desse período. Os dias com cor branca deveram-se a problemas pontuais na recolha dos tweets. Neste período foram recolhidos aproximadamente 30 milhões de tweets geolocalizados com uma média diária de aproximadamente 66 mil tweets.

Pelo procedimento descrito na Secção 3 foram integrados em média 232 novos utilizadores por dia num total aproximado de 105 mil utilizadores. Analisando o gráfico da Figura 12 observa-se que o número de novos utilizadores por dia manteve-se constante ao longo de todo o período.

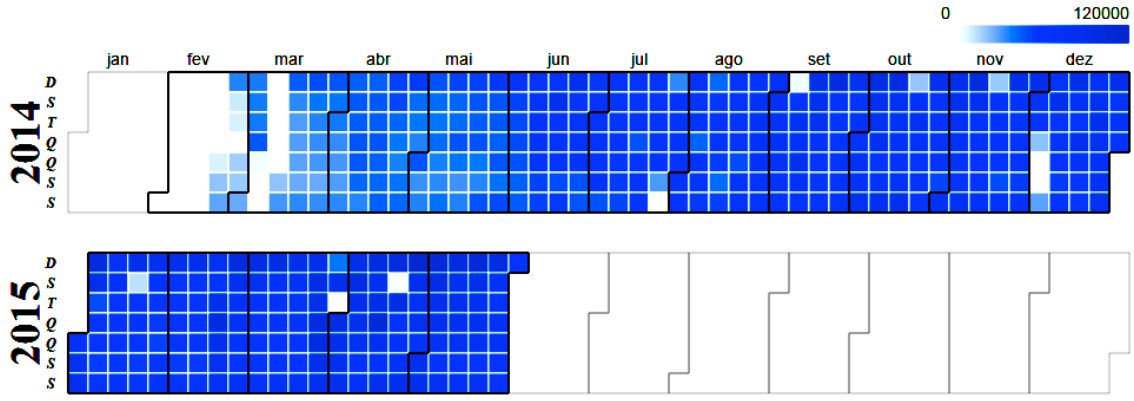


Figura 11: Distribuição da recolha de tweets geolocalizados por dia.

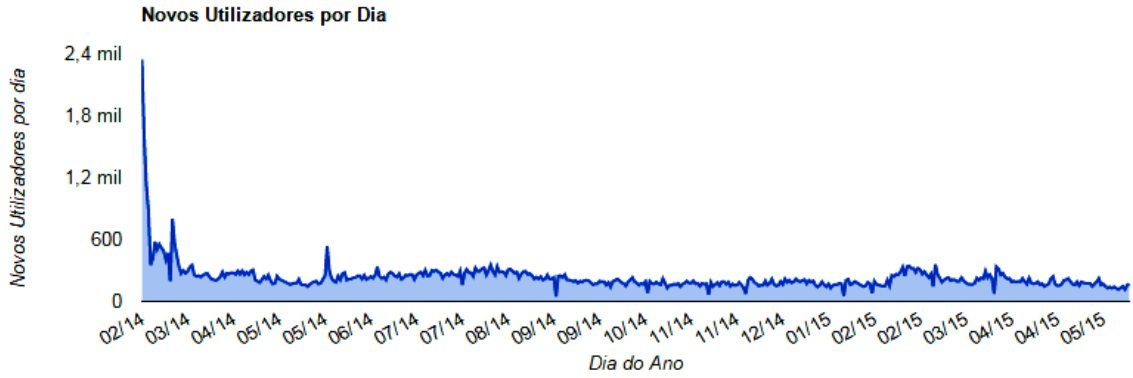


Figura 12: Novos utilizadores por dia.

Da execução do procedimento de expansão da base de dados descrito na Secção 4, resultou a recolha de cerca de 260 milhões de tweets. De entre os tweets armazenados em resultado da leitura das timelines cerca de 246 milhões correspondem ao intervalo de fevereiro de 2014 a maio de 2015, o que representa um acréscimo médio de cerca de 8 vezes relativamente ao número de tweets recolhidos do fluxo da Streaming API. Na Figura 13 apresenta-se a relação entre a quantidade de tweets geolocalizados e os tweets lidos das timelines. É notório o acréscimo de tweets recolhidos pelo mecanismo apresentado neste artigo, por comparação com volume de tweets disponibilizados pelo Twitter na Streaming API. A linha "paralela" ao eixo das abcissas representa o número de tweets geolocalizados recolhidos por dia. As restantes linhas representam a evolução do volume de tweets lidos das timelines em determinado instante, indicado pela data de cada uma das linhas na legenda da Figura 13. Cada uma das linhas resulta de uma iteração completa para a recolha da totalidade ou atualização da timeline dos utilizadores registados na base de dados.

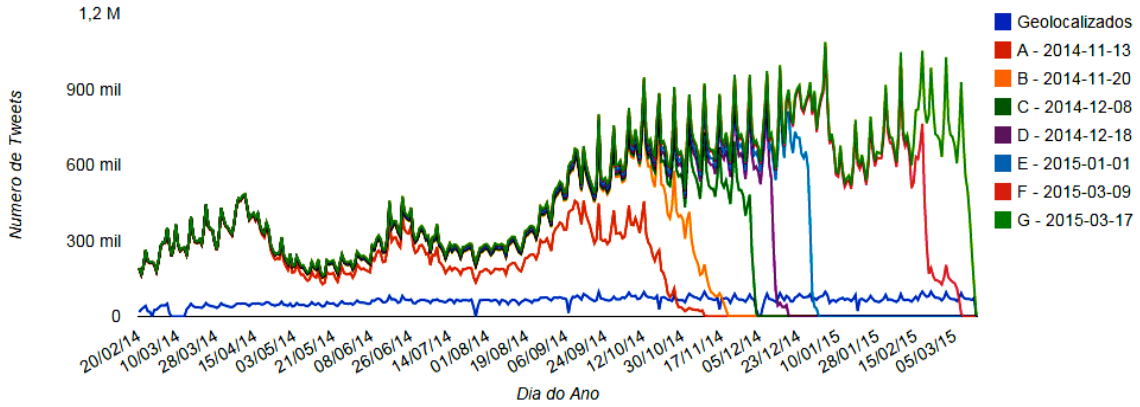


Figura 13: Comparação do volume de tweets obtidos da Streaming API e da REST API do Twitter.

Ainda relativamente à Figura 13 verifica-se que os máximos locais observados na oscilação da linha G da parte direita do gráfico correspondem aos primeiros dias da semana, nomeadamente aos dias de Domingo e Segunda-Feira e, por conseguinte, os mínimos locais correspondem aos dias do final da semana mais concretamente a Sexta-Feira. O dia 4 de janeiro de 2015 foi o dia para o qual se recolheram mais tweets pelo processo apresentado, com o total de 1.069.851 tweets enquanto que no mesmo dia foram recolhidos somente 98.366 tweets da Streaming API, representando um acréscimo de sensivelmente 11 vezes em relação ao fluxo de tweets geolocalizados.

7. Conclusões e Trabalho Futuro

Neste artigo é apresentado um Sistema de Informação que permite o desenvolvimento de uma base de dados de tweets escritos em Português Europeu. O mecanismo apresentado tenderá a armazenar grande parte dos tweets produzidos pela comunidade Portuguesa do Twitter dada a capacidade de identificar os utilizadores portugueses e de recolher continuamente as mensagens mais recentes produzidas pelos mesmos, tendo a capacidade de integrar continuamente novos utilizadores. A arquitetura desenvolvida permite a recolha de cerca de 8 vezes mais informação do que a disponibilizada em tempo real pelo Twitter no fluxo da Streaming API. A REST API implementada para acesso aos dados e visualização de estatísticas sobre os mesmos permite que a informação contida na base de dados possa ser partilhada e analisada em estudos das mais diversas áreas do conhecimento.

Na continuação deste trabalho tem-se como principal objetivo permitir a utilização do Sistema de Informação pelo método desenvolvido por Rosa *et al.* [2014] como fonte de dados para identificação e análise de tópicos, assim como dos utilizadores mais importantes relativamente a esses tópicos. Além da informação que este método poderá obter do conjunto de dados recolhido irá permitir a validação da arquitetura proposta. Também está previsto o

enriquecimento da informação obtida de cada utilizador pela adição da idade e género (masculino ou feminino) correspondente a cada utilizador, através do método proposto por Vicente *et al.* [2015]. Para tal, será necessária a implementação de um *endpoint* na REST API que possibilite a alteração e adição de dados na base de dados MongoDB.

Agradecimentos

Este trabalho foi financiado com fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia pelo projeto PTDC/IVC-ESCT/4919/2012 (MISNIS) e fundos com referência UID/CEC/50021/2013.

8. Referências

Anderson, K., e Schram, A., "Design and implementation of a data analytics infrastructure in support of crisis informatics research.", *ICSE*, ACM, (2011), 844-847.

Boyd, D. M. e Ellison, N. B., "Social network sites: definition, history, and scholarship". *Journal of Computer-Mediated Communication*, (2007), 210–230.

Brogueira, G., Batista, F., e Carvalho, J. P., Arquitetura e Desenvolvimento de um Repositório de Tweets em Português Europeu, *5ª Jornadas de Informática da Universidade de Évora*, JIUE'15, (2015), Universidade de Évora.

Culotta, A., "Towards Detecting Influenza Epidemics by Analyzing Twitter Messages", in *Proceedings of the First Workshop on Social Media Analytics* ACM, (2010), 115-122.

Fielding, R., Architectural styles and the design of network-based software architectures, *PhD thesis*, (2000), Dept. of Information and Computer Science, University of California.

Gerber , M., "Predicting crime using Twitter and kernel density estimation", *Decision Support Systems*, (2014), 115 – 125.

Kumar, S., Morstatter, F., e Liu, H., "Twitter Data Analytics". *Springer*, (2013a).

Kumar, S., Morstatter, F., Zafarani, R., e Liu, H., "Whom Should I Follow?: Identifying Relevant Users During Crises", *Proceedings of the 24th ACM Conference on Hypertext and Social Media*, ACM, (2013b), 139-147.

Lachlan, K., Spence, P., Lin, X., "Expressions of risk awareness and concern through Twitter: On the utility of using the medium as an indication of audience needs", *Computers in Human Behavior* (2014), 554-559.

Laudon, K. C. e Laudon, J. P., "Sistemas de informação: com Internet", *LTC Editora*, (1999).

Marcus, A., Bernstein, M., Badar, O., Karger, D., Madden, S., e Miller, R., "Twitinfo: Aggregating and Visualizing Microblogs for Event Exploration", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, (2011), 227-236.

Mendoza, M., Poblete, B. e Castillo, C., "Twitter Under Crisis: Can We Trust What We RT? " *Proceedings of the First Workshop on Social Media Analytics*, ACM, (2010), 71-79.

Oussalah, M., Bhat, F., Challis, K., e Schnier , T., "A software architecture for Twitter collection, search and geolocation services", *Knowledge-Based Systems* (2013), 105-120.

Paul, Michael J. e Dredze, Mark, "You Are What You Tweet: Analyzing Twitter for Public Health". *ICWSM, The AAAI Press*, (2011).

Perera, R.D.W., Anand, S., Subbalakshmi, K.P., e Chandramouli, R., "Twitter analytics: Architecture, tools and analysis", in *MILITARY COMMUNICATIONS CONFERENCE, 2010 - MILCOM 2010* (2010), 2186-2191.

Qu, Y., Huang, C., Zhang, P., e Zhang, J., "Microblogging After a Major Disaster in China: A Case Study of the 2010 Yushu Earthquake", *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* , ACM, (2011), 25-34.

Rosa, H., Carvalho, J. P., e Batista, F., "Detecting a Tweet's Topic within a Large Number of Portuguese Twitter Trends". *In Pereira, M. J. V., Leal, J. P., and Simões, A., editors, 3rd Symposium on Languages, Applications and Technologies*, volume 38 of OpenAccess Series in Informatics (OASIcs), Dagstuhl, Germany, (2104), 185–199.

Sakaki, T., Okazaki, M. e Matsuo, Y., "Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors", *Proceedings of the 19th International Conference on World Wide Web*, ACM (2010), 851-860.

Santos, J. C. e Matos, S., "Predicting flu incidence from Portuguese tweets", *In Proceedings of the International Work-Conference on Bioinformatics and Biomedical Engineering*, IWBBIO 2013, Granada, Spain, (2013), 11-18.

Santos, J. C. e Matos, S., "Analysing twitter and web queries for flu trend prediction", *Theoretical Biology and Medical Modelling*, (2014), 11(1).

Scanfeld, D., Scanfeld, V. e Larson, E., "Dissemination of health information through social networks: Twitter and antibiotics", *American Journal of Infection Control* 38, 3 (2010), 182-188.

Twitter, Inc., https://about.twitter.com/company, (2015) (acedido em 14-05-2015).

Twitter, Blog, “Introducing new metadata for Twitter”, https://blog.twitter.com/2013/introducing-new-metadata-for-tweets, (2013) (acedido em 14-05-2015).

Vicente, M., Batista, F., e Carvalho, J. P., “Twitter gender classification using user unstructured information”. *In FUZZ-IEEE 2015, IEEE International Conference on Fuzzy Systems*, IEEE Xplorer, Istanbul, Turkey, (2015).

Widener, M., e Wenwen, L., “Using geolocated Twitter data to monitor the prevalence of healthy and unhealthy food references across the {US}”, *Applied Geography* 54, (2014), 189-197.

Um Método para Identificar a Não Conformidade entre os Modelos de Processos de Negócio e o Conceitual de Informação

Rafael Batista Duarte [1], Denis Silva da Silveira [2], Fernando Wanderley [3]

1) Universidade de Pernambuco, Brasil

rbd@ecomp.poli.br

2) Universidade Federal de Pernambuco, Brasil
dsilveira@ufpe.br

3) Universidade Nova de Lisboa, Portugal
fernando.wanderley@gmail.com

Resumo

A falta de conformidade entre as ações que são realizadas em um processo de negócio e a estrutura definida no modelo conceitual de informação vem se tornando uma preocupação frequente nas organizações. Visto que, o quanto antes identificada a falta de conformidade torna-se menos custoso sua correção. Este artigo apresenta um método para identificar a não conformidade entre os modelos conceitual de informação e o de processo de negócio. A identificação é realizada por meio de animações dos objetos instanciados a partir do modelo conceitual de informação com base nas descrições estabelecidas no modelo de processo de negócio. O método aqui proposto foi automatizado por uma ferramenta que interopera com a ferramenta USE, que analisa as violações estruturais definidas no modelo conceitual de informação.

Palavras chave: *Business Process; Conceptual Model of Information; Models Conformity; Models Animation*.

1. Introdução

Atualmente, o sucesso de uma organização está condicionado à eficácia com que os seus processos de negócio são executados [Silveira 2009]. Um sistema de informação desenvolvido para dar suporte a uma organização deve estar em conformidade com os processos de negócio onde este está inserido [Silveira 2009], [Olivé 2007], [Jalali 2011]. As especificações desses sistemas de informação, frequentemente, são criadas sem que haja o real entendimento das necessidades ou os problemas da organização. No entanto, uma vez que processos de negócio não estão em conformidade com o modelo conceitual de informação existe uma perda para a organização, gerando assim uma diminuição na produtividade que os sistemas de informação poderiam trazer para a organização. Segundo Olivé, o modelo conceitual de informação constitui uma parte fundamental para o conhecimento das organizações, determinando um conjunto de suposições (conceitos) que são identificados nas regas de negócio de uma organização [Olivé 2007].

A implantação de um processo de negócio em uma organização, normalmente, consome uma grande quantidade de recursos, podendo estes serem financeiros, materiais, de mão de obra e/ou equipamentos [Silveira 2009]. Logo, é de grande utilidade para as organizações o uso de algum método que possa identificar, o quanto antes da implantação, as inconsistências entre os processos de negócio e o modelo conceitual; uma vez que a execução desses processos, quando inconsistentes com a estrutura do negócio, pode acarretar em um custo muito maior para as organizações. Visto que, os modelos de processos de negócio e o modelo conceitual de informação constituem uma parte fundamental para a base de conhecimento do negócio, determinando, por exemplo, como o mesmo funciona. Além disso, esses modelos podem ser considerados como artefatos que fomentarão o projeto de desenvolvimento de um sistema de informação, pois eles deverão ser implementados nas aplicações automatizadas que darão suporte ás operações de uma organização.

Segundo o estudo realizado pelo *Standish Group* a maioria (66%) dos sistemas de informação já implementados não atenderam às expectativas dos usuários no que se referem a seus comportamentos e/ou funcionalidades [Standish Group 2011]. Brooks já afirmava, no final da década de 80, que a etapa mais difícil no desenvolvimento de um sistema de informação é entender, compreender e definir precisamente o que se vai construir [Brooks 1987]. Além desse alto índice de insatisfação, o referido estudo ainda aponta como um dos motivos à dificuldade de estabelecer uma comunicação de comum entendimento entre o especialista de negócio, aquele que entende do negócio e o profissional da área de tecnologia, responsável por consolidar detalhes técnicos do negócio.

De fato, essa dificuldade na comunicação acarreta em informações perdidas, que implica na dificuldade em representar, de forma fiel, a organização em um conjunto de modelos que estejam em conformidade entre si. Neste contexto, com o objetivo de minimizar esse problema, este artigo propõe um método para identificar a falta de conformidade entre o modelo de processos de negócio e o conceitual de informação utilizando para isso uma abordagem de animação. Mais especificamente, este trabalho apresenta um método automatizado para apoiar a elaboração e validação de modelos de processo e conceitual de informação, tendo como alvo inicial modelos PIM (modelos independentes de plataforma) em sistemas de informação transacionais.

O presente artigo encontra-se estruturado nas seguintes seções: a seção 2 aborda a fundamentação teórica necessária para o entendimento do método; a seção 3 apresenta a método proposto; e por fim, a Seção 4 apresenta as considerações finais.

2. Fundamentação Teórica

Modelagem e Modelos

A modelagem, no sentido mais amplo, é o uso econômico de algo (o modelo) no lugar de alguma coisa real, tendo em vista algum objetivo cognitivo [Silveira 2009]. A elaboração de modelos é uma técnica amplamente empregada em diversas modalidades de engenharia [Hazelrigg 1999]. Esta prática permite o uso de algo mais simples, seguro e barato do que o sistema real, para o estudo do objetivo desejado.

Um modelo é uma representação simplificada de algum conceito ou situação, com os objetivos de sua observação, manipulação e entendimento [Mellor *et al*. 2004]. Nesse sentido, modelos derivam da necessidade humana de entender a realidade, aparentemente complexa, e são, portanto, representações simplificadas e inteligíveis do mundo, permitindo vislumbrar as características essenciais de um domínio ou campo de estudo.

No desenvolvimento de software, tal como em outras aplicações, os modelos são criados com o objetivo de diminuir a complexidade inerente aos temas de suas aplicações [Silveira 2009]. Nesse contexto, modelos possibilitam a visualização, a comunicação e a validação de aspectos de um sistema antes da sua construção.

Modelagem Conceitual de Informação

Modelos conceituais são artefatos produzidos com o objetivo de representar uma dada porção da realidade segundo uma determinada conceituação [Olivé 2007]. Nesse contexto, a modelagem conceitual de informação consiste na representação dos objetos, suas características e seus relacionamentos no contexto de determinado ambiente (organização). Esta representação independe do uso de técnicas de implementação, tecnologias ou dispositivos físicos [Olivé 2007], [Larman 2004]. A representação de uma organização, usando-se a modelagem conceitual, ocorre por meio das entidades ou classes, suas características e seus relacionamentos. Uma entidade/classe é uma categoria atribuída ao conjunto de objetos existentes na organização que estão agrupados em função de suas semelhanças. Os objetos representados passam a ser denominadas instâncias destas entidades/classes [Schimitz 2000].

De forma geral, o propósito da modelagem conceitual é classificar as entidades/classes de um domínio, definir seu vocabulário e explicitar as relações de significado existentes a fim de que possam ser interpretadas computacionalmente através de regras lógicas.

Modelos conceituais são escritos em linguagens ditas linguagens de modelagem conceitual. Neste trabalho, a linguagem utilizada para esta definição é a UML [OMG 2005], por ser ela, por meio do diagrama de classes, capaz de representar o conhecimento sobre o domínio, incluindo seus conceitos e relacionamentos.

Modelagem de Processo de Negócio

Processo é definido por qualquer atividade ou conjunto de atividades bem definidas que seguem uma sequência e que dependem umas das outras numa sucessão clara, tomando entradas adiciona valor a ela e fornecendo uma saída a um cliente específico [Gonçalves 2000]. Ainda segundo esse autor, existem três tipos de processos, sendo um deles os processos de negócio, que caracterizam a atuação da empresa e que são apoiados por outros processos internos, resultando no produto ou serviço que é fornecido a clientes externos.

Neste artigo, os modelos de processos de negócio são representados pela versão 2.0 da BPMN (do inglês: *Business Process Model and Notation*). Devido à sua manutenção pelo OMG (do inglês: *Object Management Group*) a notação BPMN vem ganhado notoriedade [OMG 2011]. Existem indícios de sua popularidade ao analisar o estudo realizado em 2008, que evidenciou o fato da BPMN já ser utilizada em mais de trinta países [Recker 2008]. Seguindo essa tendência, um segundo estudo, mais recente, confirma ainda mais os indícios da popularidade do BPMN ao afirmar que a maioria (72%) das empresas entrevistas, em um total de 559, utilizam a BPMN para modelar seus processos de negócio [Harmon 2011].

Sobre a notação gráfica da BPMN, ela possui sete elementos básicos. Porém, dentro desses elementos básicos podem ser adicionados variações e informações para suportar novas necessidades, sem mudar drasticamente o visual básico do elemento e entendimento do modelo. A Tabela 1 ilustra esses elementos.

Tabela 1- Elementos Básicos da BPMN.

Nome	Símbolo	Descrição
Piscina		É a representação gráfica de processo em si, ou dependendo da situação de um participante externo ao modelo.
Raia		Uma piscina pode ser dividia em raias que representam os atores do processo.
Atividade		Seu objetivo principal é descrever o trabalho realizado. Pode ser atômica ou não atômica (composta)
Objeto de dados		Representa a descrição visual das informações que são utilizadas e/ou produzidas ao longo do processo.
Evento		Representa os fatos que acontecem ao longo do processo. São classificados como inicial, intermediário e final.
Fluxo		É usado para mostrar a ordem das atividades na execução do processo.
Gateway		É usado para representar a ocorrência de um fluxo diferente do caminho natural do processo. Ou seja, controla as divergências e convergências no fluxo de um processo.

Animação de Modelos

Embora a especificação de um modelo possibilite a modelagem de aspectos importantes de forma clara e não ambígua, é preciso certificar que ele atende aos seus requisitos. Existem várias técnicas que podem ser empregadas a partir de um modelo de modo a aumentar a sua confiança, que consiste em capturar as propriedades fundamentais [Silveira 2009]. Uma dessas técnicas é a animação direta dos modelos. A animação é uma técnica de apoio aos projetistas na definição dos processos de negócio de uma organização. Seu uso torna-se particularmente interessante, pois, ao trazer um processo de negócio à vida, a animação possibilita um valioso *feedback* na interação entre usuários e especialistas do mesmo e, ao contrário das técnicas manuais de produção de protótipos, uma relação explícita entre a especificação e a animação pode ser mantida, o que previne eventuais divergências de uma em comparação à outra. Além disso, os cenários da animação podem ser repetidos, o que é bastante útil, pois comumente após modificações na especificação do processo, deseja-se saber se todas as partes da mesma continuam se comportando como esperado [Gargantini 2005], [Kazmierczak 1998].

Por outro lado, tal como outras abordagens orientadas a teste, a animação se baseia em encontrar contraexemplos. Assim uma animação nunca pode provar que um modelo é consistente, correto ou completo, mas pode ser utilizada para produzir um nível satisfatório de confiança na qualidade da especificação [Silveira 2009]. Logo, a técnica de animação tenta achar um equilíbrio entre completude, de um lado, e velocidade e facilidade de uso de outro. Ela é caracterizada por uma ênfase na automação e na aceitação de um nível menor de garantia em troca de um *feedback* mais rápido, além de uma menor dependência de conhecimento profundo de provas matemáticas, empregadas em métodos mais rigorosos e formais de validação.

3. Método Proposto

O método proposto neste artigo visa identificar, por meio de animação, a não conformidade entre os modelos de processo de negócio em relação ao seu modelo conceitual de informação. A estratégia de animação do referido método consiste em dá vida aos objetos do modelo conceitual a partir das descrições das ações que outrora foram definidas no modelo de processo de negócio. Espera-se que com o resultado dessa animação fornecer informações importantes para a interação entre os profissionais de tecnologia com os especialistas do negócio, dessa forma diminuir as informações perdidas na comunicação entre esses profissionais. Para atingir tal objetivo, optou-se em considerar a filosofia dos métodos *Catalysis* [D'souza, 1998] e *Animare* [Silveira 2009], isto é, cada atividade (do inglês: *activity*) do

processo deve ser especificada de forma contratual, conforme apresentado mais adiante. A Figura 1 ilustra as etapas do presente método.

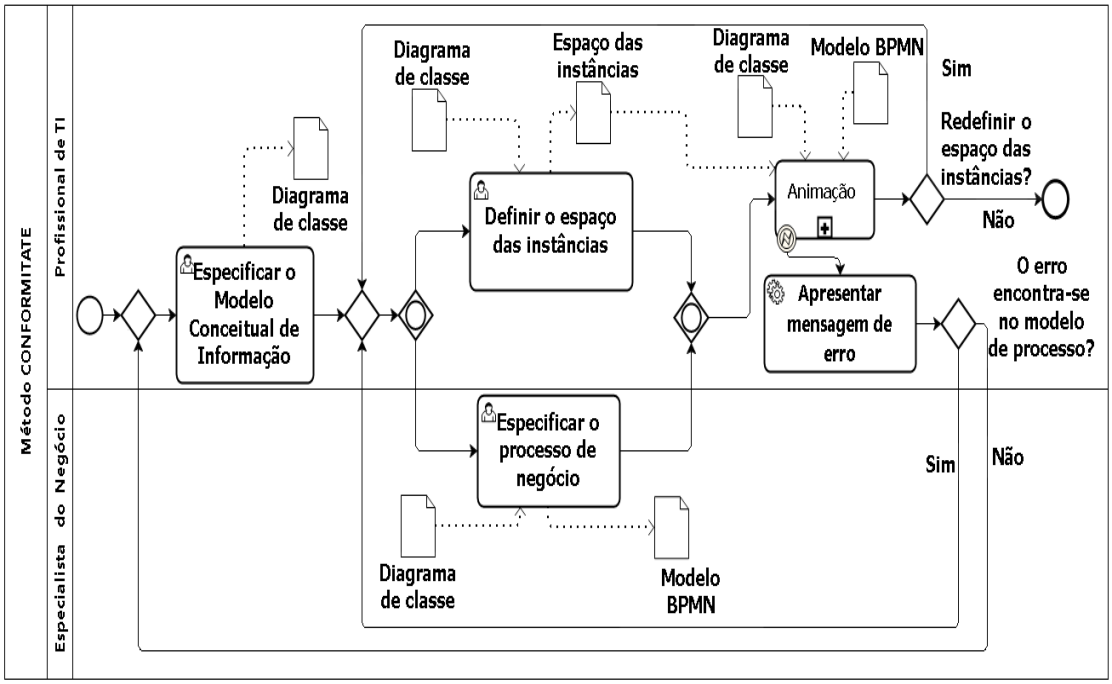


Figura 1 – As Etapas do Método.

Coforme pode ser observado na Figura 1 o método tem inicio com a etapa de *Especificar Modelo Conceitual de Informação*, que deve definir as restrições estruturais do domínio utilizando os elementos do diagrama de classe da UML. Em seguida, o diagrama de classe será utilizado como insumo para a realização das duas próximas etapas (*Definir o Espaço das Instâncias* e *Especificar o Processo de Negócio*), que podem ser realizadas em paralelo.

A etapa de *Definir o Espaço das Instâncias* consiste em definir o espaço das instâncias para as variáveis que compõem as entidades do domínio. O espaço das instâncias é utilizado para gerar o estado inicial dos objetos do domínio para uma animação. Cada animação é composta pela definição de um estado inicial dos objetos e pela realização de uma ação simples ou conjunto de ações. Em paralelo a definição do espaço de instâncias o especialista de negócio deve realizar a etapa de *Especificar o Processo de Negócio* utilizando a notação BPMN 2.0.

Dando continuidade, neste ponto o método está apto para iniciar a animação. Em outras palavras, a infraestrutura concebida para automatizar o método aqui apresentado deve começar a identificar a não conformidade entre os modelos formalizados nas etapas anteriores, apresentando *feedbacks* visuais que estão apresentados nas seções subsequentes.

Após a animação, o método permite seguir por dois caminhos. O primeiro caminho indica um sucesso que caracteriza a existência de indícios de conformidade entre as ações

descritas no processo de negócio e a restrições estruturais do negócio formalizadas no modelo conceitual para o cenário testado. Ou seja, a animação para o cenário testado foi realizada com sucesso. Porém, pelo fato de ser uma validação empírica não é possível afirmar que os modelos estão totalmente em conformidade. O que se pode afirmar é que para o cenário testado os modelos se apresentam em conformidades. Assim, é interessante que os envolvidos tentem achar um equilíbrio entre a completude das técnicas exaustivas com a velocidade e facilidade de uso do método aqui proposto realizando novos testes a partir de novos cenários. Então, cabe aqui ressaltar que o nível de confiança da conformidade entre os modelos analisados está diretamente relacionado com a qualidade dos cenários testados.

Todavia o segundo caminho indica o insucesso, ocasionado pela a ocorrência de um erro, gerado a partir da não conformidade entre uma ação definida no modelo de processo com a estrutura definida no modelo conceitual. Uma vez evidenciado a não conformidade, uma mensagem de erro deverá ser apresentada aos usuários (profissionais de TI e/ou especialistas do negócio) do método, que de posse dos estados dos objetos no momento do erro, poderá intervir no modelo conceitual de informação ou no modelo de processo de negócio. Após as interversões no(s) modelo(s) a validação poderá ocorrer novamente até se obter o sucesso esperado.

As próximas subseções ilustram o método proposto com o auxílio de um exemplo de locação de carros de uma organização denominada *EU-Rent*. As especificações envolvidas para os modelos aqui apresentados já são de domínio público e estão detalhadas em [OMG 2007].

Especificar Modelo Conceitual de Informação

A Figura 2 apresenta o diagrama de classe contendo os principais conceitos, representada como classes, para o contexto de locação de veículos. As classes *Reserva* e *Aluguel* são classes associativas, logo seus objeto são instanciados somente a partir das respectivas associações entre as instâncias de *Grupo* com *Cliente* e *Carro* com *Cliente*. Ainda existe uma associação simples indicando que cada instância de *Aluguel* está associada a uma instância de *CartaoCredido*. Por fim, existem as associações de agregação indicando que um *Grupo* agrega vários *Carros* e que *Cliente* agrega vários *CartaoCredido*.

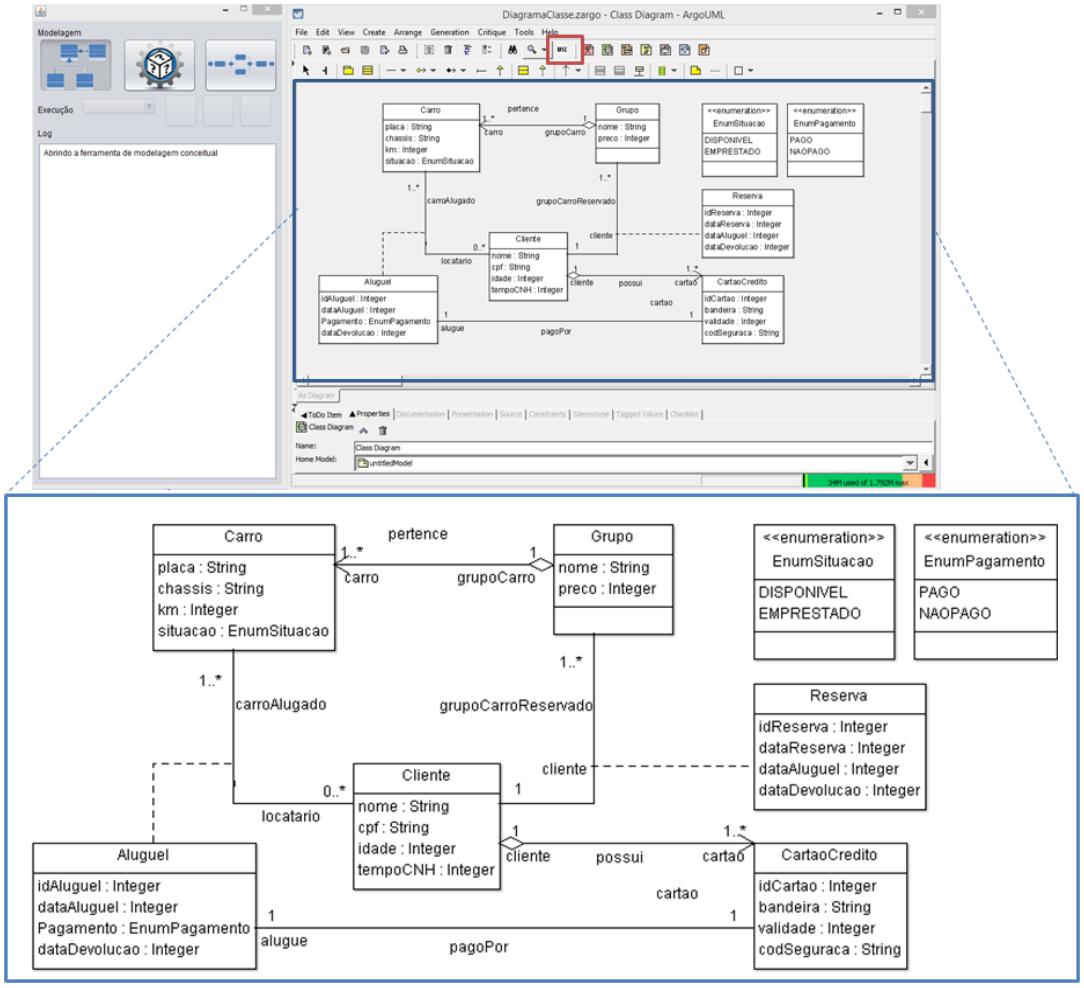


Figura 2 – Diagrama de classe usado no estudo de caso.

Após a especificação do modelo conceitual de informação, este é transformado para o formato proprietário da ferramenta USE (do inglês: *UML-based Specification Environment*). A ferramenta USE permite a verificação das restrições estabelecidas no modelo conceitual de informação. O USE é um ferramenta para especificação, analise e validação de sistemas de informação baseado na UML [OMG 2005]. Nela é possível representar um domínio por meio de uma descrição textual utilizando os conceitos do diagrama classes, que na ferramenta são tradados de forma dinâmica, por meio de comandos que possibilitam a manipulação das instâncias dos objetos definidos no diagrama. Essas manipulações podem ser acompanhadas visualmente pelos projetistas.

Em resumo, com a ferramenta USE é possível executar cenários, criando uma versão visual de instâncias do diagrama de classe para avaliar a especificação do mesmo. Assim, determinando se um dado modelo realmente representa o domínio real antes mesmo de sua implementação [Gogolla 2007]. A Figura 3 apresenta o digrama de classe para o domínio de locação de veículos e fragmentos de sua representação textual no referido formato.

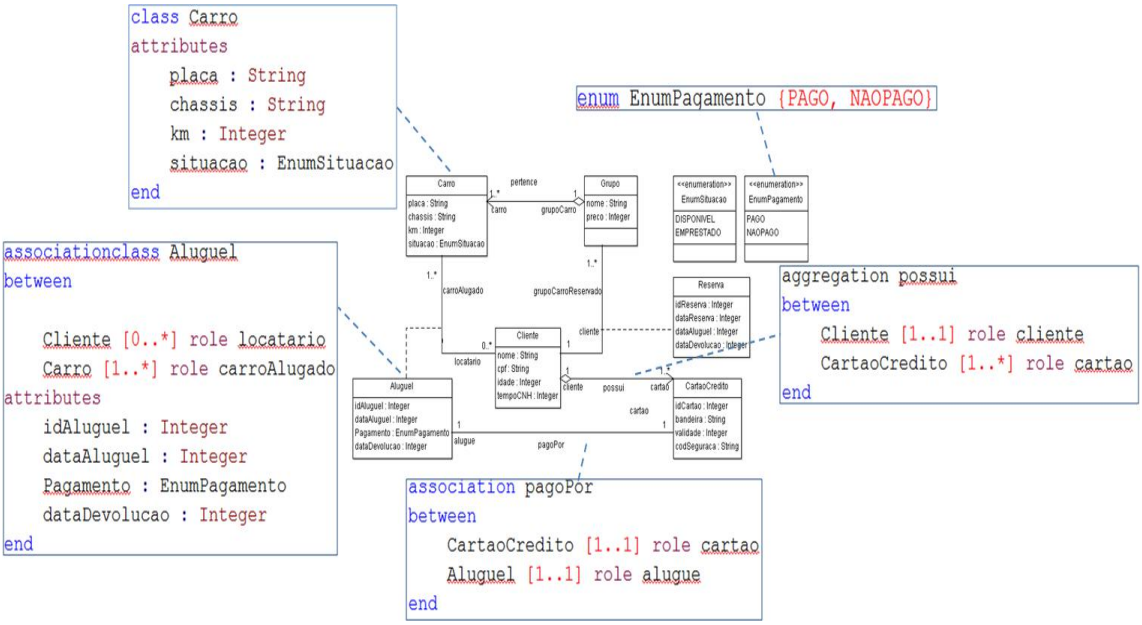


Figura 3 – Fragmentos da versão textual do diagrama de classe.

Após a transformação do diagrama de classe para o formato proprietário da ferramenta USE, esse é automaticamente passado para a ferramenta USE, que deverá instanciar os objetos seguindo as estrutura estabelecida pelo diagrama de classe. A Figura 4 apresenta a interface gráfica da ferramenta USE contendo o diagrama de classe que representa o domínio de locação de veículos.

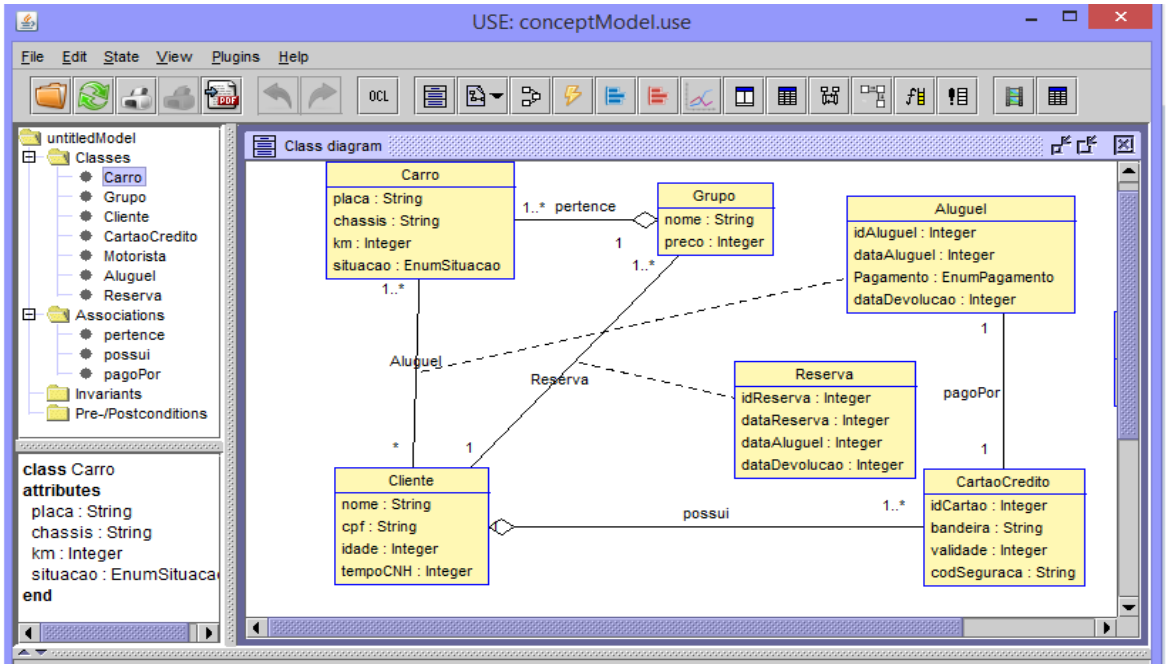
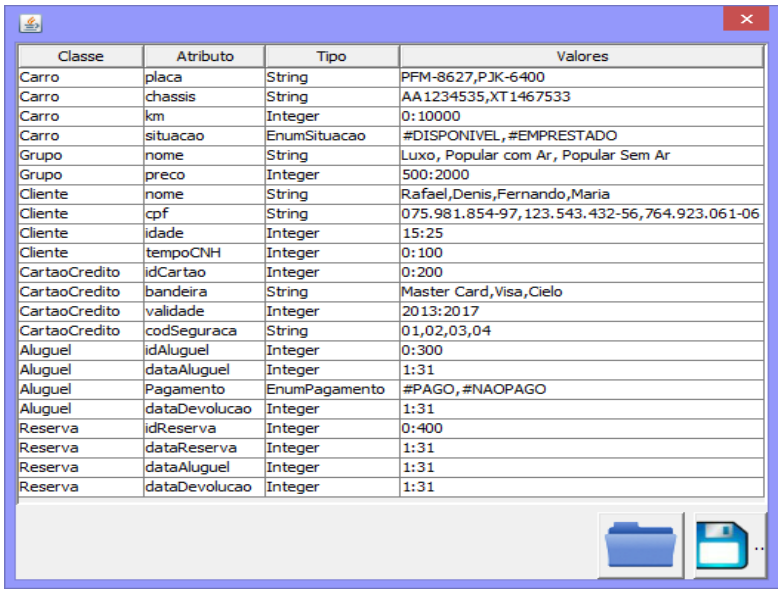


Figura 4 – Diagrama de Classe na Ferramenta USE.

Definir o espaço das instâncias

Após especificação do modelo conceitual de informação, o profissional de tecnologia deve definir o espaço das instâncias para as entidades do domínio. A Figura 5 apresenta a interface gráfica da ferramenta que ilustra visualmente as variáveis que compõem as entidades definidas no diagrama de classe. Além disso, é possível visualizar que para as variáveis do tipo numéricas, é permitido definir um intervalor de valores, utilizando a cardinalidade mínima e máxima com a seguinte sintaxe: *valor_mínimo:valor_máximo*. Assim sendo, um valor inicial para cada variáveis numéricas será escolhido aleatoriamente dentro do intervalo definido. Já para as variáveis do tipo cadeira de caracteres, o projetista deverá informar uma lista de cadeias de caracteres separada por vírgula (","). Assim, um valor (cadeia de caracteres) será também escolhido aleatoriamente dentro da lista definida.

Cada cenário é composto pela definição de um estado inicial das entidades do domínio (o espaço das instâncias) e pela execução de uma operação simples ou de uma sequência de operações definida(s) no processo de negócio.



Classe	Atributo	Tipo	Valores
Carro	placa	String	PFM-8627,PJK-6400
Carro	chassis	String	AA1234535,XT1467533
Carro	km	Integer	0:10000
Carro	situacao	EnumSituacao	#DISPONIVEL,#EMPRESTADO
Grupo	nome	String	Luxo, Popular com Ar, Popular Sem Ar
Grupo	preco	Integer	500:2000
Cliente	nome	String	Rafael,Denis,Fernando,Maria
Cliente	cpf	String	075.981.854-97,123.543.432-56,764.923.061-06
Cliente	idade	Integer	15:25
Cliente	tempoCNH	Integer	0:100
CartaoCredito	idCartao	Integer	0:200
CartaoCredito	bandeira	String	Master Card,Visa,Cielo
CartaoCredito	validade	Integer	2013:2017
CartaoCredito	codSeguraca	String	01,02,03,04
Aluguel	idAluguel	Integer	0:300
Aluguel	dataAluguel	Integer	1:31
Aluguel	Pagamento	EnumPagamento	#PAGO,#NAOPAGO
Aluguel	dataDevolucao	Integer	1:31
Reserva	idReserva	Integer	0:400
Reserva	dataReserva	Integer	1:31
Reserva	dataAluguel	Integer	1:31
Reserva	dataDevolucao	Integer	1:31

Figura 5 – Interface para definição do espaço de instâncias.

Especificar o Processo de Negócio

A etapa de especificação do processo de negócio formaliza, utilizando a notação BPMN 2.0, como as entidades do negócio são manipuladas no decorrer do processo. Foram definidas cinto ações básicas que podem ser utilizadas para definir como o processo manipula entidades do negócio: criar uma nova instância para uma entidade; remover uma instância, criar uma

ligação entre instâncias, remover uma ligação entre instâncias e modificar o estado de uma instância. Os *templates* para especificar essas ações estão apresentados da Tabela 2.

Tabela 2 - Templates utilizados para definir atividades.

Instruções	Descrição
ID_objeto:CRIAR (nome_classe);	Criar uma nova instância de uma entidade do negócio;
ID_objeto: DESTRUIR;	Remover uma instância;
ID_objeto_1,ID_objeto_2: ASSOCIAR(nome_associacão);	Criar um relacionamento entre instâncias;
ID_objeto_1, ID_objeto_2: DESASSOCIAR (nome_associacão);	Remover uma ligação entre instâncias;
ID_objeto: ATRIBUIR(nome_estado,valor);	Modificar o estado de uma instância.

A Figura 6 apresenta o modelo de processo, que descreve as ações necessárias para a locação de um veículo. O primeiro elemento do processo é evento de início, indicando onde o processo começa. A atividade *Obter dados do cliente* representa a ação de perguntar ao cliente suas informações. Após essa atividade, o *gateway* exclusivo verifica se o cliente está apto. Ou seja, verifica se ele possui idade superior ou igual a 21 anos e mais de 2 anos de habilitação para dirigir. Caso o cliente seja considerado inapto, o processo deve prosseguir sendo finalizado sem que o cliente alugue nenhum veículo. Porém, caso seja considerado apto o processo continua com as próximas atividades.

A atividade de *Selecionar Carro* representa a atividade de procurar um carro disponível. Em seguida encontra-se o *gateway* paralelo, semanticamente significa que as atividades de *Criar contrato aluguel* e *Cadastrar cartão* são realizadas de forma independente e simultânea. Por sua vez a atividade de *Criar contrato aluguel* consiste em criar um aluguel que vincule o cliente ao carro selecionado. A atividade de *Cadastra cartão* de credito consiste na ação de cadastrar um cartão de credito e associa-lo ao cliente. Após a convergência dos caminhos paralelos é realizada a atividade de *Pagar aluguel* que cria uma associação indicando o cartão de credito que pagou o aluguel. Por fim, o elemento de fim indicando a finalização do processo. As anotações de texto (destacadas em verde) conectadas as atividades apresentam as instruções que definem seus comportamentos ao manipular entidades do domínio em analise.

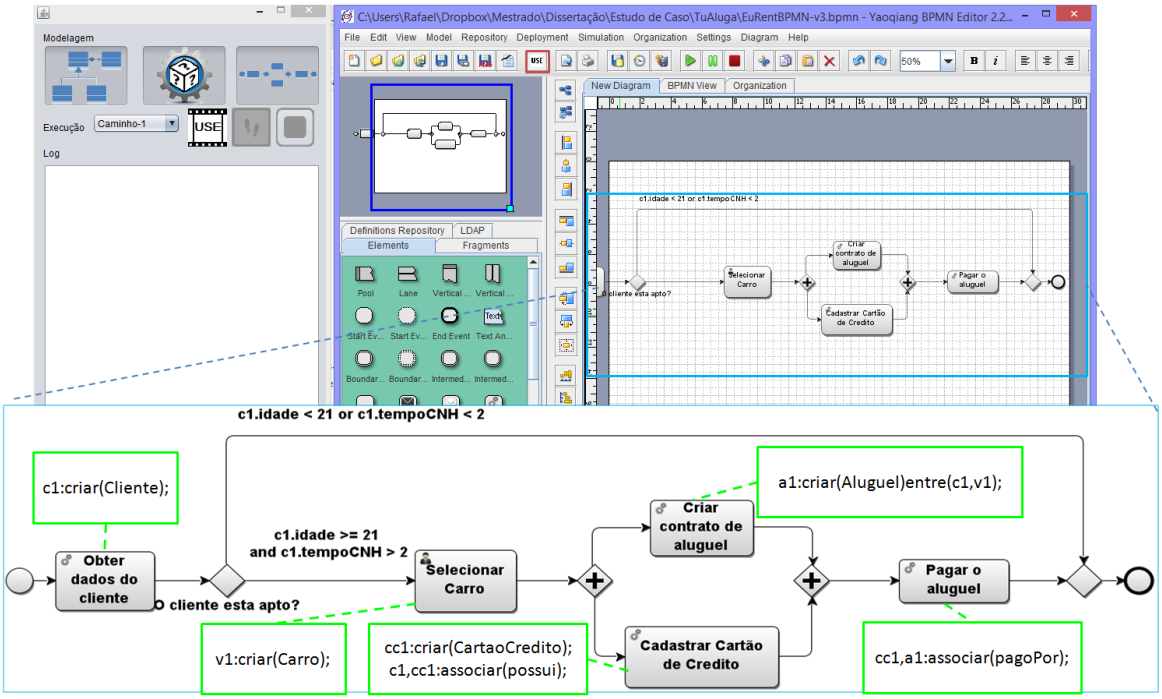


Figura 6 – Modelo BPMN usado no estudo de caso.

Gerando caminhos do Modelo de Processo de Negócio

Para realizar a análise de conformidade entre os modelos é necessário antes determinar todos os possíveis caminhos do processo que, por sua vez, deverão ser analisados individualmente. Para determinar os caminhos do modelo BPMN, o presente método adaptou o algoritmo de caminhos em grafos de fluxo orientados [Pires 2002], considerando os elementos de fluxo (atividades, eventos e *gateways*) como vértices e os fluxos de sequência como arestas, que o mesmo apresenta a ordem dos elementos de fluxo em um processo [OMG 2011].

A Figura 7 apresenta os possíveis caminhos existentes para o processo de locação. O que determina os possíveis caminhos são as condições booleanas descritas no *gateway* exclusivo, destacado pelo valor 1 em vermelho na Figura. Caso o cliente possua idade de no mínimo 21 anos e o tempo de habilitação superior a 2 anos, encontra-se apto a dirigir e, consequentemente, o processo pode seguir normalmente, caminho destacado em azul na figura 7. Caso contrário, o mesmo encontra-se inapto e o processo deve ser finalizado sem que nenhum carro seja alugado, caminho destacado em verde na Figura 7.

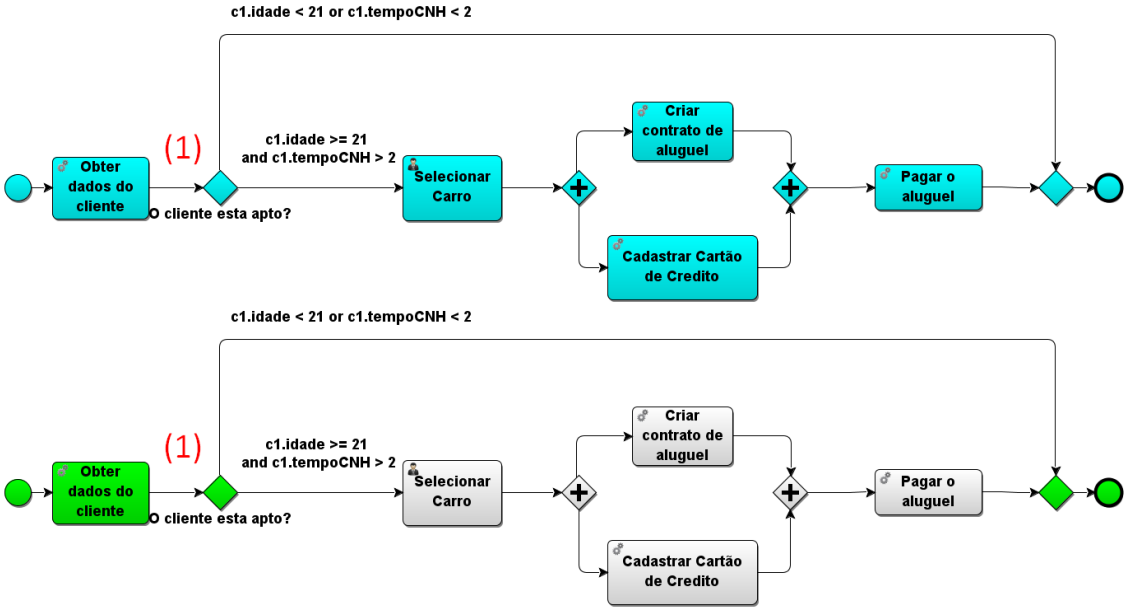


Figura 7 – Caminhos do processo BPMN usado no estudo de caso.

Selecionar um caminho do processo

O projetista deve selecionar um dos possíveis caminhos para analisar sua conformidade com o modelo conceitual de informação. A Figura 8 apresenta a interface para seleção do caminho a ser animado na ferramenta USE. Nessa interface é interessante destacar que é possível visualizar uma descrição dos elementos de fluxo contidos no caminho selecionado.

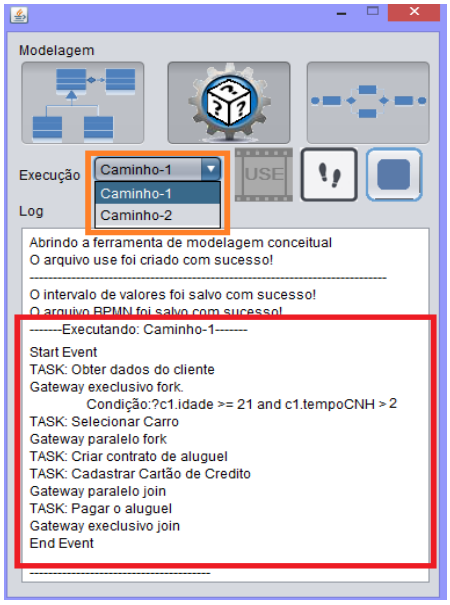


Figura 8 – Tela para seleção do caminho.

Coletar as ações

Uma vez selecionado o caminho, o mesmo será percorrido para coletar todas as ações descritas nas atividades. A ordem de coleta das tarefas segue o sentido do fluxo de sequência. A Figura 9 apresenta o caminho selecionado e o conjunto de ações coletadas das atividades.

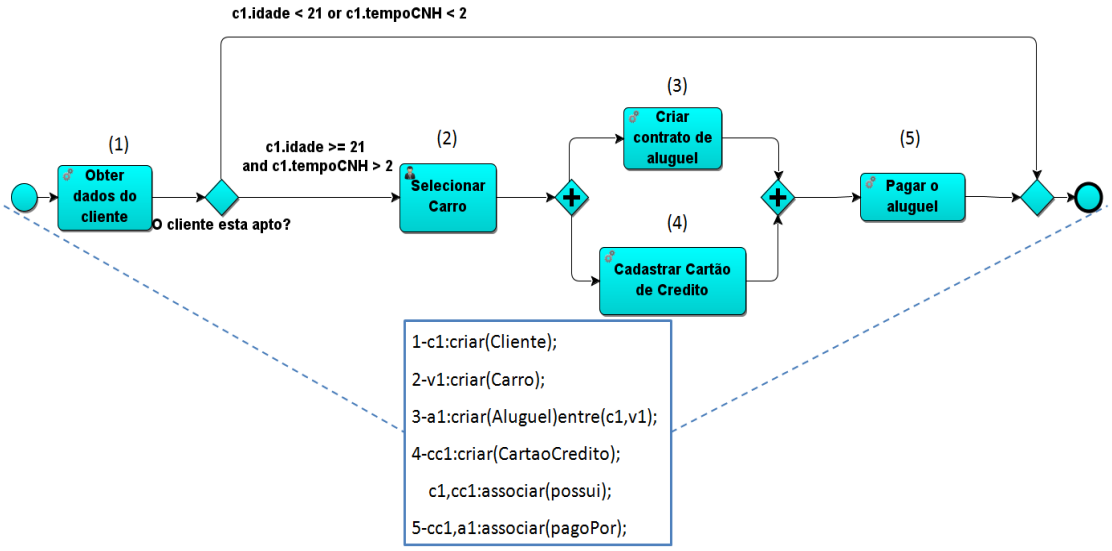


Figura 9 – Conjunto de ações coletadas do caminho selecionado.

A animação

A partir desta etapa a animação começa de fato. Logo, para cada atividade será executada sua ação ou conjunto de ações e para o *gateway* será testado sua condição, ambos na ferramenta USE. Caso a ação seja do tipo de criação de objeto os valores dos seus atributos serão gerados aleatoriamente levando em consideração o espaço das instâncias estabelecido pelo projetista.

A ferramenta USE verifica se a execução da ação não inflige nenhuma regra estrutural descrita no diagrama de classes. Caso a execução da ação, referentes a uma atividade, obtenha sucesso, significa que essa atividade encontra-se em conformidade com o modelo conceitual. Caso contrário, ou seja, a ferramenta USE retorne uma mensagem de erro, significa que a atividade possui uma falta de conformidade com modelo conceitual de informação.

A Figura 10 apresenta a animação completa do caminho selecionado. A animação é composta por várias imagens, cada uma contendo o(s) estado(s) do(s) objeto(s) logo após a execução com êxito de cada atividade.

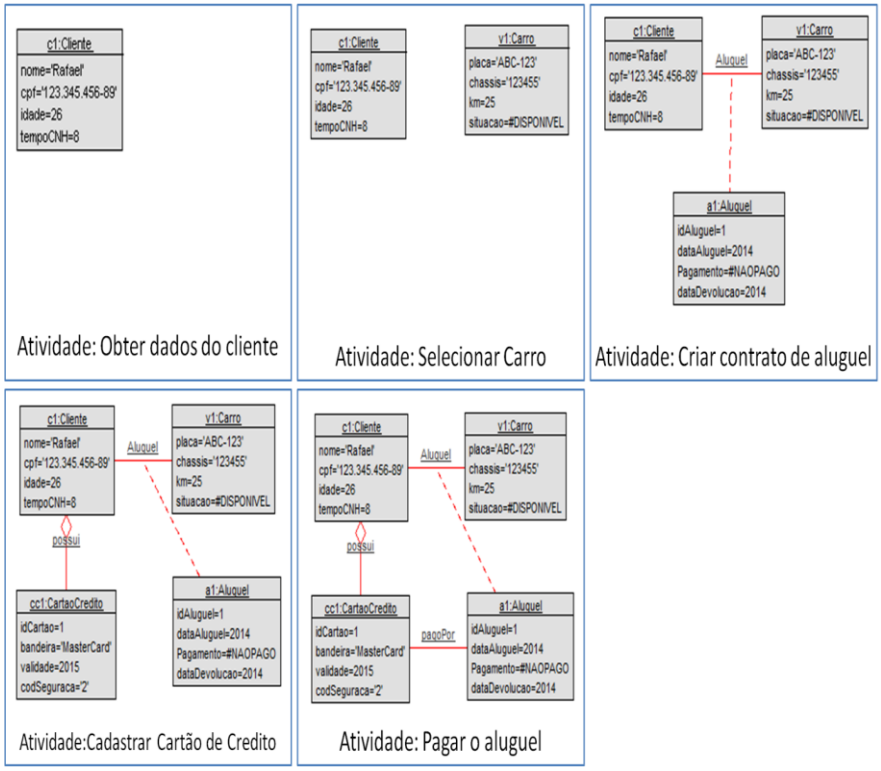


Figura 10 – Animação completa do caminho selecionado.

A Figura 11 apresenta uma animação que não obteve sucesso. Após a execução bem sucedida da atividade Obter dados do cliente foi instanciado um objeto do tipo cliente com idade igual a 18 anos. Consequentemente ao testar a condição *c1.idade >= 21 and c1.tempoCNH > 2*, o cliente irá se mostrar inapto para locar um veículo. Sendo assim, o método apresenta uma mensagem de não conformidade entre os modelos.

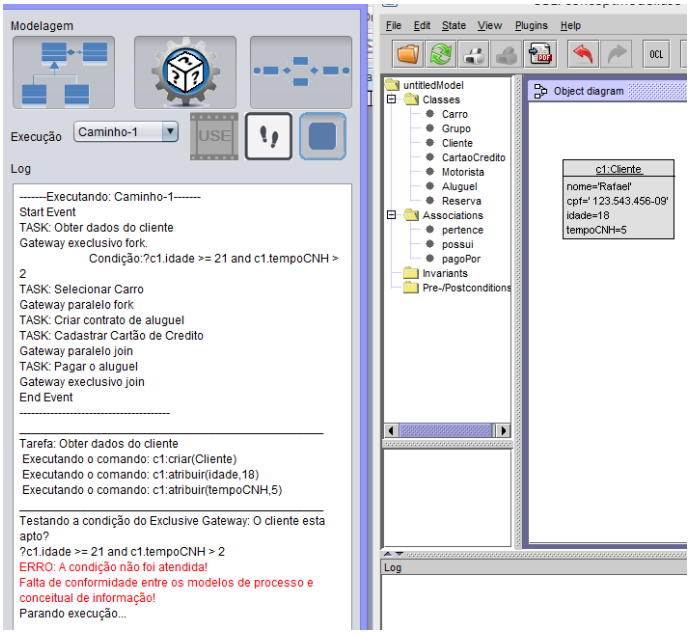


Figura 11 – Animação com erro na animação.

4. Considerações Finais

Uma vez que a modelagem dos processos é, cada vez mais, utilizada nas organizações, a necessidade de identificar sua conformidade em relação ao modelo conceitual de informação é um problema enfrentado diariamente nas organizações. Sendo assim, o objetivo principal deste artigo foi apresentar um método para identificar não conformidades entre o modelo de processos e modelo conceitual de informação utilizando técnicas de animação.

Sendo assim, o método aqui apresentado fornece uma ferramenta automatizada de suporte, que facilita essa validação dos processos em relação ao modelo conceitual, permitindo um rápido *feedback* para os envolvidos na especificação dos processos de negócio por meio da animação direta dos objetos referenciados nos processos com cenários de uso fornecidos tanto pelo analista de negócio como pelo cliente.

Embora o método de animação não garanta a corretude dos modelos, uma vez que não determina quando os modelos de processo de negócio e o modelo conceitual de informação estão em completa conformidade, ele oferece benefícios se realizados de forma apropriada. Entretanto, se a animação for realizada sem o devido planejamento, não será possível determinar se os cenários animados foram adequados para cumprir os objetivos da validação. Por fim, o método aqui apresentado encontra-se em fase final de aprimoramento, sendo empregado em casos reais por meio de um experimento que vai permitir avaliar, de forma mais criteriosa, os impactos relativos ao uso do método proposto em processos reais de desenvolvimentos de *software*. Considerando uma amostra estatisticamente relevante de participantes com diferentes níveis de experiência em modelagem.

Por fim, a materialização desse método abriu perspectivas de novas pesquisas para se agregar a esta, como por exemplo, a especificação de um verificador estrutural para os modelos BPMN que deve ser executado antes da animação. Essa verificação evitaria erros na animação que não são tratados atualmente.

5. Referências

Brooks, F., "No Silver Bullet- Essence and Accidents of Software Engineering", University of North Carolina at Chapel Hill, Computer Magazine. 1987.

D'souza, D., Wills, A. C., “Objects, Components and Frameworks with UML”, 1 ed., Addison Wesley, 1998.

Gargantini, A., Riccobene, E., “ViBBA: A Toolbox for Automatic Model Driven Animation”, 16th Conference of Simulation and Visualization – SIMVIS, Magdeburg – Germany, 2005.

Gogolla, M., Büttner, F., Richters, M., "USE: A UML-based specification environment for validating UML and OCL", Science of Computer Programming, v. 69, 2007.

Gonçalves, J., "As Empresas são Grandes Coleções de Processos", RAE - Revista de Administração de Empresas, pp. 6-19, 2000.

Harmon, P., Wolf, C., "Business Process Modeling Survey", BPTRENDS, p. 36. 2011.

Hazelrigg, G. A., "On the role and use of mathematical models in engineering design", Journal of Mechanical Design, v. 121, n. 3 , pp. 336-341, 1999.

Jalali, A., "Foundation of Aspect Oriented Business Process Management", Master's Thesis, Stockholm University, 2011.

Kazmierczak, E., Winikoff, M., Dart, P., “Verifying model oriented specifications through animation”, Pacific Software Engineering Conference, IEEE Computer Society Press, pp 254–261, 1998.

Larman, C., "Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and Iterative Development", 3.ed, New York: PrenticeHall, 2004.

Markenzon, L., Pires, O. V. S., "Percurso em Grafos, Relatório Técnico", NCE/UFRJ, 2002.

Mellor, J. S., Scott, K., UHL, A., WEISE, D., “MDA Distilled: Principles of Model-Driven Architecture”, Addison-Wesley, 2004.

Olivé, A., "Conceptual Modeling of Information Systems", Springer, 1ed, 2007.

OMG - Object Management Group, “Unified Modeling Language (UML) Superstructure Specification”, version 2.0, In: http://www.omg.org/cgi-bin/doc?formal/05-07-04, Acessado em maio de 2015, 2005.

OMG - Object Management Group, "Documents Associated With Semantics Of Business Vocabulary And Business Rules (SBVR)", Version 1.2 In: http://www.omg.org/spec/SBVR/1.2/ " , Acessado em maio de 2015, 2005.

OMG - Object Management Group. "Business Process Model And Notation (BPMN)", Version 2.0, In: "http://www.omg.org/spec/BPMN/2.0", Acessado em maio de 2015, 2005.

Recker, J. C., "BPMN Modeling – Who, Where, How and Why", BPTRENDS, pp 1–8, 2008.

Schimitz, E., Silveira, D., "Desenvolvimento de Software Orientado a Objetos", Editora BRASPORT, 2000.

Silveira, D. S., "ANIMARE: Um Método de validação dos Processos de Negócio Através da Animação", Tese de Doutorado, COPPE/UFRJ, 2009.

Standish Group; (2011). disponível em: http://blog.standishgroup.com/, Acessado em maio de 2015.

Uma reflexão sobre a evolução das funcionalidades de tecnologias e sistemas de informação na perspectiva da evolução natural

Jairo Simião Dornelas [(1)], João Álvaro Carvalho [(2)],Denis Silva da Silveira [(3)]

(1)UFPE/Capes, Brasil
jairo@ufpe.br
(2)DSI/UMINHO, Portugal
jac@dsi.uminho.pt
(3)UFPE, Brasil
dsilveira@ufpe.br

Resumo

O presente artigo resume um dos eixos de uma investigação levada a cabo com acadêmicos e profissionais da área de sistemas de informação, a fim de detectar alguma verossimilhança existente entre a evolução de funcionalidades em *software*, da família das tecnologias da informação e comunicação colaborativa, e a evolução tradicional das espécies focada na teoria da evolução seletiva de Charles Darwin. A expectativa era confrontar as linhas evolutivas de *software* com aquelas previstas pela seleção natural, a fim de conjecturar algum tipo de herança evolutiva entre tecnologias da informação e especular sobre a pertinência de usos de conceitos *darwinistas* como derivação e recombinação, por exemplo, no desenvolvimento de *software* para fins sócio-organizacionais. O relato aqui apresentado não considera as entrevistas de aprofundamento feitas com profissionais portugueses no ano de 2013, remetendo a análise apenas à etapa de lançamento de questionários.

Palavras chave: funcionalidades, tecnologia da informação colaborativa, evolução.

1. Introdução

Um conjunto de procedimentos governa as ações e atividades de uma organização e lhe permite sobrevier em um mercado altamente competitivo, no qual a atitude predadora concorrencial é algo legítimo e previsível [Godfrey; German, 2008]. Este conjunto de procedimentos é necessário para fazer que as ações administrativas, marcadamente perseguidoras de eficiência, busquem uma constante evolução, um constante aperfeiçoamento de suas diversas variantes e seus elementos funcionais [Greer; Ruhe, 2003]. Esses elementos funcionais normalmente são vislumbrados em artefatos tecnológicos, técnicas de metrificação, modelos e ferramentas de gestão e práticas de acomodação ao mercado, tal que a efetividade é uma busca constante e valor proeminente nos estudos de sobrevivência organizacional [Lehman; Ramil, 2002].

Neste sentido, a modernidade tem brindado as organizações com um forte aparato produtivo

proveniente dos *fronts* acadêmico e prático [Spek, 2010], ofertando soluções que se sucedem em gerações e vão de um plano em que reluzem desde as teorias e conceitos que as explicam, até os mecanismos e autômatos que as implementam.

Um dos sustentáculos de toda esta modernidade é sem dúvida a computação (informática, tecnologia da informação), bem mais difundida sob influências decorrentes da explosão das redes sociais baseadas em computadores e da produção de dados em todo o mundo. Estas tecnologias e suas variantes têm provocado mudanças em diversos preceitos [Daft, 2008]

No *front* pessoal, as ferramentas, dispositivos e possibilidades têm ampliado a capacidade humana de percepção, comunicação, colaboração e interação, soerguendo o homem eficiente e quiçá proativo. Esse homem em sua escala evolutiva e social descobriu e fez uso da inteligência para sobreviver, adaptando-se a dificuldades e erigindo padrões em regime de competitividade, tanto vivencialmente, trabalhando aspectos de sua competência, quanto biologicamente, tendo características mutantes trabalhadas de modo aleatório, as quais os tornaram mais fortes; hoje, graças a este ciclo, o ser humano conduz a transferência de inteligência para máquinas [Benavides; Segura; Ruiz-Cortez, 2010].

No *front* organizacional, as ferramentas, dispositivos e possibilidades foram associadas e acopladas a técnicas e modelos de gestão, a fim de tornar efetivas as ações gerenciais em um mercado cada vez mais acirrado. A organização em sua escala evolutiva e social arregimentou um conjunto de técnicas para tornar-se eficiente, atuando eficazmente para melhor controlar os recursos à disposição, tal que através da tecnologia, pôde ganhar onipresença e superar questões temporais [Benavides; Segura; Ruiz-Cortez, 2010]. Neste trâmite, preocupa-se evolutivamente com as operações, é certo, mas tende vivencialmente a valorar também seus talentos humanos, preocupando-se com o bem estar destes e evolutivamente tende a incorporar mutações que se destinam à molecularização das instâncias organizacionais, pela simbiose processo-homem-máquina.

No *front* social, as ferramentas, dispositivos e processos têm revolucionado a perspectiva social. De fato, em todos os campos de aplicação do conhecimento humano tem se presenciado uma reviravolta de comportamentos, procedimentos, atitudes e até emoções. As formas como a tecnologia da informação adentrou as ações humanas são desconcertantes e flutuam do aspecto de comunicação (lúdico) aos aspectos profissionais e de prestação de serviços [Benavides; Segura; Ruiz-Cortez, 2010]. Não raro é ver inversões de valores e reacomodações até de feitio educacional e de produção de saber. A contento, em termos evolutivos, a sociedade metamorfoseia-se com combinações e derivações alhures, fazendo surgir novos contratos filosóficos e novos paradigmas de existência, em uma seleção típica e construída em saltos cada vez mais previsíveis em origem, mas enigmáticos em resultados.

Captar e sublimar a integração destes três *fronts* é tarefa cada vez mais robusta, que requer uma interdisciplinaridade de saberes e uma disposição para buscar para além dos compêndios das ciências exatas, de onde provém a tecnologia da informação, acepções e compreensões que expliquem (ou enfoquem sob outro prisma) a evolução daquela última. No presente relato, opta-se, como se verá oportunamente, por recorrer à teoria Darwinista do segmento de ciências biológicas para averiguar este mote.

2. Contexto da pesquisa

As indicações suscitadas introdutoriamente sobre os elementos componentes da tecnologia da informação são, por definição, resultantes de conversões de modelos lógicos em funcionalidades [Fuks; Gerosa, 2003]. De fato, para se concretizar a transferência de uma função cognitiva para um procedimento passível de mecanização, executa-se, incontinenti, uma decomposição lógica, baseada em uma cadeia de codificações instituída, e sempre refinada, que institucionaliza a ação concebida em nível cognitivo, para que um autômato a execute [Zuben, 2000].

Tal decomposição é responsável também pela transmissão de estímulos entre as diversas partes do código, tal que haja uma cadeia de execução e transformação constante. De fato, em se tratando de *software*, esta cadeia decorre de instruções, que em última análise transmitem uma ordem lógica à máquina para que a ação seja executada. Trata-se, pois, de transição através de cadeia codificada de estado a estado, conduzindo, em somatório, à realização de uma tarefa.

Para isso, mirando as funcionalidades, tem-se que instruções de diversos níveis concorrem para execução de rotinas, as quais, conjugadamente, executam um procedimento, o qual por sua vez, é parte de uma lógica maior que representa, em última análise, a função a ser feita. No que concerne ao refinamento destas funções, criam-se aplicativos e as sucessivas modificações e aperfeiçoamentos nesses, funcionam como uma espiral que denota a evolução do *software* para uma versão melhorada do mesmo *software* ou, se o grau de mudança for severo, para um sucedâneo que incorpora elementos que se mostraram vantajosos e descarta, ou relega a *background*, as *features* mais desvantajosas [Mendes; Pereira, 2013].

O suporte para esta trajetória evolutiva pode ser delineado para variantes de uma mesma tecnologia por exemplo. Nota-se, nessa linha de raciocínio, variantes tecnológicas em tornos de grupos e suporte computacional em espécies como *group decision support systems* (GDSS), *computer support cooperative work* (CSCW), *groupware* (GW), *group support system* (GSS), entre outras, fizeram aflorar um novo jeito de fazer computação, uma nova ordem para o uso da TI, enfim um redimensionamento global do que se pensava como uso de ferramentas para o desempenho de funções em grupo com suporte computacional.

Pela ótica delineada na sucessão dos acrônimos de *software* listada antes, o que se mirava era um clássico avanço migratório das funções de comunicação, coordenação e colaboração, sustentáculos do clássico modelo de Ellis, Gibbs e Rein [1991], no rumo de uma construção colaborativa, que havia sido esvaziada anteriormente pela explosão de usos dos computadores pessoais [Bidgoli, 1996].

Ora, se se olha o ambiente indutor da seleção de funções, mirar-se-á no fortalecimento, explosão e massificação das redes de computadores e das funções colaborartivas, como vetores desta pretensa migração e elemento de transmutação, tanto nas linhagens de *hardware* como de *software* e procedimentos rumo à computação colaborativa [Kock; Nosek, 2005]. Todavia, não obstante este robusto perfil, os aspectos chaves da tecnologia experimentaram distintos graus de adaptação ao novo contexto.

O aspecto comunicação destacadamente disparou em uma vertente claramente necessária e vencedora, ao passo que o aspecto coordenação, grosso modo, estancou em pequenos saltos a reboque dos outros elementos, enquanto que a colaboração passou por um processo de recomposição, bifurcando-se em cooperação, uma linhagem menor, ancorada em processos abertos e de multiagentes [Lee; Buckland; Shepherdson, 2003], e evoluindo como *constructo*, uma linhagem vencedora, tornando-se mais factível e visível em todas as instâncias de referência (indivíduo, organização e sociedade).

Assim, a colaboração, dimensão original, resultante da especiação por seleção evolutiva, assentou-se (e assenta-se) muito mais nos aspectos comunicação e cooperação e se vai estabelecendo nas chamadas tecnologias de informação e comunicação colaborativas (TICC) [Kock; Nosek, 2005]. Destarte, pode-se pensar que evolutivamente os ramos da classe existente foram se reacomodando tal que um novo híbrido surgiu em forma de derivação da tecnologia anterior (GDSS), com elementos dominantes na origem. Ao ver do que se quer observar e tentar alinhavar neste estudo, esse enredo evolutivo sugere as redes sociais, como sucedâneos dos GDSS, cumprindo *script* de seleção, mesmo que de base artificial, induzida, que pode ser, hipoteticamente, mapeado nos aspectos chave da evolução natural darwinista.

Desta forma, se tentará observar através da manutenção, transformação e inovação de funcionalidades existentes em artefatos anteriores e atuais, tomando por base os GDSS, a eventual migração e a nova oferta dessas funcionalidades em redes sociais, buscando identificar itens que sejam correspondentes, complementares, suplementares, transpostos ou extintos.

Para tal firmar-se-á o olhar em quatro grandes suportes, quais sejam: a computação evolucionária, como estado da arte do desenvolvimento de pacotes e sistemas; elementos de processo ancorados em mapeamentos técnicos de projeção de etapas em funções; as diretrizes e

direcionamentos de uso das ferramentas da TI nos *fronts* de atividade humana e, por fim, os aspectos mais introdutórios da teoria darwinista de evolução das espécies.

Como cenário para este roteiro de conceitos, no qual se buscou alguma chancela empírica à idéia lançada, o estudo alinhavou, mesmo que sem maior intensidade, dois momentos em campo: um coletando opiniões de docentes da área de TI, conhecedores de SAD-G, e prováveis usuários das redes sociais, e outro entrevistando experientes desenvolvedores de *software* de viés corporativo, a fim de resgatar elementos para subsidiar as conjecturas traçadas. O resultado global evidenciou alguma propriedade da ideia lançada e a necessidade de aprofundamento na exploração do tema, mas consignou uma coerente possibilidade de enxergar sob determinada lente, a evolução de *software* como um elemento aderente às ideias de seleção natural propostas por Darwin.

3. Delimitação Conceitual

As organizações cada vez mais buscam solidificar suas posições sociais e de contato com a comunidade, através de técnicas de gestão que apontam para soluções diretivas colegiadas, decisões compartilhadas e constante gestão de mudanças [Daft, 2008], as quais, por sua vez, usufruem bastante da tecnologia da informação. Tal formato reporta-se como um retorno às origens gregárias do ser humano e contrapõe-se à explosão da individualidade provocada pelo surto dos computadores pessoais, conforme anteviram Ellis, Gibbs e Rein [1991].

De fato, de portais corporativos a controle de acesso coletivo; de computação para usuário final a metodologias participativas de concepção de sistemas; de *chats* e mensagens instantâneas em *blogs e twitters* com vertiginosa velocidade de disseminação, acrescendo as ontologias e redes de significado semântico, endereçáveis tais quais IP´s [Fonseca et al., 2008], assiste-se a uma imensa metamorfose social que tem como ponto de chegada, ao menos em delineamento, o reposicionamento da noção de coletividade como instância gestora da comunidade.

Os primeiros segmentos a alinharem estes conceitos com o *slogan* de panacéia para a computação empresarial foram a área de GDSS, e sua posterior materialização em *groupware*, pelo lado da ciência administrativa, e a computação cooperativa (*computer supported cooperative work* - CSCW) pelo lado da ciência da computação [Bidgoli, 1996]. Todavia, como nenhuma das duas vertentes tornou-se hegemônica, o que se seguiu foi uma querela conceitual entre correntes que culminou com um incremento da confusão e da coexistência dos conceitos.

Ellis, Gibs e Rein [1991] propuseram então que a computação viria a ser colaborativa, trafegaria pelo perfil multidisciplinar e envolveria tecnologias de nível organizacional para integrar processamento de informação e atividades de comunicação para interação de pessoas. Nessa

perspectiva, os GDSS foram pioneiros em fornecer funcionalidades baseadas em computador para explorar problemas não estruturados num contexto de grupo.

A partir do legado acima, Kock e Nosek [2005, p. 1] propõem que a "colaboração entre indivíduos engajados numa tarefa comum usando tecnologias eletrônicas", transforme-se em *e-collaboration*, numa clara alusão ao suporte à colaboração entre indivíduos; ou seja, há um filo comum e espécies com contornos semelhantes derivados da base, mantendo um tronco comum e especificidades, numa referência notória à Biologia [Lewis, 1996].

A massificação da Internet, com clara integração de ambientes distribuídos baseados na *web*, acionou o gatilho para declínio dos GDSS e para a assunção de parte de suas funcionalidades por ferramentas que facilitavam esforços colaborativos e comunicação paralela [Limayem; Banerjee; Ma, 2006], conforme modelagem conhecida como modelo 3C [Fuks et al, 2003].

Tal disseminação maximizou em tecnologias de informação ditas inovadoras, o aspecto comunicação, ente forte e de excepcional apelo social, em detrimento de aspectos de coordenação e de cooperação no apoio ao processo decisório. De fato, Costa, Melo e Godinho [2003, p. 7] evidenciaram a necessidade de estudar aspectos cooperativos suportado por computador, à luz das funcionalidades antes presentes nos GDSS e agora distribuídos via rede, como forma de apreender uma "potencial transferência de aspectos mais fortes para uma outra tecnologia, deixando remanescer aquelas ainda necessárias, mas sem maior respaldo individual".

Assim, a reboque dos avanços em telecomunicações e de sua associação com novas tecnologias, especialmente as tecnologias colaborativas e sociais, há de se perguntar em que formatos os demais aspectos do escopo de uma tecnologia pensada para sistemas, implementada com o auxílio de redes, através de funcionalidades padrão, foram ou estão sendo sucedidos por aplicações disponibilizadas em perfil coletivo na *web* 2.0, antenadas com as redes sociais.

Os aspectos principais destas novas grandezas de uso da TI no *front* social, que se querem como vetores da transferência, podem ser vistos tanto como elementos de variação quanto de recombinação de funcionalidades, sendo, por opção conceitual dessa pesquisa, buscados como grandezas de cooperação, comunicação, coordenação e interação, brevemente comentadas antes.

4. O suporte conceitual à luz da ciência natural

Analisando a história das ciências detecta-se alguma semelhança entre o que se ideou estar acontecendo com os GDSS rumo A TICC e às redes sociais e a construção explicativa da teoria da seleção natural proposta por Darwin [2003] no século XIX.

De fato, tendo em conta que a ciência da computação tem buscado com certa veemência

explicações no campo da biologia, em segmentos como algoritmos genéticos e redes neurais, [Sasaki; Tokoro, 2000], soa razoável buscar elementos de perfil darwinista para expor a evolução de uma tecnologia.

Uma noção prévia e importante é a hereditariedade, já explorada nos sistemas de informações (SI) tradicionais, nas características de sistema legado e herança de classes, por exemplo, noções que indicam que há, no caso dos SI, algo equiparável a uma especiação entre variedades e subvariedades ligeiramente diferentes entre si [Darwin, 2003]. As noções aqui apontadas para os SI, se tomadas em viés ordenado, permitem observar que a transmissão de características, e porque não de funções, é regra, também entre os artefatos. Assim, o primeiro passo é buscar os conceitos de variação e variedade [Darwin, 2003] e entendê-los. Em fazendo isto, perceber-se-á que é possível conjecturar uma linha de hereditariedade [Darwin, 2003] entre GDSS e TICC, quando se pensa em nível de transmissão de características [Johnson; Lam, 2010].

De fato, se se contempla a acumulação de efeitos durante os diversos refinamentos de um GDSS, assevera-se que diferentes versões do *software* foram criando espaço para exploração de funcionalidades do escopo da comunicação. Assim, isto fez que este aspecto das tecnologias para suporte a grupos se desenvolvesse mais rapidamente, por indução do contexto, que as funcionalidades de colaboração e coordenação, o que configura uma explicação plausível para a emancipação da variedade do escopo da variação e o surgimento de um novo ramo da espécie, espelhando-se na conceituação derivada da seleção natural [Darwin, 2003].

A elucubração acima conjecturada é, de fato, cogitada para as espécies por Darwin [2003, p. 50] que enuncia que "a passagem de um grau de diferença para outro pode resultar da ação da natureza ou das diferentes condições de contexto".

Raciocínio similar pode ser estabelecido em relação à evolução por estabelecimento de necessidades de adaptação. A frase "todavia quando se trata de adaptação, pode-se atribuir às ações acumuladoras da seleção natural e aos efeitos do aumento do uso e não uso" [Darwin, 2003, p. 50], encaixa-se sob medida em qualquer análise de requisitos de uma função sistêmica. Ainda a título de ilustração, tome-se, por exemplo, o *bulletin board system* (BBS) presente nas versões preliminares dos GDSS como um elemento importante para habilitar comunicação entre utilizadores remotos. Tal elemento sucumbiu em refinamento, para a *postagem* fartamente permitida hoje em dia nos mecanismos da *web 2.0,* mas funcionalmente guarda o traço de propiciar que usuários se comuniquem sem intervenções de terceiros, direta e imediatamente.

Reforçando item essencial nessa odisseia tentativa de transposição, constata-se que em termos de história natural, a seleção natural institui-se "pela conservação das diferenças individuais favoráveis e eliminação das variações noviças ou insignificantes" [Darwin, 2003, p. 75].

Contextualizando em termos de TI e de sua velocidade de mutação, "estas transformações lentas e progressivas escapam-nos até que no decorrer das idades, a mão do tempo as tenha marcado

com o seu sinete" [Darwin, 2003, p. 78]. Verifica-se assim que as gerações de *software* [Hess, 2011] são, exatamente, em visão ajustada, essa progressividade (nem tão lenta) alentada.

No entanto, para que modificações sucessivas importantes se produzam numa espécie, é necessário que uma variedade uma vez formada apresente de novo, no tempo, diferenças individuais que se conservem e se renovem. Isto, mais uma vez, é o que aparentemente ocorreu quando os GDSS tornaram-se *groupware* e passaram a sistemas e tecnologias colaborativas. Tal enunciado reforça o princípio seletivo de que "se uma espécie não se modifica nem se aperfeiçoa depressa como as concorrentes deve ser exterminada" [Darwin, 2003, p. 107].

Assim corrobora-se, ao menos em nível especulativo e semântico, que tenha acontecido algo similar com as variantes básicas do GDSS para a assunção da TICC. Com isso, aspectos frágeis de coordenação foram se esvaindo e caindo em desuso na evolução das novas funcionalidades e se fixando apenas nos GDSS; ao passo que aspectos e variações diferentes e significativas de comunicação e cooperação (em menor escala) foram se agigantando e formando um novo ramo de crescimento, restando à variedade perdedora, a manutenção do espaço obtido e por falta de variação gênica [Darwin, 2003], a (previsível) condenação à extinção.

Evidente que diversas outras analogias podem ser estipuladas, com o devido cuidado de análise, o que aparentemente justificaria persistir na busca que se quer estabelecer, esmerando-se para não incorrer nas imperfeições e exageros citadas por Nehaniv *et al*. [2006].

Enfim a síntese operacional que se quer implementar para transpor a idéia do *front* teórico para o campo aplicado está retratada na figura 1, a qual contém os diversos elementos constantes da pesquisa, muitos dos quais por questões de espaço não serão aqui comentados. A figura aponta GSS ao invés de GDSS, por questões de semântica do lançamento do inquérito.

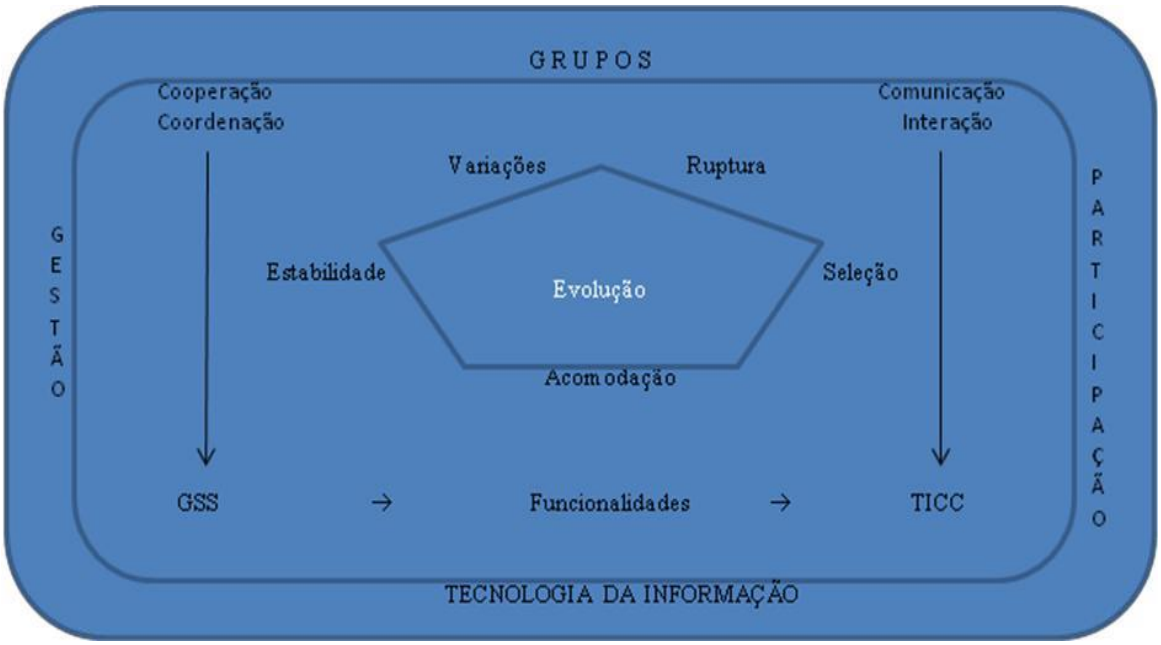


Figura 1 – Síntese da trama conceitual idealizada para a pesquisa.

5. Procedimento metodológico

A pesquisa possuiu um caráter qualitativo, na medida em que foram resgatadas impressões, percepções e sentimentos sobre fenômenos em curso, diretamente das pessoas, mas também valeu-se de uma sondagem modelada em patamares quantitativos, para capturar perfis de uso e compreensão de especialistas e profissionais sobre as tecnologias empregadas.

Desta forma, o esforço de campo concentrou-se em duas frentes: aplicação de questionários padronizados para coleta de informações junto a docentes, educadores da área, em uma primeira etapa, e em seguida efetuaram-se entrevistas com profissionais desenvolvedores, nas quais se pôde estruturar questões de aprofundamento sobre temas emergentes da 1ª etapa, ao lado de temas já estabelecidos que nela foram trabalhados. A etapa das entrevistas não será tratada.

As etapas conjugadas também possibilitaram apurar emergência, decadência e readequação de funcionalidades, o que fez possível a montagem de uma projeção evolutiva, inspirada declaradamente na teoria da evolução seletiva de Darwin, como argumentado previamente.

A partir destes itens foi possível instituir o desenho metodológico, exibido na figura 2, onde, como se pode depreender, houve um momento inicial de validação das escolhas feitas, a partir do mergulho literário, seja no nível das funcionalidades usuais em GDSS, seja nas características das tecnologias colaborativas que dão cobertura hoje em dia às atividades sócio-organizacionais de suporte às pessoas.

O público escrutinado na primeira etapa foi formado por docentes, *experts* acadêmicos, tanto da área de sistemas de informação, quanto da área de informática (ciência da computação), o que, em tese, chancelou o mapeamento e as escolhas das funcionalidades efetuadas. Tendo tudo isto em mente, de acordo com Monteiro [1998] é possível expor um desenho de pesquisa que contém os principais elementos para a realização do processo de investigação. Na montagem abaixo, nota-se que apenas a fase I está descrita e as conclsões são parciais.

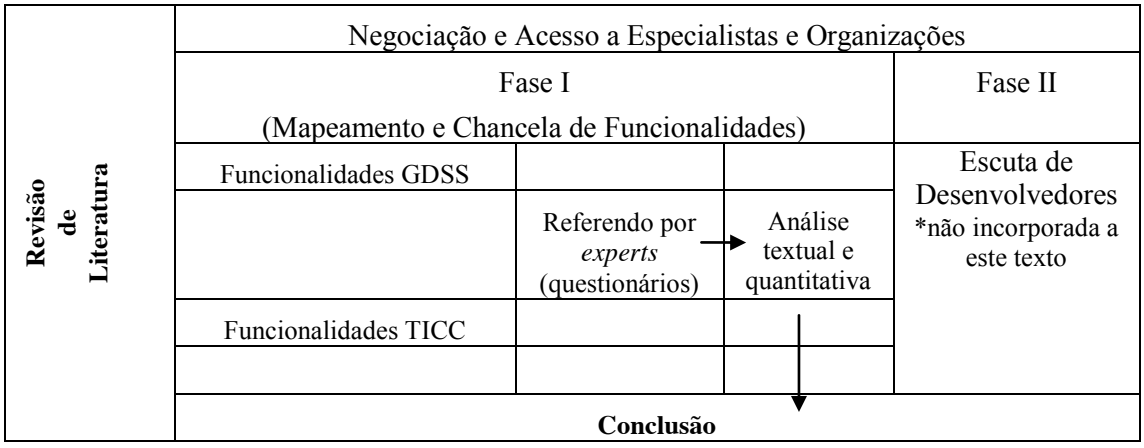


Figura 2 – Desenho da pesquisa.

Para a análise foram montadas matrizes de similaridade e quadros de consistência, buscando enquadrar as funcionalidades identificadas vis-à-vis os conceitos da teoria *darwinista*. A partir do mapeamento realizado foi possível estipular a existência da linha evolutiva que se queria mostrar em certas variantes; já em outras houve indefinição de direção e sentido e em um terceiro grupo houve imprecisão, como será mostrado.

6. Resultados alcançados (relacionados apenas à 1ª fase)

A análise global foi empreendida sobre o conjunto de respostas cedidas pelos *experts* aos questionários entabulados. A partir de suas percepções, cogitaram-se ideias sobre a transferência de funcionalidades, fundadas no modelo 3C de colaboração, tentando mitigar conceitos da teoria Darwinista aplicáveis às associações efetuadas, a reboque da computação evolucionária e de outras ideias que de alguma forma já graçam na literatura de apoio [Marks, 2002; Mehta, 2012].

6.1 Análise dos questionários

Foram resgatadas impressões de 18 docentes das Escolas de Engenharia (Departamento de Sistemas de Informação e Informática) da Universidade do Minho – Portugal, que se declararam conhecedores da temática de sistemas colaborativos e de apoio à decisão.

A questão inicial suscitava que o respondente escolhesse sua definição de funcionalidade dentre opções retiradas da literatura de apoio. Expressiva maioria denotou funcionalidade como a característica efetiva do que o *software* faz para atender uma atividade. Nesta compleição houve destaque para a adequação da rotina informatizada a uma necessidade organizacional, que não é estática e evolui. Por indução, características mutantes das necessidades a serem atendidas, compelem a arranjos de refinamento nas funcionalidades fornecidas.

A fim de evidenciar as impressões que se tinham sobre as duas tecnologias confrontadas, foi pedido ao respondente que enquadrasse de forma livre um conceito para cada uma delas, apenas suscitando a sinalização de que ambas deveriam ter perfil colaborativo. Tal perfil encaminhava o respondente ao leque de tecnologias atualmente consagradas para arranjos coletivos.

GDSS é uma tecnologia colaborativa, pois...	
Inclui ferramentas de apoio à decisão colaborativa	Viés Tecnológico e de Decisão
Induz à convergência de conhecimento	Viés de Conhecimento
Facilita a cooperação e ajuda a decidir em grupo	Viés de Cooperação e de Decisão
Ajuda a cumprir melhor os objetivos	Viés Organizacional
Avalia e pondera em decisão conjunta, coletiva	Viés Processual e de Decisão
Auxilia a gerir uma equipe de forma coordenada	Viés de Coordenação

Quadro 1 – Acepções sobre os GDSS como tecnologia colaborativa.

Ao quadro foi ancorada a percepção dos pesquisadores sobre qual aspecto do contexto literário era mais proeminente na visão fornecida. Cerca de 70% das respostas admitiram ser GDSS uma tecnologia eminentemente colaborativa. A expressiva maioria que considerou tal tecnologia com este perfil, justificou o enquadramento pela "possibilidade de fazer actos colaborativos" [Professor #1]. Algumas definições obtidas merecem destaque por indicarem certa categorização de *constructos* usuais na literatura de apoio ao tema. Ei-las em pequenos comentários.

Observa-se certa sintonia entre os aspectos processuais e organizacionais voltados à tomada de decisão com suporte de tecnologia e incentivo à coordenação e à cooperação (e por maioria na visão conceitual, também à colaboração).

Os respondentes indicaram também que os GDSS têm perdido espaço no arsenal de TI´s em uso nas organizações, o que é confirmado também pela literatura [Barkhi; Jacob; Pirkul, 2004], embora cerca de 42% das opiniões ainda atestassem a existência de um espaço para uso.

Olhados à luz do modelo 3C de colaboração, mais da metade dos ouvidos admitem como valiosos os ramos da coordenação (67%) e da cooperação (58%), indicando ainda a prevalência destes aspectos de forma nítida nos GDSS. Entretanto, já não se pode ter a mesma ênfase quando se olha o aspecto comunicação como rótulo destes sistemas, visto que metade dos ouvidos não o identificou como marcantemente presente (ou seja, não se garante que este aspecto defina hoje em dia o que é um GDSS). Nesta perspectiva, ratifica-se a ausência do viés comunicação tal qual detectado fora no exame literário prévio.

Ato contínuo procurou-se averiguar qual a conotação que se tinha na massa respondente acerca da propriedade de se taxar uma rede social como uma tecnologia colaborativa. Analisando as respostas obtidas, percebeu-se que não houve confirmação concreta desta suposição, visto que apenas 25% foram categóricos neste enquadramento. Equitativos 25% tiveram alguma dúvida no enquadramento, arguindo certa *restritividade* na consecução de objetivos organizacionais ou talvez em especificidades "sua utilização colaborativa seja menos apropriada para a gestão" (Professor #3).

Marcante, então, é que aspectos relativos à comunicação, já prenunciados como o elo de dissociação, reluzem nas opiniões colhidas, indicando, de fato, ser a comunicação o vetor conducente atrelado ao uso de redes sociais, aquilo que a faz tornar-se realmente uma TI colaborativa.

6.2 Presença de funcionalidades atreladas ao modelo 3C de colaboração

Reforçando a diretriz de escuta calcada nas dimensões do modelo 3C, o qual realça comunicação, coordenação e cooperação como dimensões da colaboração, questionou-se quanto

às funcionalidades inerentes às tecnologias colaborativas que se aninhavam sob os *macro constructos* encontrados naquele valioso *framework*.

Percebeu-se que os elementos se aconchegaram em torno dos aspectos base, conforme prevê a literatura, mas emergiram elementos associados à interação, que aparentam ser o elo inovador em termos de modelagem para os tempos de *web* 2.0 e itens estruturadores do processo, apontados como definidores de formato e uso da tecnologia, os quais contemplam eventos que se valem da melhor formatação, tais como gestão de conteúdos e seleção de pares.

Tal arranjo é significado no quadro 2, com as especificidades obtidas no exame dos dados.

Comunicação	**Cooperação**	**Coordenação**	**Interação**	**Processo**
Apoio à Comunicação síncrona e Assíncrona	Apoio à execução colaborativa de Tarefas	Coordenar tarefas entre vários colaboradores dispersos	Acessibilidade	Definição de grupos
Comunicação Partilhada	Có-criação	Facilitar negociação	Gerir pontos de vista	Fluxograma de trabalho
Facilitar o Contato	Disponibilizar Ferramentas (*tools*) Colaborativas	Permitir a partilha de informação e outros recursos (ficheiros)	Interatividade	Gestão de Conteúdos
Partilha e discussão de conteúdos	Facilitar a colaboração entre os participantes	Permitir a partilha de informação e outros recursos	Partilhar símbolos entre colaboradores	Organizar Contribuições
Permitir comunicação em tempo real	Permitir colaboração	Disponibilizar ferramentas colaborativas	Partilha de práticas	Partilha de documentos
Reunir contribuições	Partilha de Informação	Facultar trabalho em grupo síncrono e assíncrono	Sincronizar a atividade dos membros do grupo	Seleção de Pares
Repositório de documentos	Edição por Diferentes Pessoas		Memorizar interações entre Colaboradores	

Quadro 2– Conteúdos em torno das funcionalidades presentes em tecnologias colaborativas.

Ato contínuo foi solicitado aos respondentes que, uma vez indicadas as funcionalidades usualmente vistas na TICC, buscassem-nas também nas duas tecnologias concorrentes aos moldes da modelagem 3C. Isto foi obtido ideando um mapeamento entre as funcionalidades sugeridas, marcando-se aquelas que corporificaram um núcleo comum e, naturalmente, aquelas que fossem taxadas como divergentes. Dimensões emergentes ou complementares asseveraram o direcionamento real dado às tecnologias questionadas e permitiram ver funcionalidades segregadas em elementos tronco e variantes, como previsto nas espécies biológicas.

Funcionalidades Colaborativas - Comunicação	
GDSS	**Redes Sociais**
Comunicação partilhada em tempo real entre as pessoas envolvidas no processo de uso	Comunicação síncrona ou assíncrona de forma coletiva ou entre duas ou mais pessoas
	Partilha; troca ou divulgação de informações
Memorizar interações entre colaboradores	Memorizar interações entre elementos da rede social

Quadro 3 – Agrupamento das funcionalidades de comunicação (visão focada no modelo 3C).

Funcionalidades Colaborativas - Coordenação	
GDSS	**Redes Sociais**
Coordenar tarefas entre colaboradores	Execução colaborativa de tarefas
Coordenar etapas do processo de decisão quando vários decisores estão envolvidos	Suporte à coordenação de trabalho individual e em equipe
Facultar trabalho em grupo síncrono e assíncrono de colaboradores	Facultar interação síncrona e assíncrona de elementos pertencentes à rede social
Memorizar interações entre colaboradores	Sincronização entre pares

Quadro 4 – Agrupamento das funcionalidades de coordenação (visão focada no modelo 3C).

Funcionalidades Colaborativas - Cooperação	
GDSS	**Redes Sociais**
Permitir a edição de informação (simultânea) por diferentes pessoas	Criação de comunidades (grupos sociais)
Divisão do problema pelos membros mais especializados	Criação de conteúdos colaborativamente
Facultar às pessoas dispor os seus significados a outros bem como assimilar os significados	Disponibilização de conhecimentos
Partilha de documentos, práticas, símbolos, informação e outros recursos (arquivos)	Sincronização entre pares
	Partilhar símbolos entre elementos da rede social
	Partilha e discussão de conteúdos, recursos e informações

Quadro 5 – Agrupamento das funcionalidades de cooperação (visão focada no modelo 3C).

Outros vínculos associáveis a GDSS contemplaram alguns ecos de modernidade, mostrando que essa tecnologia, embora marcantemente ligada ao aspecto de decisão e fortemente restrita ao domínio do modelo 3C, ainda incorpora avanços da TI e talvez por isso persista, apesar de tomar menor espectro na linhagem das TICC. Tais aspectos manifestaram-se em menções como: acessibilidade, rastreabilidade, usabilidade da informação e transparência e partilha das

diversas perspectivas de uso e de opiniões; todos muito difundidos a partir da evolução de uso da TI em termos sócio-organizacionais e bastante pitorescos para as redes sociais.

6.3 Funcionalidades associadas às redes sociais

Já em relação às funcionalidades associadas às redes sociais, dada à dispersão de elementos delineadores na literatura de apoio, resolveu-se fazer o agrupamento considerando similitudes de significância, estipulando rótulos para representar as ações perceptivas de colaboração, como exibido na figura 3.

Assim, quer em estruturas de apoio, facilidades de uso e formas de acesso, foram destacadas funcionalidades que conectadas às possibilidades de exercer influências denotam como plausível ter redefinições de uso em feição colaborativa.

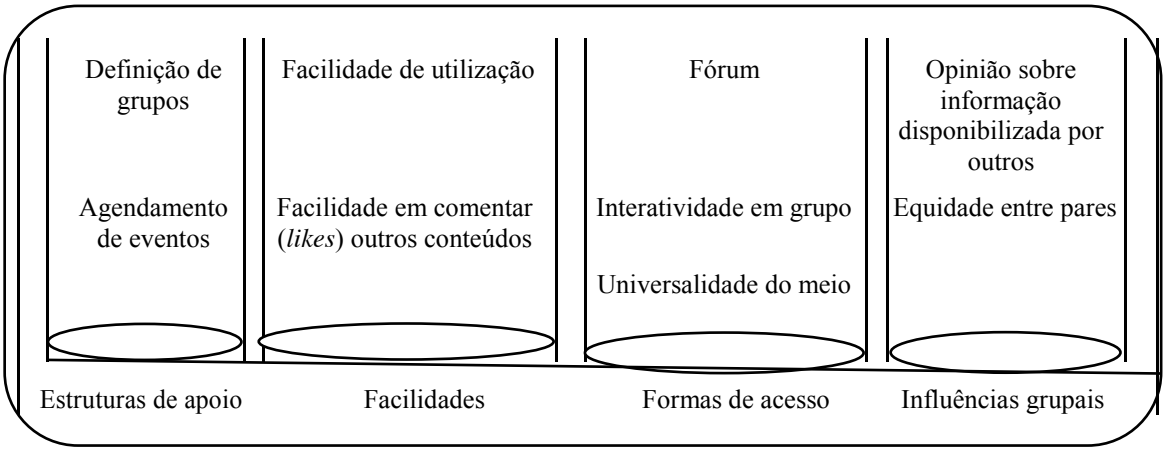


Figura 5 – Rótulos e dimensões para ações colaborativas no escopo das redes sociais.

Também se mostrou interessante arguir diretamente quais das funcionalidades dadas como congruentes para as duas tecnologias, em especial aquelas referentes às rotinas comuns, estariam em procedimento de aparente migração dos GDSS para as redes sociais. O vislumbre a destacar gravitaria em torno de linhas sucessórias

No exame evidenciaram-se campos fontes de migração, um deles nitidamente derivado do modelo 3C, qual seja a comunicação, que, como já se desconfiara, afigura-se como vetor-guia da transferência de funcionalidades. Os outros três têm a ver com:

- Interação, elemento colimado à comunicação e dotado de potencialidades ímpares na proposição *web* 2.0 (nativa às redes sociais), afigurando-se entre compartilhamentos de símbolos, opiniões recursos e perspectivas sociais;
- O apoio ao registro de fatos no processo colaborativo, mas com conotação de memória, o que propiciará eventualmente inferir roteiros migratórios nos processos de apropriação de decisão grupal;

- Apoio à estruturação do processo, a fim de aproveitar potencialidades mais amplas de processos usuais à computação grupal e prover maior uso de informação.

Tal investida propicia cogitar que a comunicação sendo o vetor guia da evolução, decorre da maior necessidade ambiental de fomentar-se interação e integração social nos processos de colaboração e decisional mirados nas novas tecnologias colaborativas. Isto ainda permite preservar dimensões de rotinas clássicas da estruturação de tarefas em grupo como *brainstorming* e provê um uso mais sofisticado da tecnologia em acesso e memória informacional.

7. Conclusões

A pesquisa debruçou-se sobre uma instigante questão, calcada na busca de transpor ideias de uma linhagem teórica para explicar fenômenos em outro segmento do conhecimento, em especial pelo apelo de uso sistêmico latente [Weber, 2011].

Sabe-se que os campos de computação e genética são relativamente distantes em sua raiz epistemológica, particularmente porque o primeiro estuda o inanimado, o racional, o produzido artificialmente sob simplificações bastante reducionistas, em que pesem algumas aventuras no campo da computação evolutiva [Mendes; Pereira, 2013]. Já o segundo busca a plenitude da explicação das espécies e consegue antagonizar a ciência e a mítica nos confrontos evolucionistas e criacionistas [Avelar, 2001]. Não obstante este aparente abismo, buscou-se neste relato uma eventual associação.

Para tanto, conjecturou-se uma linha de funcionalidades em tecnologias da informação, mormente aquelas de viés colaborativo, mesclando aqui e acolá conceitos da teoria Darwinista para relacionar fenômenos identificados e na investigação por questionários se pôde constatar ao menos coerência explicativa para se chegar às seguintes constatações:

- As funcionalidades em um dado *software* evoluem por indução seletiva, quer com base em ações de fabricantes/desenvolvedores, quer com base em requerimentos do ambiente;
- A evolução dessas funcionalidades se dá por meio de procedimentos bastante similares aos previstos no escopo da teoria Darwinista de seleção natural, em destaque quando se cogitam a aplicabilidade de conceitos como recombinação, derivação e mutação, olhando-se pelo prisma da importância relativa consignada à funcionalidade;
- As linhas explicativas sobre as características mor das funcionalidades identificadas a partir de levantamento literário encontraram boa convergência com os pilares conceituais do modelo 3C, enfatizando, portanto, a noção de colaboração e de tecnologias colaborativas.

Assim para a mutação funcional dos GDSS para redes sociais, o aspecto comunicação foi

determinante, sendo o ponto de entrecorte o avanço das funcionalidades de cooperação, restando fincadas no GDSS as funcionalidades de coordenação por maior verossimilhança com a função básica da linhagem original, a decisão.

A partir do entendimento das classes de respondentes ouvidas, a transferência de funcionalidades acontece na sequência de versões de um *software* e é decorrente tanto de prioridades técnicas e de mercado, quanto de necessidades de processos, mas principalmente de anseios sociais. Estes anseios concretizam-se em fatores econômicos e históricos e são resgatáveis como vetores que fazem plausível o uso de conceitos darwinistas para fornecer explicações da evolução das funções (e usos) de uma tecnologia da informação.

Desta forma, houve ao menos consistência na tentativa de estabelecer uma associação tendo por base a transposição de conceitos de uma teoria de outra área do saber, visto que pessoas convidadas a refletir sob esta lente tiveram suas opiniões conjugadas e isto propiciou gerar o conjunto discutido neste relato.

Estima-se com este *insight* ter suscitado a aplicabilidade da ideia darwinista de evolução, na qual espécies, variedades e subespécies, convivem um dado tempo e depois aquelas que melhor suportam as pressões externas e superam os obstáculos ambientais saem vencedoras, por um tempo, à tecnologia da informação (mesmo ciente da associação desta com as tecnologias negociais e sociais). Já que se falou em tempo evolutivo, torna-se evidente quer para o ramo da tecnologia da informação (e mesmo da computação evolucionária), este tempo é absolutamente curto em relação ao tempo da origem e evolução das espécies.

Referências

ARGYLE, M. **Cooperation - The Basis of Sociability.** London: Routledge. 1991.

AVELAR. T. **Evolução e criacionismo: uma relação impossível.** Vila Nova de Famalicão: Editora Quase. 2001,

BARKHI, Reza: The Effects of Decision Guidance and Problem Modeling on Group Decision-Making. **J. of Management Information Systems** 18(3): 259-285. 2002.

BARKHI, Reza; JACOB, V.S.; PIRKUL, H. The Influence of Communication Mode and Incentive Structure on GDSS Process and Outcomes, **Decision Support Systems**, v. 37, n. 2, p. 287– 305, 2004.

BENAVIDES, D. SEGURA, S. RUIZ-CORTEZ, A. Automated analysis of feature models 20 years later: A literature review. **Information Systems**. V. 35. 2010. p. 615–636.

BIDGOLI, Hossein. Group Support Systems: a new productivity tool for the 90's. **Journal of**

Systems Management. Jul-Ago 1996.

COSTA, J.P.; MELO, P.; GODINHO, P.. The AGAP system: a GDSS for project analysis and evaluation, **European Journal of Operational Research,** v. 145, n. 2, p. 287-303, 2003.

DAFT, R. **Administração.** São Paulo. 3ª ed. Atlas. 2008.

DARWIN, C. **A origem das espécies.** *E-book* baseado na tradução de Joaquim da Mesquita Paul, médico e professor. Porto-Portugal: Lello e Irmão (Editores). 2003.

DeSANCTIS, G. GALLUPE, R. Brent. A foundation for the study of group decision support systems. **Management Science,** v. 33, n. 5, Maio 1987.

ELLIS, C. A., GIBBS, S. J.; REIN, G. L. Groupware: some issues and experiences.

Communications of the ACM. v. 34, n. 1, p. 38-58, 1991.

ELLISON, N.; LAMPE, C.; STEINFIELD, C. Social network sites and society: current trends and future possibilities. **Interactions of ACM**. Jan-feb 2009.

FERREIRA, A.B.H. **Mini Aurélio: o dicionário da língua portuguesa**. 8. ed. Curitiba: Positivo. 2010.

FONSECA, Marcelo Jacques; GONÇALVES, Manuela Albornoz; OLIVEIRA, Marta Olívia Rovedder; TINOCO, Maria Auziliadora Cannarozzo. Tendências sobre as Comunidades Virtuais da Perspectiva dos Prosumers. **RAE Eletrônica**, v. 7, n. 2, art. 24, jul/dez 2008.

FUKS, H. GEROSA, M. The 3c Collaborative Model. **Informática na Educação: teoria & prática,** Porto Alegre, v.7, n.1, jan./jun. 2003.

FUKS, Hugo. RAPOSO, Alberto B. GEROSA, Marco A. LUCENA, Carlos J. P. Applying the 3C Model to Groupware Development. **International Journal of Cooperative Information Systems**, v. 14, n. 2-3, jun-set, 2005.

FUKS, H.; RAPOSO, A.B.; GEROSA, M.A.; PIMENTEL, M.; FILIPPO, D.; LUCENA, C.J.P. Inter- e Intra-relações entre Comunicação, Coordenação e Cooperação. **In: SBSC, 27., 2007, Rio de Janeiro-RJ. Anais...** Rio de Janeiro-RJ, XXVII SBSC, 2007.

GODFREY, M. W. GERMAN, D. M. The Past, Present, and Future of Software Evolution. **IEEE FoSM**. 2008. p. 129-138

GREER, D. RUHE, G. Software release planning: an evolutionary and iterative approach.

Information and Software Technology. 46. 2004. p. 243–253.

HESS, N. Darwinian Technology: Revolution from Evolution. **Modern advancements in technology and social dynamics**. Apr. 2011.

JOHNSON, B. R. LAM, S. K. Self-organization, Natural Selection, and Evolution: Cellular Hardware and Genetic Software. **BioScience.** Vol. 60 No. 11 December 2010.

KAPLAN, B.; DUCHEN, D. Combining qualitative and quantitative methods in information systems research: a case study. **MIS Quarterly.** v. 12, n. 4, p. 571-586. 1988.

KIOUSIS, S. Interactivity: A Concept Explication. **New Media & Society**. v.4, n.3, pp.355-383.2002.Acesso em 24 jun. 2011.

KOCK, N.; NOSEK, J. Expanding de Boundaries of E-Collaboration. **IEEE Transactions on Professional Communication**, v. 48, n. 1, march 2005.

LEE, H.; BUCKLAND, M.A.; SHEPHERDSON, J.W. A Multi-Agent System to Support Location-Based Group Decision Making in Mobile Teams. **BT Technology Journal**, v. 21, n. 1, p. 105– 113, 2003.

LEHMAN, M. M. RAMIL, J. F. **Evolution in Software and Related Areas**. ACM 2002.

LEWIS, TED. Software Architectures: Divine plan or digital Darwinism. **Bynary Critic**. Naval Postgraduate Scholl. Monterey. Ca. 1996.

LIMAYEM, Moez; BANERJEE, Probir; MA, Louis. Impact of GDSS: opening the black box. **Decision Support Systems**, v. 42, p. 945–957, 2006.

MARKS, ERIC A. **Business Darwinism: Evolve or Dissolve Adaptive Strategies for the Information Age**. New York: John Wiley & Sons, Inc. 2002.

MARKUS, M.L.; SILVER, M.S. A Foundation for the Study of IT Effects: A New Look at DeSanctis and Poole´s Concept of Structural Features and Spirit. **Journal of de Association for Information Systems**, v. 9, n. 10/11, pp. 609-632, Special Issue 2008.

McMILLAN, S.J. The Researchers and the Concept: Moving Beyond a Blind Examination of Interactivity. **Journal of Interactive Advertising**. v.5, n.2, pp.1-4. 2005.

MEHTA, Arpita. Application of Darwin's & Lamarck's theories in Business Management: a review article. **International Journal of Business Trends and Technology**. 2(6). 2012.

MENDES, R. PEREIRA, R. Current trends in bio-ontologies and data integration. In: **Distributed Computing and Artificial Intelligence Advances. AISC 217.** Switzerland: Springer. 2013. p. 579-586.

MONTEIRO, G. **Guia para elaboração de projetos, Trabalhos de Conclusão de Curso (TCCs), dissertações e teses.** São Paulo: Edicon - Editora e Consultoria, 1998.

MORGAN, G.; SMIRCICH L. **The Case for Qualitative Research.** Academy of Management Review**,** v. 5, n. 4, p. 491-500, 1980.

NEHANIV, C. L. HEWITT, J. CHRISTIANSON, B. WERNICK. P. What Software Evolution and Biological Evolution don't have in common. **Second International IEEE Workshop on Software Evolvability (SE'06).** Vienna. 2006.

OD. **Oxford Dictionaries**. Acesso em: 23 jun. 2011. Disponível em: <http://oxforddictionaries.com/>. 2011.

PEROTTONI, Rodrigo et al. Istemas de informações: um estudo comparativo das características tradicionais às atuais. **READ**: **Revista eletrônica de administração**. Porto Alegre, v. 7, n. 3, p.1-15, 2001.

PIMENTEL, M. GEROSA, M.A. FUKS, H. **Sistemas de comunicação para colaboração**. *In:*

Sistemas colaborativos. Rio de Janeiro: Elsevier, 2011, 375 p. PRESSMAN, S. **Engenharia de Software**. 6ª ed. São Paulo: McGraw-Hill.2006.

ROBBINS, S.P. **Comportamento Organizacional**. 11. ed. São Paulo: Pearson Prentice Hall, 2005.

SASAKI, Takahiro. TOKORO, Mario. Comparison between Lamarckian and Darwinian Evolution on a Model Using Neural Networks and Genetic Algorithms. **Knowledge and Information Systems 2** (201-222). Springer-Verlag London Ltd. 2000.

SCHERMERHORN, J. Determinants of interorganisational cooperation. **Academy of Management Journal**, v.18, n.4, 1975. p.846-856.

SPEK, Pieter van der. Managing software evolution in embedded systems. **Doctor aan de Vrije Universiteit Amsterdam.** VRIJE UNIVERSITEIT. 2010.

STRACHMAN, Eduardo; DEUS, Andréa Santos de. Instituições, Inovações e Sistemas de Inovação: interações e precisão de conceitos. **Ensaios FEE,** Porto Alegre, v. 26, n. 1, p. 575-604, jun. 2005.

STOLLER-SCHAI, D. **e-Collaboration**: Institut of Information Management. St. Gallen: University of St. Gallen. 2003. Disponível em: http://www1.unisg.ch/www/edis.nsf/wwwDisplayIdentifier/2767/$FILE/dis2767.pdf.

Acesso em: 12 nov. 2012.

TELLES, André. **A Revolução das Mídias Sociais. Cases, Conceitos, Dicas e Ferramentas**. São Paulo: M. Books do Brasil, 2010.

WEBER, B. H. Extending and expanding the Darwinian synthesis: the role of complex systems dynamics. **Studies in History and Philosophy of Biological and Biomedical Sciences**. 42. 2011 P. 75–81

ZUBEN, FERNANDO J. **Computação Evolutiva: Uma Abordagem Pragmática.** DCA/FEEC/Unicamp. 2000.

Value Focused Approach to Information Systems Risk Management CAPSI/2015

Sérgio Nunes [1], Gurpreet Dhillon [2], Mário Caldeira [3]

1) ISEG, Universidade de Lisboa, Portugal
snunes@advance.iseg.ulisboa.pt

2) Virginia CommonWealth University, USA
gdhillon@vcu.edu

3) ISEG, Universidade de Lisboa, Portugal
caldeira@iseg.ulisboa.pt

Abstract

Information Systems (IS) risk management is a challenge to every organization, in that they are exposed to cyber-attacks that bypass physical barriers. Organizations increase online business in order to remain competitive, but as a consequence their online exposure becomes greater. However their risk management practices and governance are inadequate in the face of increasing new threats and vulnerabilities. This paper presents a Multi-Objective Decision Model for assessing Information Systems Risks. The decision model is based on the values and perceptions of stakeholders. It uses the Value-Focused Thinking approach, as opposed to the predominant Alternative-Focused Thinking. The objectives serve as a basis for decision making in the context of Information Systems risk management in complex managerial situations..

Keywords: IS Risks, Value-Focused Thinking, Multi-Objective Decision Model

1. Introduction

Nowadays most enterprises have requirements to protect the information security from outsiders, competitors and even between employees from different departments. It has been proven that disgruntled employees are seen as a massive threat to information security as they have access to sensitive information that can put the company's reputation at stake, although they are frequently perceived by management as trusted personnel that follow safeguards [Dhillon, 2001]. The enforcing of regulatory requirements like Basel II or SOX by supervision entities force enterprises to conduct information systems risk assessments and perform strict management of information security.

Top management recognizes the need of information security, but how much of information security investment is enough? An information security budget is seen by top management as a black hole that sucks resources and investment without any returns. What are the return benefits of security investments? Although the information security manager is fighting daily to align business objectives with security objectives, it is not always possible to achieve a direct alignment between them. A key factor in influencing top management to expand the information security budget is risk.

Risk is the common language between information security and the business, by translating the technical details into business losses and bringing top management long term commitment to the security strategy [Vitale, 1986]. Security investments, most of the times, do not bring tangible added value to business, but they mitigate risk to an acceptable level by management by following a cost benefit approach. The security budget lives from those detected risks, based on the importance of the critical business assets that it protects. Information systems entail complex technological environments and are targeted by multiple threats with a probability of being exploited. A successful exploitation of a vulnerability in the information systems might expose critical business information. Top management is flooded with multiple risks and are urged by stakeholders to decide on the best alternative as soon as possible. Should the decision be focused in an alternative based approach?

[Keeney 1992] argues that alternative focused thinking limits the decision criteria by focusing only on the alternatives rather that concentrating on company's objectives that are driven by values. Alternatives are pushed by others to force a decision among measures that are not aligned with business objectives and can result on the least damaging alternative being chosen by top management without appropriate reflection on business values. Values should drive decision making as they drive intrinsically everything we do. Deciding on alternatives without understanding the values behind them and without matching the alternatives towards company's values will constrain the decision process. The right approach is a value focused thinking where values articulate different alternatives that achieve them, thus identifying better decision situations and changing a reactive decision process into a proactive process [Keeney, 1996].

2. Literature review

The literature review process started by looking at research that employed the value focused thinking in multiple areas namely energy, information systems (IS), information security and operations [Parnell et al., 2013]. The problem always resides in the ability to justify a difficult decision by weighting the multiple objectives [Keeney et al., 1986].

[Dhillon and Torkzadeh 2006] present an assessment of information systems security in organizations, in which they use a value focused approach with the major objective of maximizing information systems security in organizations. They interviewed 103 managers from multiple organizations and initially identified general values for managing information systems security and recorded them in a wishlist. In a second phase values are clustered, labeled and converted into security objectives. The third phase consists of the classification of the objectives in the fundamental and means objective group by elaborating the "why is it important?" (WITI) test. The final phase deals with the validation of the objectives list in a panel of experts. The research resulted in 86 objectives that were organized into 25 clusters with 9 fundamental and 16 means objectives. According to the authors a need to amplify the principles beyond confidentiality, integrity and availability arises for designing security, as other objectives were valued by the employees.

[Duarte and Reis 2006] use the VFT approach and the multiple attribute value theory to gather the objectives to help the Portuguese Public Administration to choose among projects implementation. The objectives considered are: "maximize the innovation, maximize the geographical impact (economics of scale), maximize the connection between partners and skills (networking effect), maximize the number of direct beneficiaries (Enterprises, Citizens and other Organizations), maximize the number of agents indirectly benefited (Enterprises, Citizens and other Organizations), maximize the technical skills of employment (jobs created and maintained), maximize the economic efficiency (economic sustainability) and maximize the synergies between actions (project integration)". The second step consisted of developing multiple attributes from the objectives to be able to weight projects and support decision making. They tested the evaluation framework against 5 project proposals, where 2 projects were terminated after evaluation, and developed a computer interface to be able to apply the same principle to other future project proposals in a simplified approach.

[Mishra and Dhillon 2008] develop control objectives based on the values of IT managers using a value focused approach. They classify controls according to scope of control and target of control. In scope of control, controls can be classified as technical, formal and informal. Technical controls target electronic information and can be for example the use of encryption algorithms or authorization mechanisms. Formal controls target organizational structure and management and an example is the enforcement of policies, standards and procedures. Informal controls deal with the importance of values and examples are the expectations of responsibility and accountability. The target of control can be input, behavioral and output. Input controls deal with the alignment of employees with business interest by applying training programs. Behavioral controls determine how work is accomplished in the organization through the

following of defined business processes. Output controls manage results by measuring the transformation process. With this classification process in mind they interviewed 54 individuals and gathered 7 fundamental objectives and 18 means objectives related to internal controls. By analyzing the outputs they conclude that technical controls by themselves cannot provide the necessary information security governance that an organization needs. They state that a formal control structure should be present in the organization by enforcing documented policies and procedures via top-down approach with top management awareness and involvement. Informal values should not be underestimated by the business as they enhance the impact of security governance and motivate employees, if the business and individual values are aligned to achieve a sense of ownership, ethical behavior and trust.

[Averill et al. 2009] discuss egress technology, efficient communication and safe evacuation during emergencies from large buildings using a VFT approach during a 3 day workshop. They gathered 7 main objectives: save lives and prevent injuries to occupants, save lives and prevent injuries to firefighters/responders, minimize property damage, minimize impact on property operations, minimize economic costs, reduce stress and reduce grief. The discovery of alternatives to respond to the multiple objectives is also a goal of the workshop, so it was divided into 2 stages. The first stage entails the individual definition of values and their objectives along with the alternatives to achieve those objectives. The second phase includes the group discussion of the proposed objectives and alternatives and the creation of multiple subgroups to reevaluate and discover new alternatives. The alternatives were classified on 3 criteria: quality, feasibility and creativity. Quality deals with effectiveness of the alternative, feasibility concerns the implementation within a 10 year period and creativity awards out of the box ideas.

[Dhillon and Chowdhuri 2013] collect individual values for protecting identity in social networks using a value focuses thinking mindset. They interviewed 147 individuals and summarized social media objectives across 19 clusters divided by 5 fundamental and 14 means objectives. The 5 fundamental objectives are: maximize end user trust, ensure development of social networking ethics, ensure authenticity of user identity, maximize identity management to make social networks useful and maximize social networking infrastructure protection. These results deliver a roadmap for individuals and organizations to be able to set up an identity protection strategy for social networks.

[May et al. 2013] define value-based objectives for Enterprise Resource Planning (ERP) systems planning. They defend there is commonly a misalignment between organizational business processes and ERP packages. To narrow this gap they use value focused thinking to

develop a list of objectives collected with 16 interviews across 3 ERP implementation case studies on Southern Europe. They argue that without determining stakeholders values prior to the ERP implementation, the project will only consider the technical implementation as the main critical success factor, disregarding other social, organizational and contextual factors. The results consisted of 13 means objectives and 4 fundamental objectives: minimize cost, ensure ERP benefits realization, enhance product and service improvement and maximize costumer relationship effectiveness. These objectives grounded in stakeholders values aid organizations to understand the complex technical and social issues related to ERP projects and provide the basis to develop an ERP strategic plan.

3. Methodological Approach

The basis of this research is value theory [Catton, 1954, 1959] and the method employed is the value focused thinking approach [Keeney 1992]. [Catton 1959] explains that valuing is defined by the intensity of the desire to obtain an object and that the preference follows a motivational pattern. [Catton 1959] defines value as: ”a conception of the desirable which is implied by a set of preferential responses to symbolic desiderata”. He states that value conceptions are “socially acquired”. He explains that some researchers advocate that values cannot be measured, but he argues that by following a judgment based on values it is possible to make predictions of decisions in a defined context. These values guide the decision makers in a decision analysis process. [Keeney 2004] defines decisions: ”as situations where the decision maker recognizes that a conscious choice can be made”. The ultimate goal by following value focused thinking in decision analysis should be to select the best alternative, but as that is not always possible due to hidden alternatives. The enumeration of values and the creation of objectives serve the principle of eliminating the bad decisions that looked good before, but do not accomplish any of the proposed objectives. The unframing of the decision process should be performed as soon as possible by defining the problem at hand and removing the psychological traps that influence our clear judgment in creating new alternatives without the anchoring in the previous alternatives.

[Keeney 1992] explains that values are principles for evaluation of the consequences of action or inaction towards a decision among different alternatives. He enumerates the uses of value focused thinking: ”uncovering hidden objectives, guiding information collection, improving communication, facilitating involvement in multiple stakeholder decisions, interconnecting decisions, evaluating alternatives, creating alternatives, identifying decision opportunities and guiding strategic thinking”.

[Gregory and Keeney 1994] argue that value focused thinking is a facilitator in a negotiation situation to reach consensus among stakeholders, as the unique list of objectives is part of the values and contributions from all the people involved in the decision process. This list of objectives forms the context to evaluate alternatives that assure the commitment from stakeholders, even if each stakeholder wants to push an alternative, he will have to justify the inclusion of the alternative as a consequence of multiple previous agreed objectives. The design of alternatives will have also to take into account the accomplishment of the type of objective, namely the division between fundamental and means objective.

[Keeney and Gregory 2005] research the process of the creation and selection of attributes that measure the achievement of the defined objectives of value focused thinking. They describe 3 types of attributes: natural, constructed and proxy attributes. The natural attributes are intuitive in nature, in which for example the number of fatalities per time frame is an attribute to the objective of setting automotive speed limits. The proxy attribute is characterized by not measuring the objective directly, but count in conjunction with other attribute to define the objective's achievement. Using the same example of setting the speed limit, the proxy attribute can be for example the number of accidents. The constructed attributes, are like the name explains, the construction of a scale when the natural attribute doesn't exist. Once the scale is known and continuously used, the constructed attribute becomes intuitive and resembles a natural attribute. Proxy attributes are most used when the intuitive natural attribute lacks information, so they apply to means objectives that influence achievement of the fundamental objective.

The combination of the different attributes into a value model follow 3 main independence concepts [Keeney 2001]: additive, preferential and utility independence. Additive independence means that the attributes do not escalate taking into account the consequences of the objective, for example saving money on product 1 and saving money on product 2 can be added to fulfill the objective saving money on products. Preferential independence deals with pairs of attributes that are independent from the other attributes. Utility independence normally involves risk situations where the risk taking strategy from one attribute is independent from the fixed levels of other attributes. These independence concepts lead to 2 value model alternatives. The additive value model functions when all attributes are additive independent and the multiplicative value model functions when each pair of attributes is preferential independent and one attribute is utility independent. The common uses of value models are the creation of alternatives, evaluating existing alternatives, product design, guiding information collection, identifying and resolving conflicts, facilitating decision making, and identifying decision opportunities.

Adapting the approach from [Keeney 1992] and [Shoviak 2001], the research process can be summarized into a 5 step methodology:

1. Collect the detailed list of values for the decision context;
2. Rewriting the values in a common form, transforming them into sub objectives and clustering them into main objectives;
3. The clustered objectives are classified into fundamental objectives and means objectives, using the "why it this important" (WITI) test;
4. Weight the fundamental objectives;
5. Develop attributes to measure the objectives.

4. Value focused IS risk management

The initial list of values for IS risk management were gathered by conducting semi structured interviews in Portugal with several security and IT professionals who represented a wide variety of job descriptions, such as the CIO, CISO or IT Manager, for example. The interviewees were representative of multiple business sectors, but were predominantly from consultancy, banking and the telecommunications industry. Should the information gathered require further explanation that was relevant for the research, we interviewed other employees from the same organization in order to clearly identify the context of the information collected. The interview data gathering approach is adequate, as values should not be constrained and should be intrinsic from the individual. Interviews were conducted within borders by using general targeted topics, broad categories and examples, but open questions were posed which allowed the respondents to reflect on their past decisions and enabled a review of their judgmental values. A total of 71 interviews were performed. No more were carried out because the total of values that were collected from the last interviews were repetitions of previous values, so we thus opted for theoretical saturation. A total of 612 IS risk management values were collected, and after the removal of duplicates, a total of 414 values were identified. These values were enlisted in a common form and followed the methodology of obtaining a wishlist from the interviewees. The values in a common form were then transformed into 114 distinct sub-objectives, and any duplicates were removed, which resulted in the same goal in different words, following a correlation and consolidation procedure. This transformation into sub-objectives is accomplished by applying an active verb which turns an objective into an effective action. These objectives were then sorted into 23 clusters, taking into account a shared common theme. These 23 clustered objectives were further classified into means and fundamental objectives, by

using the "why is this important" (WITI) test. This structured procedure is important for enabling reflection as to what individuals care about in a IS risk context, and for seeing how these objectives relate in terms of importance. The WITI test resulted in a total of 6 fundamental objectives and 17 means objectives, as can be seen in Table 1.

Means Objectives	**Fundamental Objectives**
-Ensure properly configured IT infrastructure -Promote IS risk performance metrics -Ensure ongoing monitoring of IS risks -Ensure IS risk management processes are audited -Maximize access control -Minimize IS risks related to IT service providers -Reduce human negligence -Maximize vetting of employees for IS risks -Ensure adequate internal communication regarding IS risks -Ensure adequate external communication regarding IS risks -Maximize IS risks management for critical information -Ensure information confidentiality -Ensure information availability -Ensure information integrity -Develop IS risk management competencies -Develop an IS risk awareness programme -Develop a training programme for IS risk management	-Ensure risk management governance -Maximize IS risk knowledge -Ensure IS security quality -Maximize responsibility and accountability for IS risks -Maximize compliance -Maximize the protection of human life

Table 1. Means and fundamental objectives for IS risk management

The objectives were weighted using the swing method (Kirkwood, 1997), whereby a panel of specialists in risk management and information security is asked to judge the importance of objectives designed for the global objective of minimizing risk in information systems. This approach leads to defining a local weighting to each sub-objective in a branch. All the local weightings in a branch sum up to 1, in order to fulfil the main objective. A multi-tier hierarchy is then evaluated with global weightings, whereby local weightings are multiplied to accomplish the main objective, using an additive function. For the purpose of weighting the fundamental objectives, we arranged a workshop to foment the discussion and collect the opinion of the attendees regarding the importance of each objective within the model. The workshop was

performed within an information security conference with the duration of 40 minutes. There were 31 participants with positions such as for example CISO, IT Manager, Auditor and Professor. We opted to weigh also the 3 basic pillars of information security (information confidentiality, integrity and availability) although they are considered in this model means objectives to accomplish the fundamental objective of ensuring IS security quality, because of their importance in minimizing specific risks inside information systems. The hierarchical model and the local weights are presented in Figure 1.

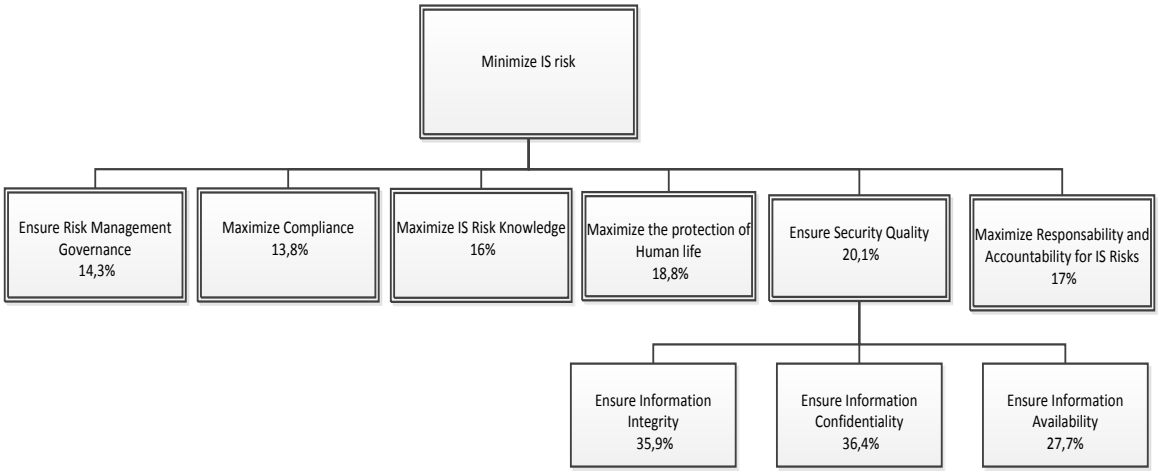


Figure 1 – Hierarchy of objectives

The next step would be to develop attributes to measure the performance of each objective within the model and transform these objectives into value functions that are situated between 0 and 1 to unify the measurement scales.

In the IS risk context, if we take as an example - the objective of "Maximize Compliance", then an attribute that needs to be measured could be the "Number of non-compliance issues per year". This is decreasing value function in which the target is 0 non-compliance issues per year for a value of 1.

5. Discussion

This section discusses each of the 6 IS risk management fundamental objectives and the 3 means objectives present in the model, taking into account existing best practices and by detailing the context in which they were structured and scored.

Ensure risk management governance includes the adoption of IT and security best practices. Adequate risk management governance entails the nomination of a IS risk committee, which has the role of discussing IS risk at top management level and which consults all relevant

stakeholders. The weight of this objective in the hierarchy is classified with 14,3% in the overall model which reflects the lack of risk governance maturity in most organizations in Portugal.

Maximize responsibility and accountability for IS risks deals with who does what, and who is ultimately responsible for risk mitigation measures. Most of the time, the person who is responsible for executing tasks is responsible for a specific delegated task, and the data owner is the person who is accountable for either accepting the risk, or deciding whether to implement additional safeguards. These data owners should be clearly identified in IS risk management and, together with their responsibilities in the risk management process, their role should be clear and objective. This identification and definition process prevents finger-pointing across the organization when a risk turns into a real situation. In the model this objective is weighted with 17% which demonstrates the need to assure adequate responsibility and accountability practices.

Maximize IS risk knowledge entails the creation of an intangible capability as a risk management organizational culture, in which each stakeholder is aware of existing IS risks and their management practices. Adequate testing of procedures and know-how should be performed in order to ensure that every employee is informed about the IS risk culture and knows what their role is in that risk management framework. In our model this objective represents 16% of the overall weight and it demonstrates that people skills in risk management are an important objective, but often left in the background in current organizations, as there are always more critical objectives appearing in risk management that need immediate attention.

Maximize compliance deals with ensuring that requirements from supervisory entities are met. This objective impacts on an organization's business directly, as sanctions are applied for lack of compliance and, in extreme cases, this may lead to the legal prosecution of management. An organization has to adopt those proven methodologies, frameworks and best practices that guarantee the maximization of compliance. Risk management has defined standards and best practices which should be adapted for IS risk management in the context of every organization. The documentation of clear IS risk policies and procedures, together with the definition of internal sanctions for their non-compliance, should be an initial step in establishing the risk management framework. Legal compliance issues should also be accounted for, such as, for example, data retention time frames, which differ for each country. Copyright management is also a compliance requirement which affects not only the software acquired by an organization, but also, for example, a IS attack with the intention of implanting illegal software. This objective was the least scored objective with 13,8%, this score details the significance of compliance in Portugal as the supervision entities are still benevolent in terms of compliance practices, functioning still as advisory when compared with mandatory requirements.

Maximize the protection of human life may seem an outside objective at first glance when dealing with IS risks. However this objective makes complete sense, after careful examination, and when considered within the mindset of those IS risks which affect critical infrastructures that may harm human life. Critical infrastructures are being attacked daily with advanced persistent threats (APT), with the goal of compromising their infrastructures, examples being energy companies and water management, for example. A recent example was when a computer worm named Stuxnet was created to attack nuclear power plants. These IS risks should be carefully managed as most critical infrastructures, although not directly connected to the Internet, are indirectly exposed to attacks and loss of human life may occur. When reflecting about cyber warfare, wars happen first in the cybersphere before a physical attack occurs. It is easier to attack a country that is rendered blind, through the lack of communications or power, for example. Recently, efforts are being secretly pursued by countries to enhance their IS competitive intelligence. At the individual level with the Internet of Things phenomenon, those risks will affect the common citizen in their houses, in their jobs and directly in their life with the adoption of e-health devices for example. The score of 18,8% details that this objective in risk management is still dormant in our lives at the moment, but it may increase dramatically as examples of human losses turn into a reality with vulnerabilities being exploited in information systems.

Ensure IS security quality objective aggregates security concepts, including explicitly the information confidentiality, integrity and availability triad. Examples of some of the fundamental concepts for assuring IS security quality are: information authenticity, reliability and non-repudiation. The ability to counteract aggressive actions can be ensured by robust authentication, complemented with strong auditing mechanisms and an adequate identity management of multiple stakeholders across multiple platforms, and also by using applications with strong access controls. Maintaining data privacy is also a critical factor for strengthening the quality of IS security. Being the model weighting workshop conducted within an information security conference we would expect that this objective is the most critical within the model. This is what happened with this objective weighting 20,1% of the model. Security practices and mechanisms represent the common safeguards to minimize information systems risks and ensuring security quality is a main concern of every information systems professional.

Ensure information confidentiality to prevent sensitive information being leaked to an unauthorized entity. This can be ensured by adopting adequate encryption measures when dealing with stored information, by encrypting laptops' hard drives or critical databases. Information transmitted can also use network protocols, which ensures that the encryption of data and data leak protection (DLP) mechanisms be implemented to prevent leakage by

disgruntled employees. Implementing an information classification program is also a crucial step for defining information's value and for protecting printed documents. In our model this objective contributes 36,4% as a means objective to ensuring IS security quality, being scored as the most important pillar in the CIA triad. This score reveals the importance of information as competitive advantage in most organizations. The global weight of this objective within the hole model is 7,3%. This global value is obtained by multiplying the local value of the means objective (36,4%) with the value of his parent fundamental objective (20,1%).

Ensure information integrity by adopting good change management practices. Change management allows for the tracking of responsibilities and prevents unauthorized and unprepared changes. Changes should be planned and a rollback plan should be available in case corruption of data occurs. The weight of this objective (35,9%) resembles the importance given to the previous objective "Ensure information confidentiality". The global weight of this objective within the hole model is 7,2%.

Ensure information availability when access to information is required by an entity. The risk of loss of information needs to be minimized, and adequate backup procedures and data recovery methods should be tested periodically. The transportation of backup information offsite should be evaluated in order to protect against disaster. Business continuity and disaster recovery best practices should be adopted by the organization, which should include the definition of recovery times and point objectives. The presence of high availability mechanisms in the infrastructure that supports critical processes is an added protection measure against failures, and protects against denial of service attacks. The score of this means objective is 27,7%, being the least prioritized objective in ensuring IS security quality. This reveals that organizations prefer to have their information systems not accessible, than to have information leakage or data integrity compromises. This result goes in opposite direction as the current trend of cloud computing being adopted by enterprises, where information availability is one of the most important benefits of cloud computing, raising issues in the integrity and confidentiality of information. The global weight of this objective within the hole model is 5,6%.

The value focused assessment of IS risks allows managers to base their decisions on stakeholders values that can be subject to analytical generalization in different cases. Yin (2003) argues along that line, and states that analytical generalization leads to theoretical propositions.

This study details the relationship between different objectives and establishes priorities for IS risk mitigation by dividing means and fundamental objectives. The fundamental objectives are weighted in a multi objective decision model. These risk objectives are focused on technical, procedural and social mechanisms. With these objectives an organization can evaluate what are

the critical ones for the current business and which safeguards will minimize the current risk level individually. Summarizing the discussion, this study makes the following major contributions:

- It uses Value Focused Thinking as a proven methodology for the exploratory study of the theme of IS risk management;
- It generates a detailed list of values and proposed objectives, based on their contribution to minimizing IS risk;
- It delivers a qualitative prioritized IS risk management decision model with fundamental and means objectives;
- It allows managers to simplify the decision process by justifying which risk management objectives are being addressed with the chosen safeguard.

6. Conclusion

This study presents a value focused assessment of IS risks across different organizations in Portugal and presents an IS risk management decision model. This model is composed of structured and weighted objectives which help management decide and justify investment in safeguards that guarantee the ultimate goal of risk mitigation. These objectives establish a roadmap for strategic planning within IS risk management and promote the involvement of stakeholders in the decision process by leveraging different objectives that may be pertinent to each stakeholder.

7. References

Averill, J. D., R. D. Peacock, R. L. Keeney, and P. D. Gallagher, "Rethinking egress: A vision for the future", Vol. Technical Note 1647. US Department of Commerce, National Institute of Standards and Technology 2009.

Catton, W. R. 'Exploring Techniques for Measuring Human Values'. American Sociological Review 19(1), 1954, 49–55.

Catton, W. R.: 'A theory of value'. American Sociological Review 24(3), 1959, 310–317.

Dhillon, G. "Violation of Safeguards by Trusted Personnel and Understanding Related Information Security Concerns." Computers & Security 20 (2), (2001). 165–172.

Dhillon, G. and G. Torkzadeh "Value-focused assessment of information system security in organizations." Inf. Syst. J. 16 (3), (2006). 293–314.

Dhillon, G. and R. Chowdhuri "Individual values for protecting identity in social networks." Thirty Fourth International Conference on Information Systems, Milan. (2013).

Duarte, B. P. and A. Reis, 'Developing a projects evaluation system based on multiple attribute value theory'. *Computers & operations research* 33(5), 2006, 1488–1504.

Gregory, R. and R. L. Keeney 'Creating policy alternatives using stakeholder values'. *Management Science* 40(8), 1994, 1035–1048.

Keeney, R. L., J. F. Lathrop, and A. Sicherman 'An Analysis of Baltimore Gas and Electric Company's Technology Choice'. *Operations Research* 34(1), 1986, 18–39.

Keeney, R. L. "Structuring objectives for problems of public interest." *Operations Research* 36 (3), (1988) 396–405. ISSN: 0030-364X.

Keeney, R. L. "Value-Focused Thinking: A Path to Creative Decisionmaking". *Harvard University Press*. (1992) ISBN: 9780674931985.

Keeney, R. L. "Value-focused thinking: Identifying decision opportunities and creating alternatives." *European Journal of Operational Research* 92 (3), (1996), 537–549.

Keeney, R. L. 'Modeling values for telecommunications management'. *Engineering Management, IEEE Transactions* on 48(3), 2001, 370–379.

Keeney, R. L. "Making better decision makers." *Decision Analysis* 1 (4), (2004), 193–204.

Keeney, R.L. and R. S. Gregory 'Selecting attributes to measure the achievement of objectives'. Operations Research 53(1), 2005, 1–11.

Kirkwood, C. W. *Strategic Decision Making, Multiobjective Decision Analysis with Spreadsheets*. Belmont:Wadsworth Publishing Company, (1997).

May, J., G. Dhillon, and M. Caldeira 'Defining value-based objectives for ERP systems planning'. *Decision Support Systems* 55(1), 2013, 98–109.

Mishra, S. and G. Dhillon "Defining Internal Control Objectives for Information Systems Security: A Value Focused Assessment." In: *ECIS*, (2008). pp. 1334–1345.

Parnell, G. S., D. W. Hughes, R. C. Burk, P. J. Driscoll, P. D. Kucik, B. L. Morales, and L. R. Nunn, 'Invited Review Survey of Value-Focused Thinking: Applications, Research Developments and Areas for Future Research.'. *Journal of Multi-Criteria Decision Analysis* 20(1/2), 2013, 49 – 60.

Shoviak, M. J. "Decision Analysis Methodology to Evaluate Integrated Solid Waste Management Alternatives for a Remote Alaskan Air Station". Wright-Patterson Air Force Base, Ohio: Air Force Institute of Technology, 2001

Vitale, M. R. “The growing risks of information systems success.” MIS Quarterly, (1986). 327–334.

Yin, R. K., “Case study research: design and methods”. *Sage Publications*, 3rd edition,2003

ARTIGOS

(APRESENTAÇÃO EM POSTER)

Alinhamento do negócio com os sistemas de informação: Médias e pequenas empresas portuguesas do setor Vitivinícola

Paulo António Salvador [1], Ramiro Gonçalves [2], Vítor Santos [3]

[1] NOVAIMS da Universidade Nova de Lisboa, Portugal

M2013055@novaims.unl.pt

[2] Universidade de Trás-os-Montes e Alto Douro, Portugal
ramiro@utad.pt

[3] NOVAIMS da Universidade Nova de Lisboa, Portugal

vsantos@novaims.unl.pt

Resumo

O alinhamento do negócio com os Sistemas de Informação tem sido encarado como um motivo de preocupação pelas organizações, principalmente pelos gestores que reconhecem a importância e benefícios que este alinhamento traz ao seu negócio. Os Sistemas de Informação devem ser encarados como sendo a combinação de procedimentos, informação, pessoas e tecnologias organizados de forma a permitirem a concretização mais eficaz dos objetivos do negócio. Neste artigo é discutido o contributo estratégico que as Tecnologias e Sistemas de Informação podem trazer a um sector tradicional em Portugal como o sector vitivinícola. Foram inquiridos as pequenas e médias empresas (PME's) deste sector e avaliados os resultados com base no modelo de Maturidade de Alinhamento Estratégico desenvolvido por Luftman. Em todos os seis critérios que fazem parte do modelo, as PME's do sector vitivinícola encontram-se ainda numa fase inicial (fase 1) de alinhamento entre o negócio e os sistemas de informação. Apesar de a maioria dos inquiridos terem consciência da importância dos sistemas de informação, estes são encarados principalmente como uma despesa e não como uma ferramenta que permitirá um aumento da competitividade e valor da empresa. Espera-se que este paradigma se altere nos próximos anos com a admissão de profissionais com maior formação para os lugares de gestão operacional.

Palavras-chave: Sistemas de informação, alinhamento estratégico, maturidade, setor vitivinícola

1. Introdução

A crescente necessidade de sobrevivência e/ou crescimento das organizações empresariais, independente do sector em estudo, leva a que haja uma restruturação das definições de estratégias de negócio por parte da gestão empresarial. Com isto surgiu a necessidade da utilização dos Sistemas de Informação (SI), tornando-se num recurso estratégico fundamental, sabendo que estes são capazes de fornecer informação que é necessária a uma boa gestão. Os SI permitem criar ferramentas para tomadas de decisão mais competentes levando assim à criação de vantagens competitivas seja na redução de custos seja na diferenciação dos produtos e dos serviços que são prestados pela organização. Estas vantagens permitem ainda, na grande maioria

das vezes, a criação de novas oportunidades e áreas de negócio [Santos 2015]. O alinhamento dos SI com as áreas de negócio das organizações foi classificado como uma das principais preocupações da gestão das organizações durante a última década [Malta & Sousa 2011].

Os gestores, apesar de conhecerem o seu potencial têm muitas vezes dificuldades em perceber como maximizar a rentabilidade do investimento em SI de modo a que a organização obtenha vantagens competitivas. A capacidade de promover este alinhamento parece depender de vários fatores entre os quais a qualidade do seu planeamento, desenvolvimento e exploração de SI e da qualidade da sua gestão.

A área de negócio que irá ser abordada de modo a avaliar qual o nível de inserção dos SI é um sector de negócio tipicamente tradicional em termos de Tecnologias de Informação (TI), o sector vitivinícola e também por sequência o sector Enológico português. Neste sector a utilização dos SI é maioritariamente, se não exclusivamente, baseada na utilização das TI para o controlo de custos e receitas através da utilização de um Software de Gestão, sendo que o mais amplamente utilizado por grande parte do sector é o INOGEST. Porém este sector encontra-se numa fase emergente na utilização de TI, como descrito por Pires [Pires 2008], tendo um maior sucesso na sua implementação que a Agricultura de Precisão devido às características específicas do sector das vinhas (parcelas de minifúndio que possuem uma elevada heterogeneidade).

2. Setor Vitivinícola

Nos últimos anos a agricultura tem sofrido a nível mundial uma profunda revolução ao nível tecnológico, quer através da utilização de sistemas de recolha e de integração, quer no tratamento da informação. Estes procedimentos quando utilizados em conjunto com os equipamentos agrícolas permitem obter uma agricultura de maior precisão. Isto permite a diminuição dos custos ambientais e de produção, levando consequentemente a um aumento da qualidade e quantidade das produções vitivinícolas. Este aumento permite um melhoramento das condições de trabalho de uma atividade que é na grande maioria dos casos uma atividade manual e de sol-a-sol, em todas as fases que constituem o ciclo de produção (plantações, fertilizações, mobilizações de solo, sementeiras e colheita) [Proffit *et al* 2006].

O sector vitivinícola, como atividade agrícola, vem assistindo nas últimas décadas à introdução de tecnologias que permitem uma monitorização de todo o ciclo da videira desde a poda, no final de uma campanha de vindimas, até ao início de uma nova campanha. Este tipo de monitorização permite seguir o vigor, rendimento e qualidade de toda a área de videiras,

acompanhando o ciclo vegetativo, permitindo ao vitivinicultor um acompanhamento mais eficaz e detalhado da vinha e das uvas.

A deteção remota é uma das tecnologias de monotorização atualmente utilizada na viticultura de precisão. A viticultura de precisão baseia-se na deteção e registo da radiação refletida a partir de um alvo localizado no solo que no caso das vinhas pode incluir a vegetação da videira, o solo e eventual vegetação do enrelvamento da entrelinha. Esta monitorização pode ser feita através de imagens obtidas por aeronaves não tripuladas (vulgarmente denominadas drones) que voam a baixas altitudes ou a partir de satélites. A quantidade de radiação que é obtida do alvo no solo é descrita através de um índice espectral de reflectância [Proffit *et al* 2006].

Um dos índices muito utilizado na viticultura de precisão é o *Normalised Difference Vegetation Index* (NDVI) para detetar áreas de vinha que se encontram infestadas por filoxera [Johnson *et al* 1996]. O NDVI possibilita também determinar a variabilidade espacial do vigor da vinha o que consequentemente permite prever a qualidade e rendimento [Hall *et al* 2002]. Os mapas obtidos através da monitorização dos alvos no solo constituem uma ferramenta muito importante para a gestão eficaz da vinha. Deste modo o vitivinicultor pode recorrer ao uso de fertilizantes, de fitofármacos ou simplesmente aos seus recursos hídricos apenas nos locais em défice conseguindo com isso não só uma maior rentabilidade dos seus recursos como também diminuir o impacto ambiental da sua exploração vitivinícola.

3. Alinhamento de Sistemas de Informação com o Negócio

Na perspetiva das organizações, os sistemas de informação são soluções organizacionais e administrativas baseadas nas TI que têm como função encontrar soluções para os desafios propostos pelo ambiente em que as organizações se inserem. As TI oferecem ferramentas aos gestores para planear, realizar previsões e também monitorizar os seus negócios com uma maior precisão [Laudon & Laudon 2004].

Os sistemas de informação reúnem, guardam, processam e facultam informação relevante de modo a torná-la acessível e útil para aqueles que a desejam utilizar [Buckingham *et al* 1987]. Poderão também ser vistos como subsistemas do sistema organizacional global, compreendendo a conceção do modo como é realizada a comunicação da informação e de que formas são executadas as ações de disseminação e transmissão da mesma [Falkenberg *et al* 1996].

São muitos os gestores empresariais, nas mais diversas áreas, que nas últimas décadas procuram corresponder os investimentos em tecnologia com as reais necessidades das suas empresas de modo a que estes investimentos em TI tragam uma fonte de mais-valias e inovação ao seu negócio [Porter, Millar, 1985; Ives et al, 1993; Henderson, Venkatraman, 1993; King, 1995;

Maes et al, 2000]. Com vista a atingir este propósito demonstrado pelos gestores, tem sido proposto e implementado o conceito de alinhamento entre as necessidades e objetivos do negócio com as TI, sendo este alinhamento estudado por muitos autores e investigadores de Alinhamento Estratégico entre Negócio e Tecnologias de Informação.

Modelo de Henderson e Venkatraman

No modelo de alinhamento estratégico de Henderson e Venkatraman [1993], o alinhamento é baseia-se em dois blocos básicos que são a integração funcional e a adequação estratégica que por sua vez abrange os domínios interno e externo na busca da melhor estratégia de negócio. No domínio interno, a organização tem uma preocupação com todos os aspetos que são críticos ao funcionamento do negócio, como a estrutura administrativa interna, os processos críticos para o negócio (engenharia ou reengenharia caso o processo já exista) e a qualificação dos recursos humanos com a obtenção de competências organizacionais. Por outro lado, o domínio externo é onde as organizações competem e têm de se empenhar de modo a terem as melhores decisões possíveis para evitar que o negócio sofra pouco revezes. Neste domínio as organizações têm também de se preocupar com os mercados onde o seu negócio está inserido, estudando as necessidades de oferta e procura do mercado e com tudo o que lhes possa tornar o seu produto mais atrativo e competitivo em relação à concorrência. Henderson e Venkatraman [1993], alertam que estes dois domínios têm que estar em alinhamento para que o modelo de alinhamento possa ser eficaz.

Modelo de Chan *et al*

De acordo com Chan *et al* [1997] o conceito de orientação da estratégia de negócio e a estratégia dos SI é muito importante, sendo que o alinhamento entre ambos permite determinar se os SI têm uma grande eficácia na performance do negócio. Os sistemas de informação e as tecnologias de informação têm sido usados como um meio de influência em certo tipo de negócios, como forma de otimizar a fusão de empresas, na reestruturação de indústrias e como forma de facilitação da competitividade global. Para os autores o alinhamento estratégico dos SI é um dos top10 desafios para os diretores de TI. Contudo o alinhamento continua a ser um conceito de difícil compreensão e de difícil medição. Com isto os autores, apresentam-nos várias maneiras de conseguirmos compreender como o alimento estratégico dos SI pode ser importante e quais as maneiras de o obtermos. Para tal concebeu um modelo conceptual como podemos ver na figura 1.

Este modelo conceptual demonstra que a estratégia de alinhamento dos SI situa-se entre a orientação estratégica de negócio e a orientação estratégica dos SI, sendo que está diretamente relacionada com a eficácia dos SI. Os autores do modelo [Chan *et al* 1997] desenvolveram

quatro instrumentos de medição, sendo que cada um destes instrumentos deve vir acompanhado por diferentes informações-chave.

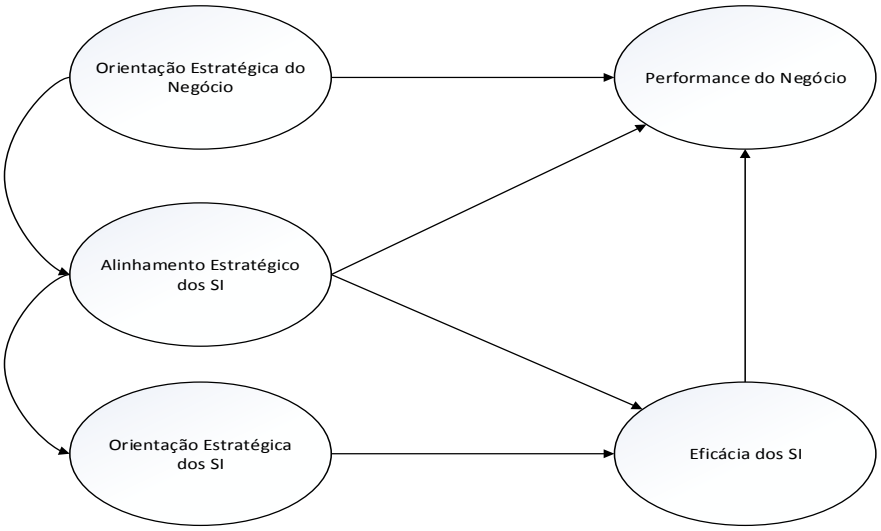


Figura 1 – Adaptação do Modelo conceptual de Chan *et al* [1997]

Estratégia de Negócios realizadas **-** O primeiro instrumento que consta na tabela de instrumentos de medição descrito por Chan, é uma versão melhorada da STROBE (orientação estratégica de organizações empresariais, sigla inglesa) descrito por Venkatraman [1985]. Estes agrupamentos emergentes descritos por Chan são muito próximos dos propostos por Venkatraman, contudo estes são formados de forma diferente, Chan propõe oito enquanto Venkatraman só propõe seis.

Performance do Negócio - Este é o segundo instrumento, sendo que de acordo com Chan *et al*, é complexo e multifacetado, com uso de diferentes partes interessadas. Tem diferentes meios de medir a performance do negócio, sendo que as principais são crescimento de mercado e a rentabilidade. Mas para os autores, as duas dimensões principais anteriormente propostas por Venkatraman e Remanujam [1986] devem ser divididas em quatro dimensões distintas, que são crescimento de mercado, rentabilidade, inovação de produtos/serviços e a reputação da empresa.

Eficácia dos Sistemas de Informação – Para Chan *et al*, a eficácia dos SI foi medida utilizando um instrumento que é composto na sua essência por duas subcategorias, a satisfação do cliente e o impacto organizacional. Para medir a satisfação do cliente foi usado como base um questionário anteriormente utilizado e validado [Pearson 1977, Bailey and Pearson 1983] contendo agora questões que permitem que os inquiridos comentem a sua satisfação com os

sistemas da empresa invés de se focarem nos sistemas pessoais, tornando o seu formato mais consistente com os outros instrumentos de análise.

Estratégia dos Sistemas de Informação Implementada – O quarto instrumento foi concebido, de acordo com Chan et al, para avaliar a "orientação estratégica do portfólio existente das aplicações dos sistemas de informação" (STROEPIS, sigla inglesa). Para acomodar o alinhamento estratégico dos SI, o STROEPIS foi desenhado sobre as mesmas oito dimensões do STROBE.

Modelo de Luftman

O alinhamento estratégico não é um evento isolado e singular, mas sim um processo contínuo em que a empresa vai adaptando as suas estratégias e infraestruturas às necessidades dos clientes e aos desafios impostos pela concorrência [Luftman 1996]. De acordo com este princípio, Luftman desenvolveu um modelo de avaliação da maturidade do alinhamento das TI com o negócio baseado em 12 elementos, que podem ser organizados no modelo de Handerson and Venkatraman [1993]. A figura 2 ilustra o modelo de avaliação da maturidade do alinhamento das TI com o negócio.

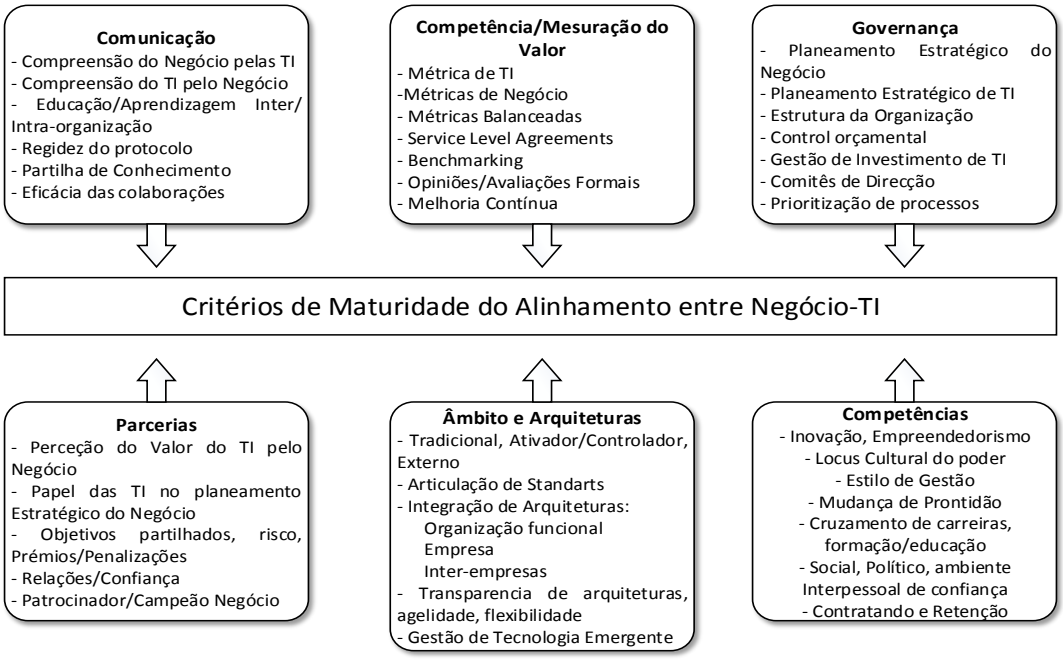


Figura 2 – Adaptação do modelo avaliação da maturidade do alinhamento do IT com o negócio [Luftman 2000].

Este modelo é constituído por 6 critérios principais: comunicação, competência / mensuração do valor, governança, parceria, âmbito e arquiteturas e habilidades. Cada um destes critérios possui

um conjunto de práticas ou atributos, sendo que cada uma destas é mensurada por cinco níveis de maturidade, onde o nível 1 corresponde a um nível inicial e o nível 5 corresponde a um processo considerado otimizado.

Comunicação - Este critério serve para avaliar a compreensão entre a parte técnica e a parte da gestão do negócio, sabendo com isto se ambas tem uma interação puramente casual ou interagem regularmente, promovendo uma troca de ideias e/ou posições sobre assuntos de alinhamento das empresas.

Competência / Mensuração do Valor - Este critério, segundo Luftman, vai servir para que as empresas possam fazer uma medição mais viável da sua performance, estimar o valor dos seus projetos e avaliar o valor do TI em termos da sua contribuição para o negócio. As empresas devem fazer avaliações após as conclusões do projetos, para identificar os pontos positivos e negativos com vista a identificar quais os processos internos a ser melhorados.

Governança - Este critério vai trazer a garantia que todos os participantes do negócio e das TI possam rever prioridades de modo a alocar os recursos das TI em alinhamento com o negócio. Na maioria das vezes, o alinhamento é apresentado como uma parte de uma boa governança, podendo ser este um inibidor ou motivador de uma boa governança e por isso foi considerado por Luftman como uma das seis dimensões do seu modelo.

Parceria - Este critério vai servir para que exista uma garantia de que o papel das TI nas estratégias de negócio seja relevante. Estas parcerias têm como objetivo final uma evolução das TI de modo a que estas possam direcionar mudanças no negócio, nas estratégias do mesmo e nos processos.

Âmbito e Arquiteturas - Quando pretendemos fazer uma análise sobre o nível de maturidade do alinhamento entre o negócio e as TI, um elemento chave é avaliar a maturidade na compreensão da parte técnica, sendo que esta vai muito além do papel mais tradicional das TI, *back office* e *front office* da organização. Neste critério, Luftman sugere que as TI assumam uma função de suporte a uma infraestrutura de negócio muito mais flexível e por outro lado transparente a todos os parceiros do negócio. Esta será obtida através de avaliações regulares e também da aplicação de tecnologias emergentes [Bernardo, A 2014].

Habilidades - Este critério procura que exista uma avaliação ao nível dos recursos humanos e quais as suas competências, muito para além das considerações mais tradicionais, como o salário, formação, *feedback* de desempenho e oportunidade de carreira. Este critério inclui também a parte social da organização e avalia o ambiente cultural que existe na mesma.

Como referido anteriormente cada um destes critérios é constituído por diversos atributos e é avaliado com base nos cinco níveis de maturidade do alinhamento estratégico (AE):

Nível 1 – nível inicial ou processo improvisado – neste nível a organização não se encontra apta para uma estratégia de alinhamento do TI com o negócio;

Nível 2 – Processo comprometido – neste nível encontram-se as empresas que estão comprometidas com o processo de maturidade do AE e que com isto começam a entender que existe uma mais-valia e oportunidades no mesmo. Este nível é muitas vezes direcionado às funções organizacionais dentro das empresas;

Nível 3 – Processo Estabilizado e Focado – neste nível as organizações estabeleceram processos focados na maturidade da AE, sendo que a este nível existe uma inserção das TI nos negócios. Neste nível, para além de um impulsionamento dos ativos de TI ao nível da organização a informação gerada pelas TI é já utilizada para tomar diversas decisões de negócio por parte da direção;

Nível 4 – Processo Gerido ou Melhorado – neste nível existe uma boa governança e os serviços que reforçam o TI como um centro de valor para a empresa. A este nível as organizações impulsionam os seus ativos de TI na extensão da empresa de modo a obter vantagens competitivas, sendo a este nível o TI visto como um contribuinte estratégico para o sucesso da organização;

Nível 5 – Processo Otimizado – neste nível as organizações apresentam processos integrados de governança onde se encontram os processos de planeamento estratégico de TI e do negócio. As organizações que se encontram neste nível, através de um impulsionamento dos ativos de TI a toda a empresa, pretendem com isto estender o alcance da organização para a cadeia de fornecimento de clientes e de fornecedores.

4. Metodologia

Este estudo utilizou como suporte à investigação o modelo de Luftman. Este modelo é o mais adequado para o estudo do alinhamento de sistemas de informação com o negócio considerando que se baseia numa avaliação quantitativa e por isso de fácil interpretação.

Os dados foram obtidos através de um inquérito realizado *online*. Este tipo de metodologia permite introduzir um grande número de variáveis típicas do sector vitivinícola. Neste sector, não só as regiões possuem características e climas diferenciados como as castas usadas variam de região para região, fazendo com que as necessidades específicas e o modo de produção varie.

Por sua vez, a utilização de inquéritos permite a utilização das ferramentas de AE descritas pelo modelo de Luftman, o que permite uma correta avaliação do alinhamento do SI com o negócio.

Todos os critérios do modelo de Luftman [2000] foram avaliados numa escala de 1 a 5 (escala de Likert). Após apreciação individual de todos os critérios do modelo de Luftman, coadjuvado pela escala de 1 a 5, os resultados obtidos foram utilizados para se avaliar a inserção e maturidade de alinhamento dos sistemas de informação com as pequenas e médias empresas do setor vitivinícola.

Para a distribuição do questionário online as organizações que fazem parte do sector vitivinícola português foram distribuídas pelas respetivas regiões vitivinícolas. Segundo dados existentes no portal do vinho português, Infovini, em Portugal em 2014 existem 1750 produtores de vinho. A distribuição por região é apresentada na tabela 1. Do universo total, as empresas com um volume de negócio superior a 100.000 garrafas por época foram eliminadas de modo a focar este estudo apenas nas pequenas e médias empresas, um universo de 1650 organizações.

Regiões Vitivinícolas	Número de Produtores
Vinhos Verdes	412
Trás-os-Montes	92
Douro	450
Bairrada	94
Dão	184
Beira Interior, Távora-Varosa e Lafões	45
Lisboa	72
Tejo	97
Península de Setúbal	69
Alentejo	204
Algarve	4
Madeira	27

Tabela 1 – Distribuição de Produtores Vitivinícolas por Região, fonte: Infovini, adaptado

5. Análise de Resultados

Do universo de organizações existente no setor vitivinícola que colaboraram no estudo, 71% dos inquiridos consideraram que é importante um alinhamento entre os sistemas de informação e a sua área de negócio, muito embora esse valor diminui para 45% quando perguntamos qual o nível de inserção dos Sistemas de Informação na sua atividade. Dos inquiridos, 68% admite

fazer parte do processo de decisão e formulação das estratégias de sistemas de informação e todos reconhecem que a sua organização investe muito pouco em sistemas de informação, um máximo de 5% das despesas da sua organização, muito embora 67% tem a perceção da importância dos sistemas de informação no melhoramento das suas atividades de negócio dentro da sua organização. Após uma análise mais detalhada aos critérios do modelo de Luftman, foram obtidos os seguintes resultados:

Critério 1 – Comunicação – como podemos verificar na figura 3 ainda existem muitas limitações quanto ao entendimento da área de negócios pelo pessoal de TI. Embora 33% dos inquiridos afirmem que existe uma proximidade entre a área de negócios e as TI, a grande maioria (47%) confirma que essa proximidade é limitada ou não existe. Estes resultados refletem o facto de muitas das pequenas e médias empresas não possuírem um departamento de Sistemas e Tecnologias de Informação, recorrendo a profissionais apenas para a resolução de problemas pontuais. Deste modo, as TI não são utilizadas como forma de mais-valia para o negócio. De acordo com os níveis de maturidade presentes no modelo de Luftman, a comunicação nas pequenas e médias empresas do sector vitivinícola português encontra-se ainda no nível 1 – inicial ou improvisado.

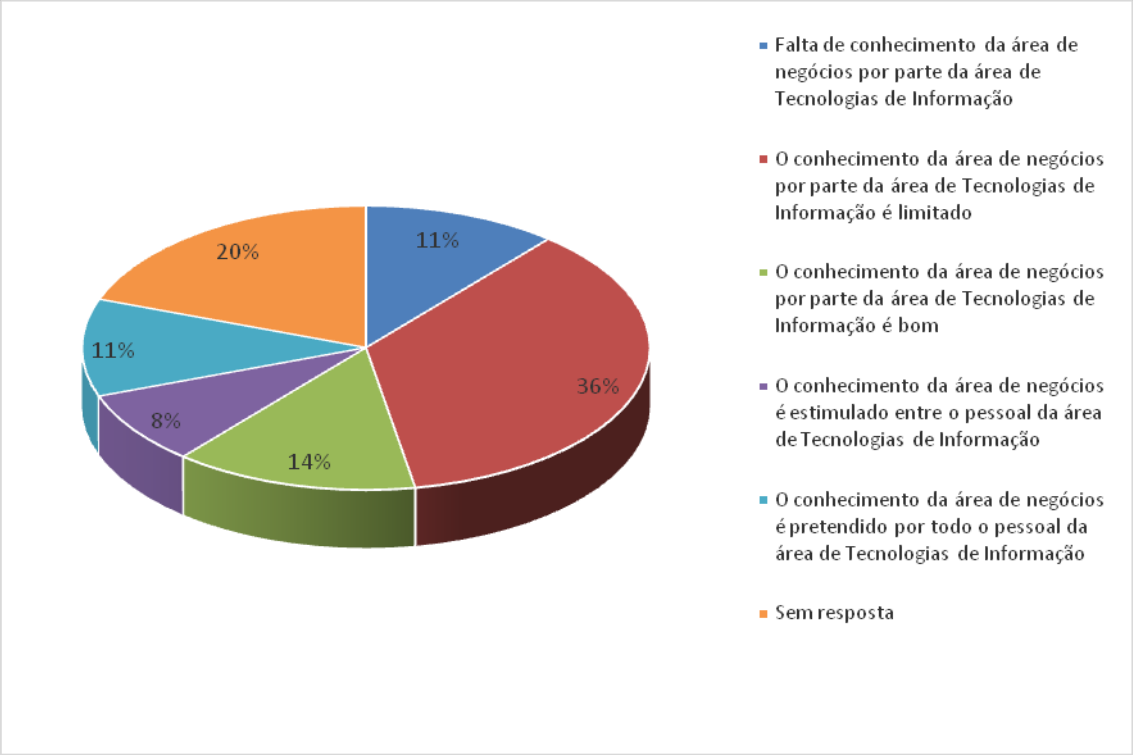


Figura 3 - Qual o nível de conhecimento sobre a área de negócio da sua organização que possui o pessoal da área de Tecnologias de Informação?

Critério 2 – Competência / Mensuração do Valor – neste critério, 47% dos inquiridos considera os ativos de TI como um ativo/investimento ou como uma forma de controlo de custos, com enfase na produtividade e/ou eficiência das operações e da infraestrutura. Na grande maioria organizações (72%) não existe uma clara preocupação com medidas que permitam medir quer internamente quer externamente o desempenho do negócio nem com a importância e facilidade que o uso dos sistemas de informação permitem a realização destas medições. Quanto à periodicidade das avaliações formais, estas são inexistentes (36%) na grande maioria das organizações ou nunca foram realizadas (28%). Estes resultados permitem inferir que neste critério o sector em estudo encontra-se no nível 1 – inicial ou improvisado.

Critério 3 – Governança – quanto à realização de um planeamento estratégico de negócios existem já uma maior dispersão nas respostas: 17% das respostas indicam que este não é realizado, 44% afirma que este é realizado ao nível da unidade funcional e com pouca colaboração das TI, 16% afirma que este existe e é um planeamento de funções, com alguma envolvência das TI e 33% dos inquiridos não sabe/não responde. Nas organizações do setor vitivinícola, as estratégias de negócio encontram-se enraizadas e limitadas às unidades funcionais, sendo que a área de TI possui pouco ou nenhuma influência neste planeamento. Quanto à formalização do planeamento estratégico dos SI, 33% das organizações afirma que este não é realizado e cerca de 16% afirma que este é realizado apenas ao nível da unidade funcional. As organizações que constituem o setor vitivinícola ainda mostram muita resistência em associar o investimento em TI com vantagens competitivas e de valor. Segundo os níveis de maturidade que estão presentes no modelo de Luftman, podemos constatar que neste critério o universo em estudo encontra-se no nível 1 – inicial ou improvisado.

Critério 4 – Parceria – as áreas de negócio do setor vitivinícola ainda não consideram as TI como um parceiro valioso, com o qual podem aumentar o seu valor e sua competitividade no sector (3%). Apesar de a área de negócio ainda não considerar e ter uma perceção do real valor das TI, cerca de 47% vêm já as TI como um ativo/investimento e não só despesa e 10% admite mesmo ter um papel facilitador de futuros negócios. No entanto, a maioria das organizações precisa ainda de melhorar o relacionamento entre as áreas de negócio e as TI tendo em conta que apenas 13% cumpre os processos de relacionamento criados. Os restantes inquiridos referem que os processos ainda não foram criados ou que estes nem sempre são cumpridos. Nas organizações do setor vitivinícola ainda é desvalorizado o apoio e patrocínio de ações das TI pela área de negócio, com uma percentagem significativa (37%) a admitirem que este é inexistente e só 16% admite que este apoio e patrocínio é feito pelo CEO. A análise das respostas dadas permite constatar que neste critério o sector em estudo encontra-se ainda no nível inicial (nível 1).

Critério 5 – Âmbitos e arquiteturas – o papel das tecnologias de informação nas pequenas e médias empresas do sector vitivinícola é ainda muito tradicional. Neste universo as TI tem como principal função o de apoio à atividade, maioritariamente na utilização do Word como processador de texto (73%), na consulta de *email* (60%) e ainda para faturação através da utilização de software específico (25%). Apenas uma pequena fração, 2% do universo de inqueridos utiliza as TI como sistema informático integrado de monitorização. A integração é baixa ao nível de toda a organização e inexistente com os parceiros de negócio. A grande maioria das organizações (83%) não reconhece que a infraestrutura de TI como um fator importante ao desenvolvimento do negócio, pelo que sem uma mudança de mentalidades este processo encontra-se muito pouco desenvolvido, encontram-se no nível 1 nos critérios do modelo de Luftman.

Critério 6 – Habilidades – quando confrontadas com a questão relativa às iniciativas de inovação e empreendedorismo, 43% dos inqueridos afirmou estas eram encorajadas ao nível da unidade funcional. Embora verificou-se que ainda não existe um forte encorajamento ao nível de toda a organização, pois estas são organizações pequenas com poucos funcionários muitas vezes sazonais. Em grande parte destas organizações, 40%, as tomadas de decisão-chave sobre as tecnologias de informação encontram-se restritas à alta gestão, embora em cerca de 21% comece a existir distribuição pela organização com o envolvimento das unidades funcionais, o que já demonstra ser um indicador motivador de que existem mudanças a ocorrer. No entanto, presentemente as pequenas e médias empresas do sector vitivinícola encontram-se também neste critério no nível inicial ou nível 1.

6. Conclusões

O setor vitivinícola é um setor ainda pouco virado para as novas tecnologias, embora com o aparecimento de uma nova geração de vitivinicultores esse paradigma esteja lentamente a mudar. As organizações que compõem este setor tão tradicional começam a reconhecer a importância dos sistemas de informação não só como um complemento mas também como um fator de diferenciação e de competitividade, quer a nível nacional como a nível internacional, criando valor e novas oportunidade de negócio. Apesar de a grande maioria dos inquiridos considerarem que é importante o alinhamento entre os Sistemas de informação e a área de negócio, a realidade é bastante diferente, pois muitos deles não o implementam na prática como foi referido no ponto 5. De facto, observou-se que a grande maioria das organizações de pequena e média dimensão não possuem uma área dedicada às tecnologias de informação.

O modelo de Luftman permitiu uma avaliação do alinhamento do negócio com as TI com base em 6 critérios. Cada critério foi avaliado individualmente e classificado em relação ao nível de maturidade, de 1 a 5, obtendo-se os resultados apresentados na tabela 6.1.

Critérios de Luftman	Níveis de Maturidade
Comunicação	2
Competência / Mensuração do Valor	1
Governança	1
Parceria	2
Âmbito e Arquitetura	1
Competências	2

Tabela 6.1 – Resultados do setor vitivinícola quanto ao nível de alinhamento entre os Sistemas de informação e o negócio

As organizações avaliadas encontram-se entre o nível 1 e 2 de maturidade dependendo do critério em estudo. No âmbito das competências, comunicação e parcerias as organizações iniciaram já um processo de AE com o pessoal de TI principalmente quando diz respeito à disponibilidade para a mudança. No entanto, apesar da motivação, as TI ainda não têm um papel ativo na definição das estratégias da empresa e são utilizadas maioritariamente como uma ferramenta de suporte à atividade (correio eletrónico, processamento de texto, faturação).

Este nível de maturidade é essencialmente devido ao facto de as organizações se encontrarem numa fase muito inicial da utilização dos sistemas de informação. Muito embora as organizações admitam investimentos em TI, este representa no máximo 5% das suas despesas. Estes resultados refletem a falta de pessoal nas áreas de TI destas organizações, sendo que estes são contratados pontualmente em caso de necessidade da organização. Apesar de a grande maioria dos inquiridos terem consciência da importância dos sistemas de informação, estes ainda representam uma despesa e não uma forma de melhoramento e de criação de valor para a organização. A maior limitação para um aumento no investimento em TI deve-se ao facto de as pequenas e médias empresas terem volumes de negócio baixos. Em comparação, as grandes empresas do setor, tem uma maior participação dos TI no negócio com cerca de 11 a 20 funcionários na área de informática.

Em resumo, existe ainda um longo caminho a percorrer pelas pequenas e médias empresas do sector vitivinícola por forma a atingirem o alinhamento estratégico entre os sistemas de

informação e o negócio. Este alinhamento só é possível com uma mudança de atitudes e de mentalidades e consequentemente um aumento do investimento na área dos sistemas e tecnologias de informação. Estas mudanças poderão ser iniciativa das gerações mais recentes que têm geralmente um maior envolvimento com as TI. Para além disso, espera-se que o aumento de formação nas áreas de gestão organizacional possa permitir que o processo seja mais célere.

Terá que ser implementado toda uma nova cultura organizacional para que este alinhamento entre os sistemas de informação e o negócio possa se tornar numa realidade neste setor tão tradicional. Quando isto se tornar uma realidade estarão dados os passos para que as organizações do setor vitivinícola se tornem mais competitivas e para que os sistemas de informação sejam considerados um fator que proporciona valor à organização.

7. Referências

Arkun, S. (2000). "Airbone remote sensing of the vine canopy. Vineyard monitoring and management beyond". Wagga Wagga, Austrália.

Andrews, D., Nonnecke, B., & Preece, J. (2003). "*Conducting Research on the Internet: Online Survey Design, Development and Implementation Guidelines*". International Journal of Human-Computer Interaction, 16(2), 185-210.

Bailey, James E. and Sammy W. Pearson, "Development of a Tool for Measuring and Analyzing Computer User Satisfaction", Management Sci., 29, 5 (May 1983), 530 – 545

Bernardo, A. (2014). "Alinhamento do Negócio com os sistemas de Informação: O caso de pequenas e médias empresas portuguesas no setor Agroalimentar, ISEGI-NOVA, Universidade Nova de Lisboa.

Boar, Bernard H. (1994). "Practical Steps for Aligning Information Technology with Business Strategies. How to achieve a Competitive Advantage." AT&T - John Wiley & Sons. EUA.

Bramley, R. G.V. (2000). "Measuring within vineyard variability in yield and quality attributes." Vineyard monitoring and management beyond. Wagga.

Buckingham, R. A., R. Hirschheim, F. F. Land e C. j. Tully (1987), "Information Systems curriculum: a basis for course design" in Information Systems education: recommendations and implementation, Cambridge University Press.

Chan, Y. E., Huff, S. L., Barclay, D. W., & Copeland, D. G. (1997). Business strategic orientation, information systems strategic orientation, and strategic alignment. Information Systems Research, 8(2), 125-150.

Deutskens, E., De Ruyter, K., Wetzels, M., & Oosterveld, P. (2004). Response rate and response quality of Internet-based surveys: An experimental study. Marketing Letters, 15(1), 21-36

Esser, A.; Ortega, R.; Santibánez, O. (2002) "Viticultura de Precisión, Nuevas tecnologias para mejorar la eficiência productiva em viñas." Revista 12, Departamento de Ciencias Vegetales, Centro de Agricultura de Precisión.

Falkenberg, E. D., W. Hesse, P. Lindgreen, B. E. Nilsson, J. L. H. Oei, C. Rolland, R. K. Stamper, F. J. M. Van Assche, A. A. Verrijn-Stuart e K. Voss (1996), "FRISCO: a FRamework of Information Systems Concepts", IFIP WG 8.1 Task Group Frisco.

Hall, A., Lamb, D.W., Holzapfel, B., Louis, J. (2002) "Optical remote sensing applications in viticulture – a review". Australian Journal of Grape and Wine Research, 8:36-47.

Henderson, J.C.; Venkatraman, N. (1993). "Strategic Alignment: Leveraging Information Technology for Transforming Organizations". IBM Systems Journal, vol.32, NO 1. Reimpressão em : IBM Systems Journal, vol.38, NOS 2&3. Pág. 472 - 484. 1999.

Henriques, P.; L Carvalho, M. L.; Costa, F.; Pereira, R.; Godinho, M. L. F. (2009) "Caracterização e Eficiência Técnica de Explorações Vitícolas da Região Alentejo", Ciência Téc. Vitiv. 24 (2) 73-80.

Ives, B. Jarvenpaa, S.L.; Mason, R.O. 1993. "Global business drivers: Aligning information technology to global business strategy". IBM Systems Journal, vol. 32, NO 1.

Johnson, L.; Bosch, D.; Armstrong, R..; Lobitz, B.; Baldy, R.; Weber, E.; Debenedicts. (1996). "Airbone imaging for vineyard canopy evaluation. California Agricultures, special issue on Phylloxera." 50 :14-18.

Johnson, L.; Bosch, D.; Roczen, D. E.; Youkhana, S.K.; Nemani, R. R.(2003). "Mapping vineyard leaf area with multispectral satellite imagery. Computers and electronics in agriculture". 38:33-44

King, J. (1995) "Re-engineering Focus Slips," Computerworld, March 13, 1995.

Laudon, K. C., & Laudon, J. P. (2004). Sistemas de Informação Gerenciais (5ª ed.). São Paulo: Pearson Prentice Hall.

Lederer, A., Burky, L.B. (Outono de 1989). "Understanding Top Management's Objectives: A Management Information Systems Concern". Journal of Information Systems. Pág. 49 – 66.

Lefevers, J. (2011). http://www.forbes.com/sites/jefflefevere/2011/10/13/where-to-buy-wine-online-the-top-eight-wine-ecommerce-sites-for-unique-and-interesting-wines/ - Site Forbes

Luftman, J. (1996). Competing in the Information Age: Practical Applications of the Strategic Alignment Model, New York: Oxford University Press;

Luftman, Jerry (December 2000). "Assessing Business-IT Alignment Maturity", Communications of AIS, Volume 4, Article 14.

Maes, Rik; Rijsenbrij, Daan; Truijens, Onno; Groedvolk, Hans; Junho de 2000. "Redefining business – IT alignment through a unified framework". Universidade de Amsterdam. Amsterdam.

Malta, P. M., & Sousa, R. D. (2011). "The organizational competences model: a contribution for business-IT alignment", International Business Information Management Association (IBIMA).

Mayson, R. (2005). "Os Vinhos e Vinhas de Portugal". Publicações Europa-América

Neto, M. C., Pinheiro, A. C., & Coelho, J. C. (2004). "Gestão da Empresa Agrícola no Século XXI: Manual III – Tecnologias de Informação e Comunicação na Gestão da Empresa Agrícola".

Oliveira, T. H., Painho, M., Santos, V., Sian, O., Barriguinha, A. (2014), "Development of an agricultural management information system based on Open-Source solutions", CENTERIS 2014 / ProjMAN 2014 / HCIST 2014, www.elsevier.com/locate/procedia

Ortega, R.; Esser, A. (2002). "Viticultura de Precisión: Fundamentos, aplicaciones y oportunidades en Chile." Pontificia Universidad Católica de Chile. 1-10.

Pearson, Sammy Wray, "Measurement of Computer User Satisfaction" Ph.D. Dissertation, Arizona State University, August 1977.

Pires, S. (2008). "Deteção Remota para aplicações em Viticultura de Precisão no Alentejo". Revista da Associação Portuguesa de Enologia, N.os 51/52.

Porter, M. e V. E. Millar (1985), "How information gives you competitive advantage", Harvard Business Review, July – August, pp. 149-160.

Proffit, T.; Bramley, R.; Lamb, D.; Winter, E. (2006). "Precision Viticulture. A new era in vineyard management and wine production." Winetitles, Adelaide.

Santos, V.; Amaral, L.; Mamede, H.; Gonçalves, R. (2015). "Creativity in the Information Systems Planning Process". In Handbook of Research on Innovations in Information Retrieval, Analysis, and Management. IGI-Global. (aceite para publicação).

Venkatraman, N. and Vasudevan Ramanujam, "Measurement of Business Economic Performance: An Examination of Method Convergence", J. Management, 13, 1 (1987), 109 – 122.

Venkatraman, N. (1985). "Research on MIS Planning: Some guidelines from Strategic Planning Research". Journal of Management Information Systems, Vol. 2, No. 3, pp 65 – 77.

CONTRIBUTOS PARA A IMPLEMENTAÇÃO DE UMA FERRAMENTA DE *SERVICE DESK* NO INSTITUTO DA MOBILIDADE E DOS TRANSPORTES, I.P.

Carlos Carvalho[1], Henrique O'Neill[2]

1) Instituto da Mobilidade e dos Transportes, I. P.; Portugal
carlosdecarvalho@sapo.pt

2) ISCTE - Instituto Universitário de Lisboa; Portugal
henrique.oneill@iscte.pt

RESUMO

O presente estudo, realizado no Departamento de Apoio ao Utilizador (DAUT) do Instituto da Mobilidade e dos Transportes, I.P. (IMT), procurou analisar o processo de implementação de uma ferramenta de *service desk* em curso nesse departamento, com um triplo objetivo: (i) garantir o alinhamento da solução escolhida com as orientações do Governo quanto à implementação de novas tecnologias pela Administração Pública (AP); (ii) avaliar a adequação da solução escolhida às reais necessidades dos utilizadores e (iii) parametrizar um conjunto de métricas que permita disponibilizar informação relevante para o funcionamento do DAUT e do próprio IMT. Para tal, efetuou-se o enquadramento das mais recentes orientações para a adoção de TIC pela AP, bem como uma caracterização estatística das solicitações recebidas dos utilizadores. Desta análise resultou quer a proposta de um conjunto de indicadores a monitorar quer um conjunto de sugestões sobre como efetivamente prestar de um melhor serviço ao utilizador.

Palavras chave: Administração Pública, TIC, IMT, *Service Desk*, *Desing Thinking*

1 INTRODUÇÃO

O presente estudo foi realizado no Departamento de Apoio ao Utilizador (DAUT) do Instituto da Mobilidade e dos Transportes, I.P. (IMT), onde está integrado na Direção de Serviços de Sistemas de Informação (DSSI). Entre outras atribuições, cabe ao IMT licenciar o acesso de empresas e pessoas às atividades de transporte terrestre, habilitar os condutores de veículos rodoviários, certificar os profissionais dos transportes terrestres, homologar os veículos rodoviários, atribuir e cancelar matrículas ou promover a realização de inspeções técnicas. Compete ao DAUT assegurar o serviço de apoio aos utilizadores internos e externos do IMT, seja no âmbito das TIC seja numa perspetiva de apoio aos serviços disponibilizados aos utentes.

O processo de reestruturação do atendimento ao utilizador em curso no IMT resultou da decisão de implementar um serviço de *service desk* que, integrando o canal de *e-mail*, permita (i) criar um ponto único de contacto com os utilizadores, (ii) registar todas as solicitações que exijam resposta por parte do IMT e (iii) responder adequada e atempadamente, de acordo com os níveis de serviço acordados [Branco *et al*. 2011]. Deste processo resultou a escolha do *software* a adotar para operacionalizar o *service desk* de apoio via *e-mail* aos utilizadores internos e externos, tendo sido contratada a sua instalação e parametrização. O *software* escolhido – OTRS *Help Desk*, desenvolvido pela companhia *Open Technologies Real Services* (OTRS) – consiste numa solução *open source* e *on premise* que permite gerir consultas, reclamações, pedidos de suporte, relatórios de defeitos e outras comunicações. O OTRS *Help Desk* segue o modelo de boas práticas ITIL e tem uma estrutura modular que permite a instalação de *add-ons* (alguns gratuitos, outros pagos) segundo as necessidades do cliente [OTRS 2013].

Atendendo ao tempo decorrido desde a escolha do *software*, e admitindo-se para breve a sua entrada em funcionamento, ganhou relevância a necessidade de (i) garantir o alinhamento da solução escolhida com as posteriores orientações do Governo quanto à implementação de novas tecnologias pela Administração Pública (AP); (ii) avaliar a adequação da solução escolhida às reais necessidades dos utilizadores e (iii) parametrizar um conjunto de métricas que permita disponibilizar informação relevante para o funcionamento do DAUT e do próprio IMT.

O presente estudo procura dar resposta a estas necessidades. Importa contudo salientar algumas limitações quanto ao seu âmbito: (i) não é vinculativo, pelo que o seu impacto no projeto de implementação do *software* poderá ficar aquém do proposto (ii) a recolha de dados não poderia afetar o funcionamento do apoio aos utilizadores; (iii) a necessidade de obter contributos em tempo útil limitou o período de elaboração do estudo; e (iv) a escolha do *software* foi tomada como um dado adquirido, focando-se o estudo na sua implementação.

2 IMPLEMENTAÇÃO DE NOVAS TECNOLOGIAS PELA AP: O DESIGN THINKING (DT) COMO FORMA DE ABORDAGEM À MUDANÇA

Apesar da opção de não investir na modernização dos serviços públicos através das TIC poder parecer a mais adequada em tempos de *stress* fiscal, a longo prazo o progresso sustentável exige um investimento continuado em inovação, para que não se consumam recursos públicos em modos de atuação menos eficientes, eficazes ou económicos do que os implementados no setor privado ou na sociedade civil [Dunleavy 2010]. Contudo, este investimento leva frequentemente os governos a promover uma visão idílica da utilização das TIC [OECD 2009], eventualmente com o intuito de melhorar a sua aceitação por parte da opinião pública. Este tipo de discurso, apesar de permitir desenvolver formas inovadoras de pensar e de trabalhar, pode originar um distanciamento face à realidade, turvando o discernimento entre o que pode ou não ser realisticamente esperado e gerando muitas das vezes clivagens entre as metas ambiciosas e a débil implementação dos programas [Bekkers e Homburg 2009; Bekkers 2011].

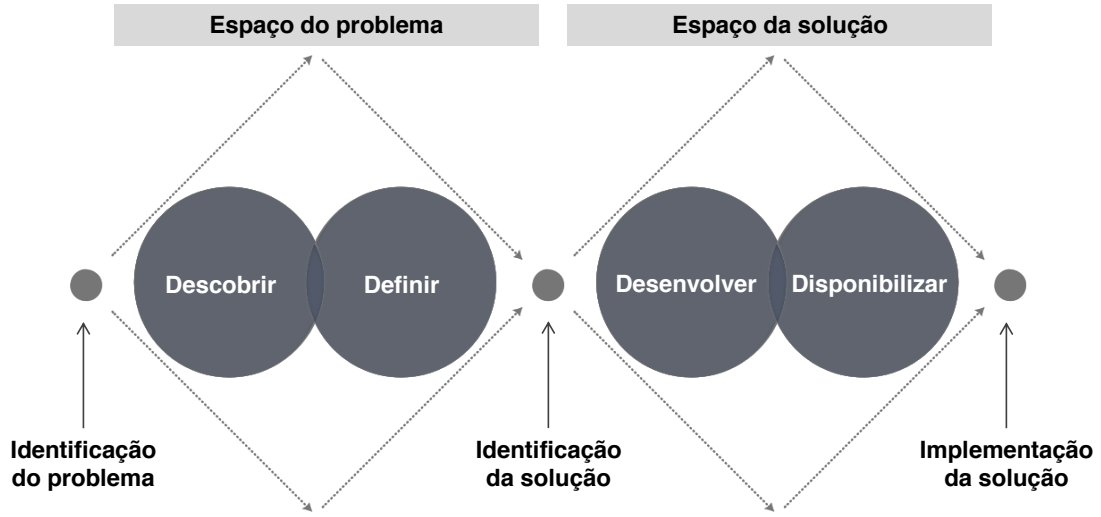
A OCDE registou níveis de utilização dos serviços de *e-government* pelos cidadãos abaixo do esperado face à qualidade e quantidade da oferta disponível [OECD 2009]. Esta discrepância reforça a necessidade de desenvolver soluções centradas no cidadão, não devendo o ponto de vista dos utilizadores ser afastado das fases de conceção, desenvolvimento e implementação destes serviços. Afinal, um serviço difícil de usar é um serviço que não será usado, pelo que o ignorar dos utilizadores obrigará a investimentos futuros na sua correção [Bertot e Jaeger 2008].

Esta atenção ao utilizador tem vindo a ser defendida pelo movimento do DT, que propõe a aplicação de práticas e princípios do *design* ao desenvolvimento e melhoria dos produtos e serviços produzidos ou disponibilizados por uma organização, procurando encontrar soluções simultaneamente exequíveis (perspetiva tecnológica ou organizacional), viáveis (perspetiva de negócio) e desejáveis (perspetiva do cliente) [Brown 2008]. O DT é encarado como um método para obter vantagens competitivas a partir da inovação, através do desenvolvimento de produtos e serviços que vão ao encontro das necessidades do mercado. Gozando de maior divulgação no setor privado, o DT tem também vindo a despertar o interesse do setor público [Young 2010], embora aqui o foco passe a ser o cidadão. O DT caracteriza-se assim por colocar quer o utilizador (cliente, cidadão) quer a organização (empresa, entidade pública) no centro do processo produtivo e de desenvolvimento [Whicher *et al*. 2013].

Apesar de haver inúmeras definições do que é o DT, há uma série de ideias que lhes são transversais: (i) procura criar produtos e serviços que sejam úteis, utilizáveis, desejáveis, eficientes e eficazes; (ii) centra-se na pessoa, focando-se nas suas necessidades, experiências e expectativas logo a partir das fases iniciais do desenvolvimento de novos produtos e serviços;

(iii) defende uma abordagem holística aos problemas, procurando resolvê-los através de uma visão ampla e integrada de estratégia, sistemas, processos, contexto e pontos de contacto com o utente; (iv) incentiva o recurso à pesquisa (técnicas qualitativas, etnográficas e observacionais) como forma de enfrentar os desafios de *design*; (v) incentiva a colaboração e a multidisciplinariedade, procurando envolver *stakeholders* e especialistas de várias áreas na procura da melhor solução para os problemas; (vi) é iterativo, recorre à prototipagem e procura a recolha rápida de *feedback* dos utilizadores de modo a avaliar e fazer evoluir as múltiplas ideias intermédias rumo à solução desejada [Young 2010; Whicher *et al*. 2013].

É possível encontrar na literatura inúmeros modelos de DT, propostos por diversos autores e instituições. As fases em que se encontram organizados estes modelos devem ser encaradas mais como um sistema de espaços do que como uma série predefinida de passos ordenados. Estes espaços demarcam diferentes tipos de atividades relacionadas que, em conjunto, formam o *continuum* da inovação [Brown 2008].



Fonte: Adaptado de McNabola *et al*. [2013]

Figura 1 – Descrição do modelo de *double diamond*

O *Design Council* propõe um modelo de DT apelidado de *double diamond* (Figura 1). Neste modelo começa-se por recolher uma multitude de ideias (pensamento divergente) para, em seguida, eleger a(s) melhor(es) (pensamento convergente). Estes dois tipos de pensamento podem ser graficamente representados como formando um losango (ou *diamante*). O modelo prevê a existência de dois *diamantes*: no primeiro examinam-se várias maneiras de olhar para um problema (espaço do problema), no segundo procura-se implementar a melhor forma de garantir a sua resolução (espaço da solução) [McNabola *et al*. 2013]. Este modelo é composto

por quatro fases: *descobrir*, *definir*, *desenvolver* e *disponibilizar*. Na fase *descobrir* procura-se "olhar o mundo" de modo a identificar as necessidades do utilizador e desenvolver algumas ideias iniciais. A fase *definir* consiste em analisar as possibilidades identificadas na fase anterior, de modo a enquadrar o desafio fundamental para a organização. A fase *desenvolver* marca um período onde as soluções são criadas, prototipadas e testadas, num processo iterativo de tentativa e erro. Finalmente, na fase *disponibilizar* o produto/serviço é finalizado e implementado [Design Council *s.a.*].

3 DESCRIÇÃO DO ESTUDO

3.1 ENQUADRAMENTO COM O MODELO DE DT APRESENTADO

Face ao modelo exposto, é possível alinhar cada um dos objetivos deste estudo com cada um dos espaços apresentados. Quanto ao alinhamento do *software* escolhido com as orientações do Governo, este é genericamente enquadrável no espaço do problema. Com efeito, não seria possível implementar uma solução que contrariasse, no essencial, essas orientações, pelo que teria de ser descartada. Já a adequação da solução escolhida às necessidades dos utilizadores pode ser enquadrada quer no espaço do problema quer no da solução. Sendo o estudo das reais necessidades dos utilizadores uma parte essencial do processo de parametrização da solução escolhida (espaço da solução), a constatação de uma desadequação flagrante a essas necessidades acabaria por inviabilizar a implementação da solução escolhida (espaço do problema). Finalmente, a parametrização de um conjunto de métricas relevantes insere-se claramente no espaço da solução, sendo a recolha sistematizada desta informação uma das mais-valias da implementação deste *software*.

Apesar deste enquadramento, há um conjunto de diligências adicionais indispensáveis para que o processo de implementação de *software* esteja, no essencial, de acordo com os ditames do DT. Há, por exemplo, de promover uma maior e muito mais precoce interação com os utilizadores antes da escolha do *software*, bem como a sua participação ativa no desenvolvimento, na parametrização e implementação das soluções a disponibilizar.

3.2 ALINHAMENTO DO *SOFTWARE* ESCOLHIDO COM AS ORIENTAÇÕES DO GOVERNO

Em Portugal, os vários governos têm vindo a promover a adoção de novas tecnologias pela AP, visando simplificar, agilizar, desburocratizar e desmaterializar processos; promover a racionalização, eficiência e inovação organizacionais; incentivar a abertura da AP e a participação dos cidadãos; melhorar a qualidade e a proximidade do atendimento e aumentar a qualificação dos colaboradores [Fonseca e Carapeto 2009]. Estes objetivos têm levado os vários

governos a condicionar a liberdade de contratação da AP, obrigada que está a respeitar as orientações do Governo quanto ao tipo de soluções tecnológicas a adotar. Esta liberdade foi também restringida por motivos de natureza orçamental, especialmente relevantes no contexto da vigência do Programa de Assistência Económica e Financeira. Para além das condicionantes genericamente aplicáveis a todas as contratações, têm vindo a ser implementadas algumas alterações ao quadro legislativo especificamente incidentes sobre a contratação de TIC.

Centrando a atenção no período de vigência do XIX Governo Constitucional, é possível identificar duas grandes orientações quanto à adoção de TIC pela AP. Uma primeira orientação consiste na tentativa de redução de custos com as TIC. Esta orientação foi operacionalizada através de várias medidas, entre as quais se destacam (i) o dificultar da contratação de TIC por ajuste direto, deixando o Código dos Contratos Públicos (CCP) de considerar como contratação excluída a aquisição de serviços informáticos de desenvolvimento de *software* e de manutenção ou assistência técnica de equipamentos; (ii) o incentivo à adoção de *software* aberto nos sistemas de informação do Estado; (iii) a necessidade de, em regra, obter uma avaliação positiva prévia da Agência para a Modernização Administrativa (AMA); (iv) a necessidade de verificar, no catálogo de aplicações já adquiridas pelo Estado, a inexistência de solução semelhante à que se pretende contratar, promovendo assim a homogeneização das soluções informáticas utilizadas pelas várias entidades da AP; e (v) a centralizando a função informática (*help desk*, licenciamento, microinformática), em cada ministério, das várias entidades sob a sua tutela (no caso do IMT, essas funções foram sendo progressivamente assumidas pela Secretaria-Geral do Ministério da Economia).

Uma segunda orientação consistiu no reforço da importância da prestação digital de serviços públicos, que passou a ser tida como regra, reforçando-se o seu valor legal e promovendo-se a maior interação dos *backoffices* das várias entidades da AP no serviço ao cidadão. Decorreu também desta orientação a criação de uma rede capilar de Espaços do Cidadão, visando apoiar e complementar a interação eletrónica entre o cidadão e a AP. No caso do IMT, a participação nestes Espaços do Cidadão, em colaboração com a AMA, traduziu-se na disponibilização de vários serviços relacionados com a carta de condução em centenas de lojas dos CTT e de instalações municipais espalhadas por todo o país.

Perante o exposto, constata-se que o *software* escolhido respeita genericamente as orientações anteriormente descritas. Com efeito, a sua instalação não implica qualquer custo para o IMT relacionado com a aquisição de licenças ou de *hardware* (atendendo ao parque informático disponível no IMT), acarretando porém alguns custos relacionados com a contratação de uma

empresa (nacional) para acompanhar as fases de instalação, parametrização e formação inicial, bem como com a eventual celebração de um contrato de manutenção.

A centralização na Secretaria-Geral do Ministério da Economia das funções de *help desk* relacionadas com o apoio informático ao utilizador interno do IMT veio retirar pertinência à utilização deste *software* como ferramenta de suporte a essa função. Pelo contrário, a participação do IMT nos Espaços do Cidadão veio criar uma nova oportunidade para a sua utilização, uma vez que compete ao IMT assegurar, em *backoffice*, o suporte a estes Espaços.

3.3 ADEQUAÇÃO ÀS NECESSIDADES DOS UTILIZADORES

As limitações do modo de funcionamento do apoio ao utilizador externo dificultam a caracterização das solicitações apresentadas pelos cidadãos. Estas dificuldades não são fáceis de ultrapassar utilizando apenas os meios disponíveis, e nem todas foram ultrapassadas neste estudo, nomeadamente as relacionadas com a caracterização quantitativa das mensagens recebidas. Procurou-se, em alternativa, proceder à sua caracterização qualitativa.

A recolha de amostras foi realizada durante o período laboral (dias úteis, entre as 9 e as 20 horas), entre os dias 17 de outubro e 17 de maio, com exceção do período entre os dias 13 e 30 de dezembro. Foram aleatoriamente definidos períodos de 5 minutos (contagem iniciada com a receção da primeira mensagem) em que foram recolhidas todas as mensagens entradas. Realizaram-se no total 131 recolhas de amostras, num total de 422 mensagens recolhidas. Atendendo à dimensão da amostra, presumindo a sua aleatoriedade e considerando uma população de dimensão desconhecida, para um nível de confiança de 95% obtém-se uma margem de erro de 4,8% [Carvalho 2014].

Quanto ao tipo de intervenção requerida, há três ações básicas que o operador do *service desk* pode executar: arquivar/apagar, responder ou reencaminhar. Tendo em conta estas ações, bem como a organização estratificada de um serviço de *service desk*, apresenta-se na Figura 2 a distribuição percentual das mensagens recebidas por tipo de intervenção efetuada. Como se pode constatar, mais de 60% das mensagens são respondidas pelos operadores de primeiro nível, sendo a intervenção de operadores de níveis superiores exigida em 22,7% dos casos.

As mensagens podem ser também agregadas por tipo de intervenção requerida, ou seja, não pela forma como atualmente são processadas mas pela forma como o seriam num cenário ideal (Figura 3). Atendendo à sua natureza, verifica-se que 46,7% das mensagens seriam, à partida, potencialmente evitáveis (lixo, *spam*, informação que o utilizador poderia obter sem recorrer ao *service desk*, etc.). As restantes dividem-se entre aquelas em que a intervenção do operador de

primeiro nível é indispensável e suficiente (30.6%) e aquelas em que é imprescindível recorrer a operadores de níveis superiores (22,7%).

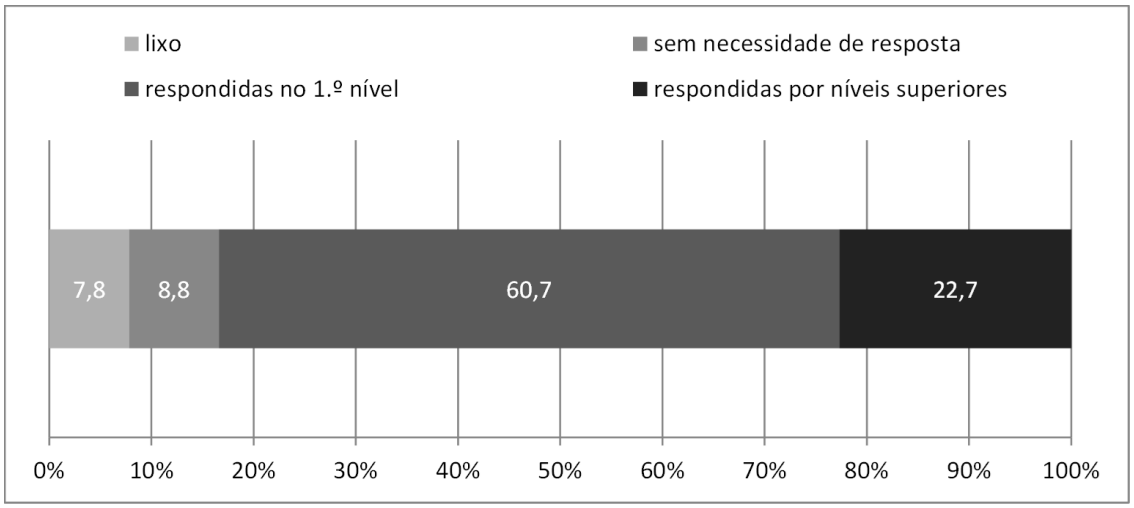


Figura 2 - Tipos de intervenção efetuada

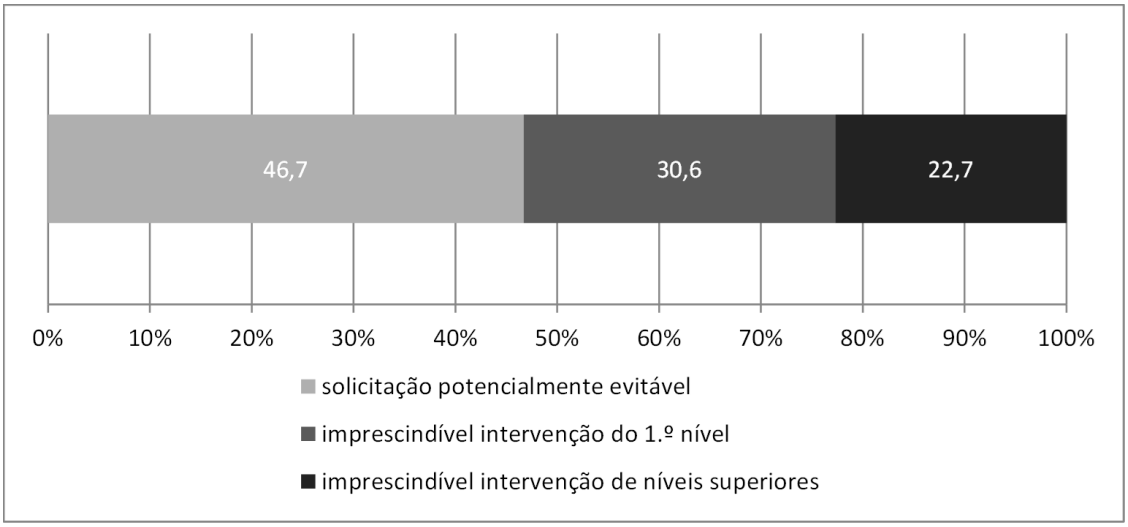


Figura 3 - Tipos de intervenção requerida

Verificou-se ainda que 17,5% das mensagens recebidas dizem respeito a pedidos de *verificação do estado de emissão da carta de condução* (representando 60% dos pedidos que podem ser resolvidos através da intervenção do operador de primeiro nível), o que indicia a existência de problemas com este procedimento, que, uma vez resolvidos, permitirão evitar a receção de grande parte destas mensagens. Esta é uma das principais vantagens da análise dos dados do funcionamento do *service desk*: a identificação precoce de problemas nos serviços prestados pelo IMT e o fornecimento de pistas para a sua eventual correção.

3.4 DEFINIÇÃO DE MÉTRICAS PARA O SERVIÇO DE APOIO AO UTILIZADOR EXTERNO

É possível, com a implementação da aplicação escolhida, definir e monitorizar uma série de métricas que ajudem a avaliar o desempenho do serviço de apoio ao utilizador externo. Só que, mais do que a sua quantidade, interessa avaliar a sua relevância para organização. Daí que se deva procurar identificar um número relativamente restrito de indicadores relevantes (KPI) que ajudem a avaliar o que é verdadeiramente importante para o seu desempenho. Estes indicadores devem ter correspondência com os fatores críticos de sucesso (FCS) do serviço de apoio. Segundo O'Neill [2005], os FCS são "um conjunto limitado de aspetos internos da organização que condicionam o seu desempenho efetivo e que devem ser melhorados", só devendo ser identificados depois dos objetivos estratégicos ficarem bem delineados.

Objetivo	FCS	KPI
Conter custos de funcionamento	Melhorar a eficiência do atendimento	custo médio dos pedidos fechados (€)
		taxa de pedidos fechados no primeiro nível (%)
	Diminuir custos do *software* utilizado	variação anual de custos (%)
	Diminuir pedidos desnecessários	taxa de pedidos desnecessários (%)
Melhorar o apoio ao utente	Aumentar satisfação do utente	grau de satisfação do utente (escala 1-5)
	Aumentar nível de resposta	taxa de pedidos respondidos (%)
	Aumentar rapidez de resposta	tempo médio de resposta (horas)
		taxa de pedidos não respondidos em 72 horas (%)
	Melhorar a qualidade das respostas	taxa de reabertura de pedidos (%)
		taxa de respostas incorretas (%)
		n.º de incumprimentos do nível de serviço
	Aumentar partilha de conhecimentos	n.º de entradas na *knowledge base*
		avaliação média das entradas (escala 1-5)
		n.º visualizações/*downloads* das FAQ
	Aumentar padronização de processos	taxa de pedidos fechados por processos padronizados (%)
		taxa de pedidos mal encaminhados (%)
	Melhorar tipificação de pedidos	taxa de pedidos tipificados pelo utente (%)
		taxa de pedidos bem tipificados pelo utente (%)
	Promover formação dos operadores	tempo de formação por operador (hora/operador)
		avaliação das ações de formação (escala 1-5)

Tabela 1 – Objetivos, FCS e KPI identificados

Perante o exposto, identificam-se dois objetivos estratégicos para o serviço de apoio ao utilizador externo: *conter os custos de funcionamento do serviço* e *melhorar o apoio prestado ao utente*. Identificam-se também onze FCS, relacionados com os objetivos de acordo com o exposto na Tabela 1, bem como 20 KPI que permitem operacionalizar cada um desses FCS.

A contenção dos custos de funcionamento deve passar pela melhoria de eficiência do processamento dos pedidos, pela diminuição dos custos relacionados com a posse e manutenção das aplicações utilizadas e pela diminuição dos pedidos resolúveis sem a intervenção do serviço de apoio. A melhoria do apoio ao utente deverá passar por aferir a sua satisfação, pelo aumento do número de pedidos respondidos, pelo aumento da rapidez e da qualidade das respostas dadas, pela promoção da partilha de conhecimentos (dentro do serviço de apoio e com o utente), pela maior implementação de processos padronizados, pela promoção de uma melhor tipificação dos pedidos pelos utilizadores e pelo maior investimento na formação dos operadores.

Estes FCS são operacionalizados de acordo com os KPI propostos (fórmulas de cálculo em anexo). Ainda que alguns destes indicadores impliquem a recolha de dados junto de outros departamentos do IMT ou dos utentes, a maioria poderá ser calculada através da aplicação escolhida, seja automaticamente (como, por exemplo, as taxas de pedidos respondidos) seja com posterior intervenção de um operador (como, por exemplo, a taxa de respostas incorretas, a calcular por amostragem). Fica assim patente a importância que esta aplicação poderá ter para uma avaliação adequada do desempenho do serviço de apoio ao utilizador externo.

4 CONCLUSÃO

Este estudo permitiu constatar que a ferramenta de *service desk* a adotar pelo IMT cumpre genericamente as especificações definidas pelo Instituto e pelo Governo no que concerne à adoção de novas ferramentas TIC pela AP. Contudo, a sua utilidade não dependerá tanto da sua legalidade mas da forma como efetivamente vier a ser implementada pelo IMT.

A existência de um grande número de solicitações potencialmente evitável deve levar o IMT a encarar a adoção de uma ferramenta de *service desk* como uma oportunidade para promover uma maior integração entre os vários canais de comunicação com os cidadãos (potenciando sinergias, por exemplo, com a informação disponibilizada no *site* do IMT), cuja melhoria seria idealmente prioritária em relação à entrada em funcionamento do *software* escolhido.

Por outro lado, a crescente participação do IMT na rede de Espaços do Cidadão abriu uma nova janela de oportunidade para a utilização do *software* escolhido. A importância destes Espaços para o Governo e para o IMT, a maior especificidade (e menor dispersão) dos assuntos aí abordados e o menor número de utilizadores poderão levar a que a utilização deste *software* no apoio a estes Espaços se assuma como prioritária face à sua utilização no apoio ao utilizador externo. Esta decisão implica contudo um estudo das solicitações dos utilizadores do *software*, devendo estes ser envolvidos na sua operacionalização. Daqui deverá resultar um novo conjunto de KPI especificamente definidos em função destes utilizadores.

No que concerne ao utilizador externo, os KPI que vierem a ser adotados deverão dar origem à elaboração de relatórios periódicos sobre o funcionamento do serviço de apoio ao utilizador, permitindo assim transformar os dados sobre o funcionamento do *service desk* em informação relevante para toda a organização, bem como sustentar a tomada de decisões quanto aos recursos a alocar à sua atividade. Assim, poderá o DAUT funcionar como um "alarme" que desperte as áreas de negócio do IMT para a existência de problemas no seu campo de atuação (como se constatou neste estudo com a emissão de cartas de condução). Para que tal aconteça, deverá o DAUT promover a disseminação destes relatórios periódicos por todo o IMT.

Apesar das suas limitações, este estudo permitiu evidenciar a utilidade de adotar uma perspetiva consentânea com os princípios do DT na implementação de uma ferramenta de *service desk*, promovendo a importância do utilizador em todas as fases do processo de mudança de modo a não se perca de vista o que é fundamental (a satisfação das necessidades do utilizador) em prol do que, sendo importante, é apenas instrumental (o funcionamento do *software*). Assim, é pretensão deste estudo poder servir para catalisar um debate que, a propósito da implementação de um *software*, contribua efetivamente para melhorar o serviço que o IMT presta ao cidadão.

5 REFERÊNCIAS

Bekkers, V., "Three faces of e-government: innovation, interaction, and governance", *in* A. V. Anttiroiko *et al* (eds.), *Innovations in Public Governance*, IOS Press, Amsterdão,2011.

Bekkers, V. e Homburg V., "The myths and ceremonies of e-government: beyond the hype of a new and better government?", *in* A. Meijer *et al* (eds.), *ICTs, citizens and governance: after the hype!*, IOS Press, Amsterdão, 2009.

Bertot, J. C. e Jaeger, P. T., "The e-government paradox: better customer service doesn't necessarily cost less", *Government Information Quarterly*, 25 (2008), 149–154.

Branco, A., Mendes, L. e Falcão, M. J., *Serviço de suporte ao utilizador (SSU)*, Lisboa, IMTT, 2011.

Brown, T, "Design thinking", *Harvard Business Review*, (junho 2008), 84-92.

Carvalho, C., *Reestruturação do atendimento ao cidadão na administração pública – o caso da implementação de uma ferramenta de service desk no Instituto da Mobilidade e dos Transportes, I.P.*, Dissertação de Mestrado em Administração Pública, ISCTE, Lisboa, 2014.

Design Council, *Design Methods for Developing Services*, s.a., (*online*), consult. 17-05-2015. Disponível em: http://www.designcouncil.org.uk/sites/default/files/asset/document/Design%20methods%20for%20developing%20services.pdf

Dunleavy, P., *The future of joined-up public services*, 2020 Public Services Trust, Londres, 2010.

Fonseca, F. e Carapeto, C., *Governação, inovação e tecnologias: o estado rede e a administração pública do futuro*, Edições Sílabo, Lisboa, 2009.

McNabola, A., Moseley, J. Reed, B., Bisgaard, T. Jossiasen, A. D., Melander, C., Whicher, A., Hytönen, J. e Schultz, O., *Design for Public Good,* SEE Platform, *s.l*., 2013.

O'Neill, H., "Planeamento estratégico e arquitectura de sistemas de informação", *in* A. Casaca (coord.), *A sociedade da informação e a administração pública*, INA, Oeiras, 2005.

OECD, *Rethinking e-government services: user-centred approaches*, OECD Publications Service, Paris, 2009.

OTRS, *OTRS 3.3 - Admin Manual*, 2013, (*online*), consult. 17-05-2015. Disponível em: http:// ftp.otrs.org/ pub/otrs/ doc/doc-admin/3.3/en/pdf/otrs_admin_book.pdf

Whicher, A., Swiatek, P. e Cawood, G., *An overview of service design for the private and public sectors*, SEE Platform, *s.l*., 2013.

Young, G., *Design thinking and sustainability*, 2010, (online), consult. 17-05-2015. Disponível em: http://zum.io/wp-content/uploads/2010/06/Design-thinking-and-sustainability.pdf

ANEXO – FÓRMULAS DE CÁLCULO DOS KPI PROPOSTOS

KPI	Fórmula
Custo médio dos pedidos fechados (€)	$\frac{\sum_1^3(tempo\ trabalho\ nível\ n\ \times custo\ médio\ trabalho\ nível\ n)}{n.º\ pedidos\ fechados}$
Taxa de pedidos fechados no 1.º nível (%)	$\frac{n.º\ pedidos\ fechados\ no\ 1.º\ nível}{n.º\ pedidos\ fechados} \times 100$
Variação anual de custos (%)	$\frac{custo\ software\ (ano\ n) - custo\ software\ (ano\ n-1)}{custo\ software\ (ano\ n-1)} \times 100$
Taxa de pedidos desnecessários (%)	$\frac{n.º\ pedidos\ desnecessários}{n.º\ pedidos} \times 100$
Taxa de pedidos respondidos (%)	$\frac{n.º\ pedidos\ respondidos}{n.º\ pedidos} \times 100$
Tempo médio de resposta (horas)	$\frac{tempo\ necessário\ para\ fechar\ pedidos}{n.º\ pedidos\ fechados} \times 100$
Taxa de pedidos não respondidos em 72 horas (%)	$\frac{n.º\ pedidos\ não\ respondidos\ em\ 72\ horas}{n.º\ pedidos} \times 100$
Taxa de reabertura de pedidos (%)	$\frac{n.º\ pedidos\ reabertos\ pelo\ utente}{n.º\ pedidos\ fechados} \times 100$
Taxa de respostas incorretas (%)	$\frac{n.º\ pedidos\ respondidos\ incorretamente}{n.º\ pedidos\ fechados} \times 100$
Taxa de pedidos fechados por processos padronizados (%)	$\frac{n.º\ pedidos\ fechados\ por\ processos\ padronizados}{n.º\ pedidos\ fechados} \times 100$
Taxa de pedidos mal encaminhados (%)	$\frac{n.º\ pedidos\ mal\ encaminhados}{n.º\ pedidos} \times 100$
Taxa de pedidos tipificados pelo utente (%)	$\frac{n.º\ pedidos\ tipificados\ pelo\ utente}{n.º\ pedidos} \times 100$
Taxa de pedidos bem tipificados pelo utente (%)	$\frac{n.º\ pedidos\ bem\ tipificados\ pelo\ utente}{n.º\ pedidos} \times 100$

Does IS/IT Investments Impact on Firm Financial Performance: "Stakeholder" *versus* "Shareholder" Orientation

António Guerreiro [1]

1) University of Évora, Évora, Portugal and CEFAGE

ahmg@uevora.pt

Abstract

This research project addresses a central question in the IS business value field: Does IS/IT investments impact positively on firm financial performance?

IS/IT investments are seen has having an enormous potential impact on the competitive position of the firm, on its performance, and demand an active and motivated participation of several stakeholder groups. Actual research conducted in the Information Systems field, relating IS/IT investments with firm performance use transactions costs economics and resource-based view of the firm to try to explain and understand that relationship. However, it lacks to stress the importance of stakeholder management, as a moderator variable in that relationship. Stakeholder theory in its instrumental version, argues that if a firm pays attention to the stakes of all stakeholder groups (and not just shareholders), it will obtain higher levels of financial performance.

With this premise in mind, the aim of this research project is to discuss and test the use of stakeholder theory in the IS business value stream of research, in order to achieve a better understanding of the impact of IS/IT investments on firm performance (moderated by stakeholder management).

Keywords: IS/IT Investments, Impacts, Financial Performance, Stakeholder Orientation.

1. Introduction

The introduction of IS/IT in organizations is likely to have a significant impact within the organization. IS/IT can be used in restructuring organizational activity, in strengthening the competitive position of the firm [Ward & Peppard, 2002], and to transform entire business processes [Al-Mudimigh et al 2001; Brynjolfsson & Hitt, 1998].

In the 1980s IS/IT was herald as a key to competitive advantage [McFarlan, 1984; Porter & Millar, 1985]. Porter and Millar [1985] concluded that IS/IT has affected competition in three ways: it has led to changes in industry structure and competition, it was used to support the creation of new business and companies using IT outperformed their competition. Earl [1989] suggests that IS/IT has the potential to be a strategic weapon.

Despite increasing expenditure on IS/IT [Ballantine & Stray, 1999; Ryan & Gates 2004, Willcocks & Lester 1999] and the belief that IT has a significant impact on organizational performance [Osey-Bryson & Ko, 2004], the effect of such investments on firm productivity has been unclear [Dasgupta.et al., 1999; Farbey et al. 1999] and has given rise to a 'productivity paradox' [Love & Irani, 2004]. Many organizations find themselves in a "Catch 22"[1], for competitive reasons they cannot afford not to invest in IS/IT, but economically they cannot find sufficient justification for it [Willcocks 1992].

During the past four decades a great deal of attention has focused on the impact of IT investment. However studies have frequently generated controversial or inconsistent results [Kivijärvi & Saarinen, 1995].

After revising the literature in the IS business value field, where the weak use of theory is pointed as a major gap in the field, namely in the first years in which the phenomena has been studied, and one of the responsibles for the contradictory findings, Transaction Cost Theory (TCT) and Resource Based Theory (RBT) are presented and their use in the IS business field is reported.

However those theories say nothing on the relation with several stakeholder groups who interact daily with the firm, and from which the success of the IS/IT depends. Stakeholder theory is introduced in the next part as a possible good theory candidate to moderate and help to shed light on the relation between IS/IT investments and firm performance. The paper ends with the presentation of the empirical model, results and conclusions.

2. Impact of IS/IT on Firm Performance

2.1. IS/IT investments, firm performance and the 'productivity paradox'

A growing body of research into the firm performance effects of IT investment has emerged and is sometimes referred to as IT business value research. The problem researchers face is identifying robust methods to gain insight into how IT business value is created [Kauffman & Weill 1989]. The crux of the problem is whether IT investment really makes a difference in firm performance. Prior researchers have reached contradictory conclusions when studying the relationship between IT investment and firm performance.

[1] See Joseph Heller's 1961 novel.

The shortfall of evidence concerning the productivity of IT became known as the 'productivity paradox' (Solow, 1987). As Robert Solow, the Nobel laureate economist state "we see computers everywhere except in the productivity statistics [in Brynjolfsson 1992, p.2).

The early studies tended to address the question of computer use [Lucas 1975] and the relationship between performance and computerization intensity [Cron & Sobol 1983]. The studies by PIMS [1984] and Bender [1986] measured the proportion of expenses dedicated to IT in firms, while Breshniham [1986] and Roach [1987] measured amounts of resources dedicated to IT in a sector.

Early work in the field is based on some notion of productivity drawn from accounting (which basically ignore the process by which inputs are converted into outputs) or on methodologies from economics. In this case the process that links inputs to outputs is modelled, but very simply using computed ratios of input to output transformation [Crowston & Treacy, 1986].

Simply empirical studies, without a strong theory-base, difficultly will reveal the heart of the IT pay-off question. In the view of Crowston & Treacy [1986] we must look for a strong theory about the process in organizations to guide our choice of variables and to generate testable hypothesis about them.

Once a theory base and methodology have been chosen and the unit of analysis has been decided upon to measure IT impact and its locus, the next logical step in the progression is to select a set of performance measures [Kauffman & Weill 1989]. With respect to performance measures, at firm level, we can find two sets of measures: accounting based measures (ROA, ROE, ROI, ROS) and market measures (as Tobin's q).

Transaction Cost Theory (TCT) and Resource Based Theory (RBT) are two widely used theories of the firm, that have started to be used in the IS field in general, and are also defended and used to study the impacts of IT on business. Next we will briefly describe both theories and their fundamental assumptions.

2.2. The Transaction Cost Theory and the Resource Based Theory

2.2.1. The Transaction Cost Theory

Coase [1937] refuted the idea from the economic theory of price-mechanism as the key to resource allocation within the firm, and call to a new theory of the firm, actually known as the TCT. As he states, "Outside the firm, price movements direct production, which is coordinated through a series of exchange transactions on the market. Within a firm, these market

transactions are eliminated and in place of complicated market structure with exchange transactions is substituted the entrepreneur-co-ordinator, who directs production" [Coase 1937, p.388]. In his view the main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism in the market place (information costs, market regulation costs, negotiating costs, contract costs, costs of monitoring the contracts). According Coase [1937] a firm becomes larger as additional transactions are organized by the entrepreneur (inside the firm) and becomes smaller as he abandons the organization of such transactions (and goes to the market). TCT intends to answer questions like "Why do firms exist" and "Why is there vertical integration?" [Demsetz 1988, p.151]. The firm is seen as a nexus of contracts [Demsetz 1988] or as a governance structure [Rindfleisch & Heide, 1997].

Over the past four decades, Williamson [1975, 1979, 1985, 1996] has added considerable precision to Coase's general argument by identifying the types of exchanges that are more appropriately conducted within the firm boundaries than within the market. Opportunism and bounded rationality are presented as the key behavioural assumptions on which transaction cost economics relies. He argues that opportunism is a central concept in the study of transaction costs and focuses his attention on the exchanges in which opportunistic potential is relevant.

In the IS field, the interest in TCT increased with the rising interest in studying the options of insourcing or outsourcing the IS/IT function [Grover et al., 1998; Willcocks & Lacity, 1998].

Crowston & Treacy [1986] state that Williamson's studies of markets and hierarchies can help to explain the enterprise and industry level impact of IT by explaining changes in production and transaction costs. He points out that the boundaries between industries arise at those points where a market's advantage of production efficiencies outweigh the transaction cost superiority of internal organization. Simply put, separate and specialized industries exist because at some points it is cheaper to buy a product or service in the market than to make it. IT has the potential to radically alter cost structures and transform the structure of industry boundaries. In some cases, functions that were once integrated into the firm may be eliminated and alternatives may be purchased in a market. In other cases, products and services that were once purchased now may be created by functions within the firm. IT can have this impact on industry structure by altering the relative production efficiencies and transaction costs of market and organization mechanisms, and the specificity of assets that create products.

Kauffman & Weill [1989] argue that the use of strong theory bases will improve the likelihood of achieving meaningful IT impact analysis results, future IS research should tap a broad range of applicable theories and methods, TCT is one of them.

TCT looks to firm as a 'nexus' of contracts and assumes 'opportunism' as a central issue, and lacks to address the importance of stakeholders (and their cooperative relations) to the prosperity and sustainability of the company.

2.2.2. The Resource Based Theory

The resource-based view argues that firms possess resources, a subset of which enables them to achieve competitive advantage. This theory focus on the idea of costly-to-copy attributes of the firm as sources of business returns and hence an essential way to achieve superior performance and competitive advantage [Barney 1991, Conner 1991, Conner & Prahalad, 1996; Prahalad & Hamel, 1990]. According the RBT, the firm looks for unique attributes that may provide superior performance. The firm is seen as a collection of productive resources.

According the RBT, competitive advantage occurs only when there is a situation of resource heterogeneity (different resources across firms) and resource immobility (the inability of competing firms to obtain resources from other firms) [Barney 1991]. The RBT treats companies as potential creators of value added capabilities. The development of such capabilities and competencies involves a knowledge-based perspective [Conner & Prahalad, 1996; Prahalad & Hamel, 1990].

Unlike TCT, a resource-based view of the firm does not depend on opportunistic behaviour. It focuses on developing internal knowledge and competencies to enable the firm to improve its competitiveness. It accepts that attributes related to past experiences, organizational culture and competencies are critical for the success of the firm [Hamel and Prahalad, 1996].

RBT has been widely used in the IS field [Mata et al., 1995; Caldeira 1998; Grover et al., 1998; Caldeira & Ward, 2003]. Mata et al. [1995] argue that managerial IT skills were an attribute of IT that can provide sustainable advantage (they are usually developed over long periods of time, through learning and experience). In the view of Grover et al. [1998], and according to resource-based theory, outsourcing is a strategic decision which can be used with the purpose of filling the gap between the desired IS/IT capabilities of the firm and the actual ones. Caldeira [1998] and Caldeira & Ward [2003] defend a resource-based approach to the understanding of IS/IT adoption and use in manufacturing SMEs.

The RBV has been proposed to investigate the impact of IT investments on firm performance [Santhanam & Hartono, 2003]. Researchers have shown that a firm's ability to effectively leverage its IT investments by developing a strong capability can result in improved firm performance. For instance, Bharadwaj [2000] provided evidence that firms with IT capability tend to outperform a control sample of firms on a variety of profit and cost-based performance

measures. Santhanam & Hartono [2003] indicate that firms with superior IT capability exhibit superior current and sustained firm performance when compared to average industry performance, even after adjusting for effects of prior firm performance.

RBT recognizes the importance of manage stakeholders, but puts it as one of the competences to transform the firm resources into capabilities, in order to achieve a competitive advantage.

3. Stakeholder Theory and IS/IT investments

The idea that corporations have stakeholders has become commonplace in the management literature, both academic and professional [Donaldson & Preston 1995]. The ruling paradigm of corporate governance holds that those who invest their capital into whatever kind of business, and, by that token, those who risk losing their investment in parts or in total, have an entitlement (and an obligation) to govern the business they have invested into.

Freeman's [1984] landmark work provided a solid and lasting foundation for many continuing efforts to define and to build stakeholder models, frameworks, and theories. According his work, strategic management of private sector firms could become much more effective and efficient, if managerial efforts regard various stakeholders' concerns. In other words, shareholders benefit long-term if other legitimate interests in the firm do not fall by the wayside.

Stakeholder theory establishes a framework for examining the connections, if any, between the practice of stakeholder management and the achievement of various corporate performance goals [Donaldson & Preston 1995]. The principal focus of interest here has been the proposition that corporations practicing stakeholder management will, other things being equal, be relatively successful in conventional performance terms (profitability, stability, growth, …). Instrumental uses of stakeholder theory make a connection between stakeholder approaches and commonly desired objectives such as profitability. Stakeholder management requires, as its key attribute, simultaneous attention to the legitimate interests of all appropriate stakeholders, both in the establishment of organizational structures and general policies and in case-by-case decision making.

In the information systems field, there is an extensive work about a wide range of issues such as IS/IT evaluation, design, implementation and management of IS/IT investments, using stakeholder theory.

However, the main focus is about the use of the "stakeholder" concept and with their identification. This research also stresses the importance of including stakeholders on several

tasks such as evaluation and IS/IT design in order to achieve the expected levels of performance.

The benefit of exploiting IS/IT not only relates to making business processes and tasks more efficient. Instead, IS/IT also enables the creation of products, services, distribution channels, and links with customers, suppliers, and other stakeholders. Remenyi [1999] defends that IT has no direct value in its own right. IT investment has a potential for derived value. More than any other factor of success or failure of the IT investment is a function of the skill and commitment of the information systems principal stakeholders. Only when IT is coupled with other resources, and especially the principal stakeholders, can any benefits or value be perceived. Any organization ultimately makes investments in IS/IT to create value for its stakeholders, whether they are shareholders, customers, employees or others with a vested interest in sharing in its success [Ward & Peppard, 2002]. The literature cites many examples of IS/IT projects in which multiple stakeholder groups are involved, with substantial influence. Farbey et al. [1999] found that external stakeholders could play a decisive and crucial role in many IS/IT investments.

To achieve the expected impact from an IT investment, we argue that firms need a strong commitment from these stakeholder groups, which lead us to the need of a stakeholder orientation.

According to ST main proposition, it is possible to put the following research questions:

(1) Does the IS/IT investment of firms practicing stakeholder management will, other things being equal, be relatively successful in terms of firm performance?

(2) Is there a relation between a firm's corporate social responsibility (CSR) and the impact of IS/IT investments on firm performance?

The proposed conceptual model put a new construct, 'stakeholder vs shareholder orientation', as a moderating variable in the relationship between IS/IT investments and firm performance (figure 1).

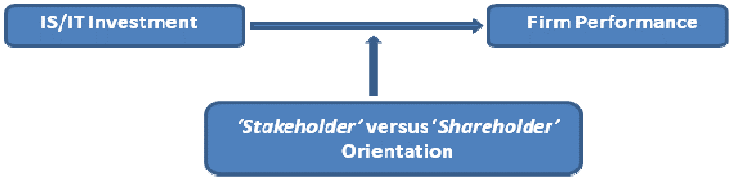


Figure 1 – IT investments and firm performance: A stakeholder approach.

Stakeholder Orientation has been assessed by stakeholder theorists in the strategic management field using KLD index (a corporate social performance index), which relies on public records of notable socially responsible activities [Berman et al., 1999), or the Dow Jones Sustainability

Index (DJSI) in USA, or the Footsie for Good Index (FTSE4Good) in the UK. When it is not possible to use those indexes (for example to use firm level data of other countries), researchers should develop efforts to identify firms with good practices of corporate social responsibility (CSR), and create a 'dummy' variable (1 for firms with a 'stakeholder orientation'; 0 for the others) as a good proxy of stakeholder management practices.

Firm performance should be assessed by accounting measures, such as return on assets (ROA), return on equity (ROE), return on investments (ROI), or return on sales (ROS), the world famous "language" of business and management, particularly when data sets include firms that are not present in the stock markets.

4. Data Description and Empirical Model

4.1. Data Description

This section provides a brief description of the data used in this paper. The Portuguese National Institute of Statistics (INE) runs annually two surveys to Portuguese companies, the Harmonized Firm Survey (IEH) which collects accounting data, and the Survey on the Use of Information and Communication Technologies (IUTIC) where we can find information about IT expenditure.

Both surveys are exhaustive for firms with more than 250 employees (all population of Portuguese firms is inquired), so we have requested data on those companies, for the years of 2004 and 2005 (2004 was the first time in which the question "how much your company spent in IT" appeared in the IUTIC survey).

The sample is constituted by large firms with more than 250 employees mainly from the private sector and has a total of 1186 observations (581 firms inquired in 2004 and 605 in 2005) from the sectors of extracting and manufacturing industry (sector C/D), electricity (sector E), construction (sector F), wholesale and retail trading and repair (sector G), Hotels and Restaurants (sector H) transport and communications (sector I), real estate and business service activities (sector K) and other collective, social and personal activities (sector O).

4.2. Variables and Model

The **dependent variable**, financial performance will be accessed by the most common financial indicators: return on equity (ROE), calculated by taking the net result over shareholders' equity for each specific year; Return on Assets (ROA), calculated by taking the EBIT over total (net) assets for each specific year; Return on Investment (ROI), calculated by taking the EBIT over

total investments for each specific year; and at last Return on Sales (ROS), calculated by taking the net result or EBIT over total sales for each specific year.

ROE represents what return the company is making on the shareholders' funds invested in the company.

According Brown et. al., 1995, ROE, ROA and ROI are all closely related and are widely accepted as profitability measures.

Return on Assets (ROA) measures the company's ability to generate profits from its assets, ignoring how they were financed [Stickney, 1990]. Return on investment (ROI) measures the company's ability to realize value from their investments. Return on sales (ROS) measures the net margin of the company on the turnover.

These profitability indicators are quite common being used by researchers in the field of information systems, to study of the impacts of investments in IS/IT on business performance [Kivijärvi & Saarinen, 1995; Rai et al., 1997; Li &Ye, 1999; Stratopoulos & Dehning 1999, 2000; Bharadwaj, 2000; Shin, 2001; Hitt et al.,2002; Lee & Boose, 2002; Byrd & Davidson, 2003; Santhanam & Hartono, 2003;Dimovsky & Skerlavaj, 2004; Zhu, 2004; Tanriverdi, 2005; Ravichandran & Lertwongsatien, 2005; Aral et al., 2006; Aral & Weill, 2007; Guerreiro & Serrano, 2007a, 2007b; Altinkemer et al., 2007; Dehning et al., 2007; Lee, 2008; Stoel & Muhanna, 2009; Ravichandran et al., 2009].

As **independent variable**, we used IS/IT investment. The IS/IT investment concept is operationalized in many different ways by different researchers. In this paper we use the concept of IS/IT investment which is asked to Portuguese companies in the IUTIC survey. This concept is closed to the concept defined by the MIT researchers Aral & Weill [2006: 23]: "total expenditures on IT (all computers, software, data communications, and people dedicated to providing IT services), including both internal and outsourced expenditures".

The Portuguese IUTIC survey provides us that data into two separate variables:

- $\mathbf{IT_{Assets}}$= All expenses in computers, software, and data communications dedicated to providing IT services;
- $\mathbf{IT_{HR}}$= Human Resources expenditure related to computers, software, and data communications dedicated to providing IT services;

The IS/IT investment variable will be the sum of both items.

In the model we divided these variables by total sales, in line with Aral & Weill [2006], with the aim of control for the relative production size of firms.

To assess the contribution of "stakeholder orientation" of the firm, a **moderate variable** (dummy) was introduced in the model. Field research was conducted to identify corporate strategy practices of companies to evaluate their stakeholder orientation, using Business Council for Sustainable Development Portugal (BCSD Portugal), Corporate Social Responsibility Portugal RSE Portugal) and GRACE Portugal (Group of Reflection and Support to Corporate Citizenship) member lists. All those non-profit organizations' mission is to develop CSR among Portuguese companies. Firms listed in these public list were considered as having stakeholder orientation.

As **control variables**, two firm level variables were introduced to control for their effects on performance, advertising expenditures and firm size [Aral & Weill, 2006]. According Montgomery & Wernerfelt [1988], advertising expenditures are positively related to firm performance. Firm size will be controlled by the natural logarithm (ln) of the nº of employees and advertising expenditures will be operationalized as the ratio that expenses to sales, to control for the relative production size of firms [Aral & Weill, 2006]. Also we will introduce p-1 control variables for the different sectors present in the sample (p=number of sectors).

At last, the model is introduced as follows (in line with Aral & Weill [2007] :

$$FP_i = \beta_0 + \beta_1\left(\frac{Total\ IT}{Sales}_i\right) + \beta_2 STO_i + \beta_3 STO_i\left(\frac{Total\ IT}{Sales}_i\right) + \beta_4 ln\left(NPS_i\right) + \beta_5\left(\frac{Advertising}{Sales}_i\right) + \beta_{ij} Sector_{ij} + \varepsilon_i, \quad \textbf{(1)}$$

where :
FP – financial performance (ROA, ROE, ROI, ROS)
ROA – return on assets
ROE – return on equity
ROI – return on investment
ROS – return on sales
$Total_{IT}$ - total IS/IT investments (euros)
Sales – turnover (euros)
STO – *dummy* variable "*stakeholder orientation*"
NPS – workforce (number of employees)
Advertising - Advertising expenses (euros).

The coefficient β_3 aims to measure the impact of the moderating effect of 'stakeholder orientation' on the relationship between spending on IS / IT and business performance.

According Aral & Weill [2007] the variable Total IT can be divided into two distinct components, namely IT_{Assets} and IT_{HumRes}:

$$FP_i = \beta_0 + \beta_1\left(\frac{TI_{Assets}}{Sales}_i\right) + \beta_2\left(\frac{TI_{Hum\,Res}}{Sales}_i\right) + \beta_3 STO_i + \beta_4\left(\frac{TI_{Assets}}{VN}_i\right)STO_i + \beta_5\left(\frac{TI_{Hum\,Res}}{Sales}_i\right)STO_i + \beta_6 ln\left(NPS_i\right) + \beta_7\left(\frac{Advertising}{Sales}_i\right) + \beta_{ij} Sector_{ij} + \varepsilon_i, \quad \textbf{(2)}$$

where:
FP – financial performance (ROA, ROE, ROI, ROS)
ROA – return on assets
ROE – return on equity
ROI – return on investment
ROS – return on sales
IT_{Assets} - total IS/IT assets expenditure (euros)
$IT_{Hum\,Res}$ - total IS/IT human resources expenditure (euros)
$Sales$ – turnover (euros)
STO – *dummy* variable "*stakeholder orientation*"
NPS – workforce (number of employees)
$Advertising$ - Advertising expenses (euros).

The β_4 and β_5 coefficients are intended to assess the impact of the moderator effect of 'stakeholder orientation' in the relationship between IS / IT expenditure and business performance.

The models presented are estimated by the method of the ordinary least squares (OLS). The Gauss-Markov base assumptions of this research methodology will be subject to analysis and in order to achieve efficient and consistent estimators may be possible to make corrections due to the model.

i) absence of autocorrelation of the residuals, analysis carried out by the Ljung-Box test;

ii) homogeneous variance analysis through the White and ARCH tests;

iii) the normality of the residuals analysed by Jarque-Bera test.

Multicolinearity will also be the subject of study by analysing the correlation map between variables.

5. Results and Discussion

The equations (1) and (2) were estimated, the first of which uses as an independent variable the total spending amount in IS/IT, and the second allocating these expenses between "IS/IT assets" and expenditures associated human resources to IS/IT.

For each equation were run 5 multiple regression by the method of least squares, as many as the variables that are used to measure the financial performance of the company (ROE, ROA, ROI, net ROS and operational ROS).

Total spending on IS/IT revealed a negative impact on financial performance, when measured by return on assets (ROA) and return on sales (ROS). When it breaks down spending on IS/IT, it appears that spending on IS/IT assets continue to have a negative relationship with the ROA and ROS, however the impact on the return on investment (ROI) is positive.

We also noted that the variable stakeholder orientation' (STO) moderates the relationship between IS/IT investment and performance, and in the case of ROA and ROS, the working hypothesis is validated, i.e., business strategies oriented to the satisfaction of all stakeholders of the company have a positive impact on the relationship between investments in IS/IT and financial performance.

Equation	Variables	ROE	ROA	ROI	ROS (Net)	ROS (Oper.)
(1)	Total IT	Model NS	(-) ***	Model NS	(-) ***	(-) ***
	'STO'		(+) *		NS	NS
	IT*'STO'		NS		NS	NS
(2)	IT_{Assets}	Model NS	(-) ***	(+) ***	(-) ***	(-) ***
	IT_{HumRes}		NS	NS	NS	NS
	'STO'		NS	NS	NS	NS
	IT_{Assets}*'STO'		(+) ***	(-) ***	(+) *	(+) ***
	IT_{HumRes}*'STO'		NS	NS	NS	(-) ***

*** (99%); ** (95%); * (90%); Model NS: Model without statistical significance; NS: The variable is not statistically significant.

Table 1 – Impact of IS/IT and the 'Stakeholder Orientation' on Firm Financial Performance.

Finally, we introduced a one year lag in the equations according IS literature (for. ex. Brynjolfsson 1992) in order to obtain more robust results.

$$FP_{(n+1)i} = \beta_0 + \beta_1\left(\frac{Total\ IT_{(n)}}{Sales}_i\right) + \beta_2 STO_i + + \beta_3 STO_i\left(\frac{Total\ IT_{(n)}}{Sales}_i\right) + \beta_4 ln\left(NPS_i\right) + \beta_5\left(\frac{Advertising_{(n)}}{Sales}_i\right) + \beta_6 FP_{(n)i} + \beta_{ij} Sector_{ij} + \varepsilon_i, \quad \textbf{(3)}$$

$$PF_{(n+1)i} = \beta_0 + \beta_1\left(\frac{IT_{Assets(n)}}{Sales}_i\right) + \beta_2\left(\frac{IT_{HumRes(n)}}{Sales}_i\right) + \beta_3 STO_i + \beta_4\left(\frac{IT_{Assets(n)}}{VN}_i\right)STO_i$$
$$+\beta_5\left(\frac{IT_{HumRes(n)}}{Sales}_i\right)STO_i + \beta_6 ln\left(NPS_i\right) + \beta_7\left(\frac{Advertising}{Sales}_i\right) + \beta_8 FP_{(n)i} + \beta_{ij} Sector_{ij} + \varepsilon_i, \quad \textbf{(4)}$$

The equations (3) and (4) were estimated, the first of which uses as an independent variable the total spending amount in IS/IT, and the second allocating these expenses between "IS/IT assets" and expenditures associated human resources to IS/IT. For each equation were run 5 multiple regression by the method of least squares, as many as the variables that are used to measure the financial performance of the company (ROE, ROA, ROI, net ROS and operational ROS).

In fact, the model that incorporates a "lag" of one year presents more robust results, which corroborates the thesis that there is an organizational learning process to obtain the benefits of this type of investment, that the IS/IT must be properly used, or that IS/IT expenditure must be converted into assets to generate value.

Total spending on IS/IT showed a negative impact on financial performance, when it is measured by return on assets (ROA) and return on sales (ROS). When spending on IS/IT is divided, it appears that spending on IS/IT assets continue to have a negative relationship with ROA and ROS, and on the return on investment (ROI).

We also noted that the variable 'stakeholder orientation' (STO) moderates the relationship between IS/IT investment, and in the case of ROA, ROI, and ROS, the working hypothesis is validated, i.e., the corporate governance model proposed by 'Stakeholder Theory' reveals a positive impact on the relationship between investments in IS/IT and financial performance.

Equation	Variables	ROE	ROA	ROI	ROS (Net)	ROS (Oper.)
(3)	Total IT	NS	(-) ***	NS	(-) ***	(-) ***
	'STO'	NS	NS	NS	NS	NS
	IT*'STO'	(-) ***	NS	NS	(+) ***	(+) ***
(4)	IT_{Assets}	NS	(-) ***	(-) ***	(-) ***	(-) ***
	IT_{HumRes}	NS	NS	NS	NS	(-) ***
	'STO'	NS	NS	NS	NS	NS
	IT_{Assets}*'STO'	(-) ***	(+) *	(+) *	(+) ***	(+) ***
	IT_{HumRes}*'STO'	NS	NS	NS	NS	NS

*** (99%); ** (95%); * (90%); Model NS: Model without statistical significance; NS: The variable is not statistically significant.

Table 2 – Impact of IS/IT and the 'Stakeholder Orientation' on Firm Financial Performance (one year lag model).

6. Conclusions

The impact of information systems/information technologies (IS/IT) investments on firm financial performance continues to be a source of heated discussion and debate.

We found a negative direct relation between IT expenditure and financial performance. Our results are consistent with the conclusions researched by others IS business value researchers, namely those who rise the problem of the "productivity paradox".

It can also be true that the stock of IS/IT capital of the Portuguese companies is not enough to produce positive impacts, they are in the learning adjustment process. Portuguese managers could not be investing in complementary organizational investments to get better results from there IS/IT investments.

This paper aims to discuss and test the use of stakeholder theory in the IS business value stream of research, in order to achieve a better understanding of the impact of IS/IT investments on firm performance (moderated by stakeholder management). To achieve the expected impact from an IS/IT investment, it is argued that firms need a strong commitment from these stakeholder groups, which lead us to the need of a "stakeholder orientation".

When firm financial performance is measured by returns on assets (ROA), returns on investments (ROI) and returns on sales (ROS), the results show that "stakeholder orientation" impact positively in the relation between IS/IT and firm performance, using a sample of Portuguese large companies.

7. References

Al-Mudimigh A, Zairi M, and Al-Mashari M. "ERP software implementation: an integrative framework". European Journal of Information Systems, 10, (2001), 216-226.

Bakos, J., "Dependent variables for the study of firm and industry-level impacts of information technology", In Proceedings of the 8th International Conference on Information Systems, Pittsburgh, 1987.

Ballantine J and Stray S., "Information systems and other capital investments: evaluation practices compared", Logistics Information Management, 12, 1, (1999), 78-93.

Barney, J., "Firm resources and sustainable competitive advantage", Journal of Management, 17, 1, (1991), 99-120.

Bender, D., "Financial Impact of Information Processing", Journal of Management Information Systems, 3, 2, (1986).

Berman, S.; Kotha, S. and Jones, T., "Does Stakeholder Orientation Matter? The Relationship between Stakeholder Management Models and Firm Financial Performance", Academy of Management Journal, 42, 5, (1999), 488-506.

Bharadwaj, A., "A resource-based perspective on information technology capability and firm performance: An empirical investigation", MIS Quarterly, 24, 1, (2000), 169-196.

Breshnihan, T., "Measuring Spillovers from Technical Advance: Mainframe Computers in Financial Services", American Economic Review, 76, 4, (1986), 742-755.

Brown R, Gatian A. and Hicks J., "Strategic information systems and financial performance", Journal of Management Information Systems, 11, 4, (1995), 215-249.

Brynjolfsson, E., "The Productivity Paradox of Information Technology: Review and Assessment", MIT Sloan School of Management Working Paper, 1992.

Brynjolfsson E and Hitt L., "Beyond the productivity paradox", Communications of the ACM, 41, 8, (1998), 49-55.

Brynjolfsson, E. and Yang, S., "Information Technology and Productivity: A Review of the Literature", [online], MIT, http://ccs.mit.edu/papers/CCSWP202.html, 1996.

Caldeira, M., "Understanding the adoption and use of information systems / information technology in small and medium-sized manufacturing enterprises: a study in Portuguese industry", PhD. thesis, Cranfield University, 1998, https://dspace.lib.cranfield.ac.uk/handle/1826/4516.

Caldeira, M. and Ward, J., "Using resource-based theory to interpret the successful adoption and use of information systems and technology in manufacturing small and medium-sized enterprises", European Journal of Information Systems, 12, (2003), 127-141.

Coase, R., "The nature of the firm", Economica, 4, 16, (1937), 386-405.

Conner, K., "A historical comparison of resource-based theory and five schools of thought within industrial organization economics: do we have a new theory of the firm?", Journal of Management, 17, 1, (1991), 121-154.

Conner, K. and Prahalad, "A resource-based theory of the firm: Knowledge versus opportunism", Organization Science, 7, 5, (1996), 477-501.

Cron, W. and Sobol, M., "The Relationship between Computerization and Performance: A Strategy for Maximizing the Economic Benefits of Computerization", Information & Management, 6, 3, (1983), 171-181.

Crowston, K. and Treacy, M., "Assessing the Impact of Information Technology on Enterprise Level Performance", Sloan School of Management Working Paper N.1835, 1986.

Dasgupta S, Sarkis J. and Talluri S., "Influence of information technology investment on firm productivity: a cross-sectional study", Logistics Information Management, 12, 1, (1999), 120-129.

Demsetz, H., "The theory of the firm revisited", Journal of Law, Economics, and Organization, 4, 1, (1988), 141-161.

Donaldson T and Preston L., "The stakeholder theory of the corporation: Evidence and implications", The Academy of Management Review, 20, 1, (1995), 65-91.

Earl M., *Management Strategies for Information technology*, Prentice-Hall, Englewood Cliffs, New Jersey, 1989.

Farbey B, Land F and Targett D., "Moving IS evaluation forward: learning themes and research issues", Journal of Strategic Information Systems, 8, 1999, 189-207.

Francalanci C and Galal H. "Aligning IT investments and workforce composition: the impact of diversification in life insurance companies", European Journal of information Systems, 7, (1998), 175-184.

Freeman R., *Strategic management: A stakeholder approach*, Pitman/Ballinger, Boston, 1984.

Grover, V.; Teng, J. and Cheon, M., "Towards a theoretically-based contingency model of information systems outsourcing", in Strategic Sourcing of Information Systems, L. Willcocks and M. Lacity (Eds), Wiley, Chichester, 1998, 79-101.

Harris S E and Katz J L., "Predicting organizational performance using information technology managerial control ratios", In Proceedings of the Twenty-Second Annual Hawaii International Conference on System Sciences, Computer Society Press, Los Alamitas, CA: 197, 1989.

Hitt L, Wu D and Zhou X., "Investment in Enterprise resource Planning: Business impact and productivity measures", Journal of Management Information Systems, 19, 1, (2002), 71-98.

Jones T., "Instrumental stakeholder theory: A synthesis of ethics and economics", The Academy of Management Review, 20, 2, (1995), 404-437.

Kauffman R and Weill P., "An evaluative framework for research on the performance effects of information technology investment", In Proceedings of the 10th International Conference

on Information Systems, J. I. DeGross, J. C. Henderson, and B. R. Konsynski (Eds), Boston, 1989, 377-388.

Kivijärvi H and Saarinen T., "Investment in information systems and the financial performance of the firm", Information & Management, 28, (1995), 143-163.

Lee J and Bose U., "Operational linkage between diverse dimensions of information technology investments and multifaceted aspects of a firm's economic performance", Journal of Information Technology, 17, (2002), 119-131.

Love P.E.D and Irani Z., "An exploratory study of information technology evaluation and benefits management practices of SMEs in the construction industry", Information & Management, 42, (2004), 227-242.

Lucas, H., "The use of an accounting information system, action and organizational performance", The Accounting Review, October, (1975), 735-746.

Mata, F.; Fuerst, W. and Barney, J., "Information technology and sustained competitive advantage: A resource-based analysis", MIS Quarterly, 19, 4, (1995), 487-505.

McFarlan F.W., "Information technology changes the way you compete", Harvard Business Review, 62, 3, (1984), 98-103.

Osey-Bryson K and Ko M., "Exploring the relationship between information technology investments and firm performance using regression splines analysis", Information & Management, 42, (2004) 1-13.

PIMS Program, "management productivity and information technology", The Strategic Planning Institute, Cambridge, Massachusetts,1984.

Porter M and Millar V.E., "How information gives you competitive advantage", Harvard Business Review, 63, 4, (1985), 149-160.

Prahalad, C. and Hamel, G., "The core competence of the corporation", Harvard Business Review, May-June, (1990), 79-91.

Remenyi D., "The elusive nature of delivering benefits from IT", Electronic Journal of Information Systems Evaluation, 3, 1, 1999.

Rindfleisch, A. and Heide, J., "Transaction cost analysis: Past, present and future applications", Journal of Marketing, 61, 4, (1997), 30-54.

Roach, S., "Technology and the Services Sector: America's Hidden Competitive Challenge", Economic Perspectives, Morgan Stanley, 1987.

Ryan S and Gates M., "Inclusion of Social Subsystem Issues in IT investment decisions: An empirical assessment", Information Resources Management Journal, 17, 1, (2004).

Santhanam, R. and Hartono, E., "Issues in Linking Information Technology capability to firm performance", MIS Quarterly, Vol 27, No.1, (2003), 125-153.

Shin N., "The impact of information technology on financial performance: the importance of strategic choice", European Journal of Information Systems, 10, (2001), 227-236.

Simon, H., "Organizations and markets", Journal of Economic Perspectives, 5, 2, (1991), 25-44.

Solow, R.M, "We'd Better Watch Out," New York Times Book Review, 12 July, (1987).

Stickney, C. "Financial Statement Analysis: A Strategic Perspective", Harcourt Brace Jovanovich, São Diego, USA, 1990.

Strassman P., *The squandered computer*, Information Economics Press, New Canaan, (1997).

Ward J and Peppard J. , *Strategic Planning for Information Systems*, 3rd Edition, John Willey & Sons, Chichester, 2002.

Willcocks L., "IT Evaluation: Managing the Catch 22", European Management Journal, 10, 2, (1992), 220-229.

Willcocks, L. and Lacity, M. "The sourcing and outsourcing of IS: Shock of the new?", in Strategic Sourcing of Information Systems, L. Willcocks and M. Lacity (Eds), Wiley, Chichester, 1998, 1-41.

Willcocks L and Lester S., *Beyond The IT Productivity Paradox*, Wiley Series on Information Systems, Wiley, Chichester, 1999.

Williamson, O. *Markets and hierarchies: Analysis and antitrust implications*, The Free Press, New York, 1975.

Williamson, O. "Transaction-cost economics: The governance of contractual relations", The Journal of Law and Economics, 22, (1979), 233-261.

Williamson, O. *The economic institutions of capitalism: Firms, markets, relational contracting*, The Free Press, New York, 1985.

Williamson, O, *The mechanisms of Governance*, Oxford University Press, New York, 1996.

Estruturas Organizacionais para Gestão do Valor das TI: Um Modelo Conceptual para as Universidades Públicas em Portugal

Cristiano Pereira [1], Carlos Ferreira [2], Luis Amaral [3]

1) STIC - Universidade de Aveiro, Portugal
cristiano@ua.pt

2) DEGEI/IEETA – Universidade de Aveiro, Portugal
carlosf@ua.pt

3) Universidade do Minho, Portugal
amaral@dsi.uminho.pt

Resumo

A Governança Corporativa das Tecnologias de Informação é uma temática primordial reconhecida pela comunidade académica e profissional. A necessidade de obter valor, seja financeiro ou estratégico para a organização, resultante dos investimentos realizados em TI implica uma Gestão do Valor das TI adequada. No contexto das Universidades Públicas Nacionais esta problemática é acentuada pelas suas especificidades organizacionais que dificultam a implementação dos modelos e práticas mais relevantes como o COBIT 5 ou Val IT. Este trabalho tem como objetivo identificar as Estruturas Organizacionais de apoio à Gestão do Valor das TI devidamente enquadradas no modelo de Governo das Universidades Públicas em Portugal usando uma metodologia de investigação baseada numa revisão da literatura, que suportará o desenho de um modelo conceptual para identificação de tais Estruturas.

Palavras chave: Governança das TI, Gestão do Valor, Estruturas Organizacionais, COBIT 5, Val IT 2.0.

1. Introdução

As organizações, sejam elas públicas ou privadas, vêm-se deparando nos últimos anos com uma cada vez maior dependência das Tecnologias de Informação (TI), essenciais para a sua sustentabilidade e desenvolvimento, no suporte, operação e prospeção do seu negócio. O incremento dos investimentos em TI aumenta a preocupação das organizações em garantir os benefícios esperados, pois são vários os estudos [Cook 2007], [Johnson and Hencke 2008], que apontam para casos de insucesso com os investimentos realizados em TI [Wilkin et al. 2013]. Este é um dos dilemas mais comuns com que as organizações e os seus responsáveis se vêm confrontando, isto é, como é que se garante a realização de valor dos elevados investimentos realizados em TI [De Haes et al. 2013]. A criação de valor significa obtenção de benefícios com otimização de custos dos recursos enquanto se otimiza os riscos [Moeller 2013], aquilo que

deve ser o objetivo das práticas de Gestão do Valor das TI (GVTI). As organizações que adotam boas práticas de GVTI veem facilitada a identificação, criação e obtenção de valor dos investimentos realizados em TI [Keyes-pearce 2005]. Aquelas que apresentam um nível de maturidade da Governança das TI superior à média têm, pelo menos, 20% mais retorno dos investimentos em TI do que uma organização com um grau de maturidade inferior [Lunardi et al. 2014].

Dada a centralidade das TI no negócio e na criação de valor para as organizações surgiu ao longo das duas últimas décadas um foco específico em Governança Corporativa das TI (GCTI) [De Haes et al. 2013]. A Governança das TI passou a ser vista como parte integrante da Governança Corporativa (GC) [De Haes and Van Grembergen 2015].

Conhecedoras desta realidade, as organizações procuram soluções para resolver a problemática descrita através da adoção de *Frameworks* de Governança das TI propostas pela comunidade profissional, ou em alternativa, optam por desenhar e implementar modelos próprios adaptados à sua realidade organizacional. Isto justifica-se, por um lado, pelo elevado nível de complexidade de implementação percecionado pelas empresas [Bartens et al. 2014] e por outro pela falta de conhecimento sobre as *Frameworks* [Zhang and Fever 2013].

O referido anteriormente para as organizações em geral, não é de menor importância no contexto das Instituições de Ensino Superior (IES), também nestas a GCTI e a GVTI são uma questão pertinente e identificada como prioritária. A corroborar esta afirmação o estudo realizado pela EDUCAUSE [Grajek 2014], destaca duas questões relacionadas com a temática, i) A necessidade de estabelecer e implementar o conceito de Governança das TI na instituição de modo a priorizar os investimentos em TI e ii) Estabelecer um alinhamento estratégico entre as lideranças de TI e as lideranças institucionais.

Organizações com estruturas descentralizadas como são tipicamente o caso das universidades, necessitam avaliar e renovar regularmente as suas estruturas de Governança das TI, para se adaptarem à evolução do negócio e do ambiente tecnológico. Quaisquer transformações a operar neste âmbito devem considerar uma visão holística da organização [Pereira and Ferreira 2014]. É neste contexto de transformação organizacional que o modelo conceptual (Figura 2) enquadra as Estruturas Organizacionais para a GVTI.

A secção 2 descreve a metodologia de investigação. Posteriormente, na secção 3, apresenta-se um estudo do trabalho relacionado que serve de suporte ao modelo conceptual apresentado na secção 4. Com base na análise dos resultados da revisão da literatura, descrevem-se na secção 5 as Estruturas Organizacionais de Gestão do Valor das TI propostas. Finalmente na secção 6, são apresentadas as conclusões e trabalho futuro.

2. Objetivos e Metodologia de Investigação

Este trabalho tem como principal objetivo, desenvolver um modelo conceptual que identifica as Estruturas Organizacionais para a GVTI no âmbito da Governança das TI enquadradas nas atuais estruturas de Governança das Universidades Públicas nacionais. A abordagem metodológica é baseada numa revisão da literatura. De acordo com [Webster and Watson 2002] a revisão da literatura em temáticas suficientemente exploradas, deve propor um modelo conceptual que sintetize e contribua para o conhecimento existente. Considera-se que o grau de maturidade da investigação nesta temática é suficiente para seguir as orientações propostas por Watson e Webster.

A Figura 1 apresenta as etapas da metodologia que suportam este trabalho. Cada uma das fases do processo, baseado na *Framework* de [Vom Brocke et al. 2009], contribuem para a investigação com os resultados destacados a negrito. A primeira fase do processo tem como objetivo enquadrar e definir o âmbito da revisão da literatura. Para [Vom Brocke et al. 2009] a revisão da literatura deve ser precedida de uma descrição suficientemente detalhada para possibilitar uma avaliação do grau de profundidade aplicado e em simultâneo permitir a outros investigadores uma maior confiança na utilização das conclusões noutros trabalhos de investigação. Desta premissa resulta a descrição, apresentada posteriormente, apoiada na taxonomia proposta por [Cooper 1988]. Esta fase fica concluída com um estudo do trabalho relacionado com a Governança e a Gestão do Valor das TI (secção 3). Na fase II desenvolveu-se um modelo conceptual (secção 4) que mapeia os conceitos chave em estudo, e em simultâneo colocou a descoberto as lacunas existentes, isto é, quais as Estruturas para a GVTI que as IES devem implementar, abrindo assim oportunidade para a realização deste estudo. Na terceira etapa procedeu-se a um processo de seleção da literatura com referência às componentes de GCTI e às práticas de GVTI, enquadráveis no modelo conceptual (secção 5). A última etapa, Fase IV, consiste na análise e síntese dos resultados alcançados (secção 6.).

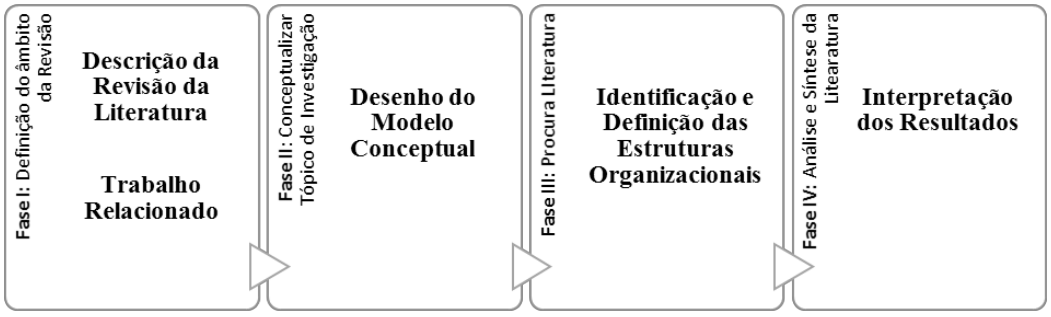


Figura 1 – Componentes Integradas da Metodologia de Investigação

A revisão da literatura focou-se em publicações académicas e de âmbito profissional. A pesquisa realizou-se em diversas bases de dados de referência (EBSCO, JSTOR, WILEY e ScienceDirect). Efetuaram-se pesquisas, em todo o texto da publicação, com termos como "*IT Value Management*", "*IT Governance*", "*ITG Mechanisms*", "*Higher Education*", sem qualquer restrição no período temporal. Esta pesquisa foi reforçada com uma procura no "Google Scholar" considerando os mesmos termos. Os dados recolhidos foram organizados de acordo com o modelo conceptual apresentado na secção 4. Consideram-se os resultados obtidos nesta investigação de particular interesse para académicos e profissionais que desenvolvem a sua atividade nesta área de conhecimento, particularmente no âmbito de IES.

3. Trabalho Relacionado

A Governança das Tecnologias de Informação é um conceito que se tornou uma questão fundamental na área das TI. Devido ao foco do termo TI neste conceito, a sua discussão restringia-se à área de conhecimento das Tecnologias de Informação. É no entanto óbvio que o valor para a organização não será alcançado apenas pelas TI, é necessário envolver o negócio [De Haes and Van Grembergen 2015]. Têm sido vários os autores que defendem esta abordagem [Thorp 2003], [Weill and Ross 2009], [Peppard 2010], similarmente o ITGI (*IT Governance Institute*), afirma explicitamente que a Governança das TI é uma parte integral da Governança Corporativa (GC).

A literatura apresenta uma diversidade significativa de definições de Governança Corporativa [Shleifer and Vishny 1997], [ANAO 2003] ou [Webb et al. 2006]. Pode resumir-se o conceito de Governança Corporativa como o conjunto de atividades de avaliação, direção e monitorização dos recursos da organização, isto é estabelece as responsabilidades e autoridades sobre as decisões tomadas na empresa ou organização. No atual contexto de competitividade económica, é objetivo da GC fornecer uma estrutura que garanta a definição, obtenção e monitorização dos objetivos organizacionais, recorrendo para tal aos meios e recursos considerados adequados [OECD 2004].

O incremento estratégico e funcional do papel das TI, e das questões relacionadas com aspetos táticos e de tomada de decisão nas organizações originou uma compreensão mais ampla da Governança das TI, designada Governança Corporativa das TI (GCTI) [Wilkin and Chenhall 2010]. Na definição de GCTI apresentada por [De Haes and Van Grembergen 2015] depreende-se que esta vai além das responsabilidades específicas das TI, expandindo-se aos aspetos de negócio relacionados com as TI, necessários à criação de valor para a organização.

3.1 Governança nas Instituições de Ensino Superior em Portugal

Apesar da diversidade de definições existentes na literatura para o conceito de Governança da Educação Superior, o termo governança não tem sido muito usado quando se trata a educação superior em Portugal [CNE 2013]. Para este trabalho adotou-se a definição apresentada por [Santiago et al. 2009] que considera o termo governança com o seguinte sentido "*estruturas, relações e processos através dos quais as políticas para a educação superior são desenvolvidas, aplicadas e revistas, tanto a nível nacional como institucional*".

Em Portugal o modelo de governança institucional está consagrado no Regime Jurídico das Instituições de Ensino Superior (RJIES), instituído pela Lei nº 62/2007, de 10 de Setembro. Esta lei atribui às IES o estatuto de pessoas jurídicas de direito público, sob a tutela do Ministério da Ciência, Tecnologia e Inovação, conferindo-lhes maior autonomia académica e de gestão de recursos. Em resumo o modelo de governança adotado em Portugal assenta na atribuição de responsabilidade da governança estratégica institucional a um Conselho Geral, que elege o/a reitor/a ou presidente da instituição. [Pedrosa et al. 2014]

O Conselho Geral é constituído por representantes do corpo docente, investigadores, estudantes e por membros cooptados externos à instituição, com conhecimentos e experiência relevante neste âmbito. As competências do reitor são amplas e muito importantes, cabendo-lhe desde logo dirigir e representar a universidade [Oliveira et al. 2014]. Entre outras matérias cabe ao reitor presidir e nomear os membros do Conselho de Gestão.

Do estudo empírico acerca da Autonomia e Governança das Instituições de Ensino Superior [CNE 2013], importa destacar o modo de funcionamento e de organização interna dos Conselhos que de acordo com os autores são merecedores de referência e que poderão ser sujeitos a ajustamentos futuros. "*Identifica-se a necessidade de agir sobre as vias e processos de comunicação entre o Conselho, outros órgãos de governança interna existentes em vários níveis e a comunidade académica. (...) alguns Conselhos a optarem por estruturas permanentes dedicadas a áreas precisas (recursos, ensino e investigação, assuntos jurídicos, finanças, plano estratégico ...) e outros a criarem comissões para tratar assuntos da agenda sem caracter permanente (reestruturação dos saberes ...) "*. Em síntese a estruturação do Concelho Geral em comissões ou comités especializados não é uma prática generalizada e em alguns casos não é prestada a atenção devida às condições exigidas para um bom funcionamento de um órgão com o estatuto e autonomia do Conselho Geral.

Se no passado a missão das IES no essencial se restringia ao domínio do Ensino e Qualificação dos seus alunos e ao domínio da Investigação e Produção Cientifica, atualmente o destaque dado a outras funções vem revelar domínios que servem como catalisadores para realçar a

importância destes novos modelos de Governança, nomeadamente funções no domínio do Desenvolvimento Económico e Social das Regiões onde se inserem e no domínio da Internacionalização e Mobilidade dos seus estudantes e colaboradores. Esta multiplicidade de papéis cria tensões internas nas Instituições pelo que são fundamentais novos mecanismos de Governança e Gestão para as combater [Machado 2014].

3.2 Domínios da Governança Corporativa das Tecnologias de Informação

Para o ITGI, a GCTI preocupa-se fundamentalmente com dois componentes: entrega de valor ao negócio por parte das TI e mitigação de riscos associados às TI. Os componentes de Entrega de Valor e Gestão de Risco são os resultados da Governança das TI enquanto o Alinhamento Estratégico, a Gestão de Recursos e Gestão do Desempenho são catalisadores para alcançar os resultados. Esta abordagem holística conduziu à definição de cinco domínios de Governança Corporativa das Tecnologias de Informação: i) Alinhamento Estratégico; ii) Entrega de Valor; iii) Gestão de Risco; iv) Gestão de Recursos; v) Gestão de Desempenho [ITGI 2003].

3.2.1 Alinhamento Estratégico

O Alinhamento Estratégico é considerado um dos fatores fundamentais na criação de valor das TI [Kohli and Grover 2008]. Vários estudos demonstram as implicações do alinhamento estratégico, no desempenho do negócio e das TI, por exemplo, [Chan et al. 2006], [Oh and Pinsonneault 2007], [Yayla and Hu 2012], [Gerow et al. 2014]. Estes trabalhos concluem que uma organização ao implementar com sucesso o alinhamento da estratégia de negócio com a estratégia de TI terá um desempenho superior àquelas que não o façam. A evolução e inovação dos processos de negócio, cada vez mais dependentes das TI, tornam o Alinhamento Estratégico um mecanismo chave através do qual as TI criam valor [Coltman et al. 2015].

A liderança das TI, o processo de Gestão de Portfólios para priorizar investimentos e projetos em TI, a implementação de comités que integram elementos das diversas áreas de negócio incluindo TI e a inclusão das TI na agenda dos conselhos de administração, são fatores identificados em [Luftman et al. 1999] e confirmados por [De Haes and Van Grembergen 2009] que afetam positivamente o Alinhamento Estratégico.

3.2.2 Entrega de Valor

A necessidade das organizações conhecerem qual o valor entregue pelos investimentos em TI, revela a importância do domínio da Entrega de Valor. A Entrega de Valor concentra-se na otimização dos custos demonstrando o valor associado das TI. A Entrega de Valor considera o risco e o impacto das atividades nos recursos e processos do negócio [Rivard et al. 2006] e

depende de uma diversidade de fatores organizacionais, nomeadamente dimensão da organização, estruturas, cultura, conhecimento e competências [Cao 2010]. A escolha adequada dos investimentos em TI, a sua gestão ao longo do ciclo de vida do investimento e as capacidades das TI para responder às necessidades do negócio são outros aspetos a considerar neste domínio [ITGI 2005].

Um instrumento importante na entrega/criação de valor é o caso de negócio (*Business Case*) [Maes et al. 2014]. Este instrumento não deve ser visto apenas como um documento formal para tomada de decisão de um determinado investimento, mas antes deve ser um documento utilizado ao longo de todo o ciclo de vida do investimento.

3.2.3 Gestão do Risco

A gestão de risco pretende assegurar que os objetivos estratégicos do negócio não são comprometidos por falhas ocorridas nas TI [ITGI 2003]. O risco associado à tecnologia sempre existiu, seja este detetado e reconhecido ou não pela organização [ISACA 2009]. Esta constatação torna a gestão de risco uma preocupação crescente. Uma gestão de risco eficaz implica uma análise inicial da propensão que a organização tem ao risco. Esta abordagem tem impacto nos investimentos a realizar em tecnologias de proteção dos ativos de TI e no nível de segurança necessário. Conhecida a realidade da organização (apetência e nível de exposição ao risco), podem ser definidas as estratégias de gestão do risco. Dependendo do tipo de risco e do seu impacto para o negócio, este pode ser mitigado, transferido ou aceite [ITGI 2003].

A investigação nesta área de conhecimento tem-se focado no desenvolvimento de modelos de identificação do risco [Gewald and Helbig 2006], gestão do risco [Young and Jordan 2002] e modelos de avaliação e medição do risco [Du et al. 2006].

3.2.4 Gestão de Recursos

A atribuição de recursos de TI adequados aos investimentos são um fator de sucesso para o desempenho das TI, entenda-se como recursos adequados, infraestruturas tecnológicas atualizadas, recursos humanos com experiência e conhecimentos técnicos e recursos financeiros suficientes para a manutenção dos serviços prestados [ITGI 2003]. Um dos desafios atuais com que as organizações se confrontam é a decisão de fornecer serviços de TI internamente ou contratar a prestadores de serviços externos (*Outsourcing*) com a garantia de que o valor entregue pelos fornecedores seja o contratado. A capacidade para equilibrar custos de recursos TI, com qualidade de serviço pretendida é essencial na GVTI, este equilíbrio consegue-se recorrendo a ferramentas de Gestão de Desempenho.

3.2.5 Gestão de Desempenho

Seja pelas necessidades referidas na secção anterior, seja pela importância de demonstrar a eficácia das TI e o valor que acrescentam ao negócio, a implementação de práticas de Gestão de Desempenho é fundamental para as organizações.

Os métodos de avaliação de desempenho tradicionais, baseados em avaliações financeiras como o retorno de investimentos (*ROI - Return of Investiment*), valor presente líquido (*NPV - Net Present Value*) ou período de retorno do investimento (*PB - Payback Period*) apenas refletem uma parte do valor das TI (componente tangível). Para integrar não só os fatores tangíveis mas também os fatores intangíveis da utilização das TI [Van Grembergen and De Haes 2005] propõem um método baseado no BSC [Kaplan and Norton 1992] devidamente adaptado às TI. Este método permite reportar aos decisores da organização o valor das TI numa "linguagem" mais orientada ao negócio.

3.3 Mecanismos de Governança das Tecnologias de Informação

Desenvolver um modelo de Governança das TI é só o primeiro passo; a sua implementação prática na organização, como uma solução sustentável, é o próximo desafio. De que modo podem as organizações implementar na prática um modelo de GCTI é uma das principais questões colocadas pelos executivos e académicos. [De Haes and Van Grembergen 2015] defende, que na implementação prática de um modelo de Governança das TI uma organização deve recorrer a um misto de mecanismos:

- Estruturas Organizacionais – Organização e atribuição de funções de TI a pessoas ou departamentos específicos, definição clara de papéis e responsabilidades.
- Processos – Tomada de decisões estratégicas, planeamento estratégico de sistemas de TI, gestão e monitorização dos serviços, ferramentas de definição de processos.
- Mecanismos Relacionais – Suporte da relação que deve existir entre as TI e o negócio como o diálogo estratégico, troca de experiências, comunicação e troca de conhecimento.

A literatura identifica um conjunto diversificado de práticas ou mecanismos de Governança das TI no entanto, a decisão de quais devem ser implementados tem que atender ao contexto e especificidades da organização, nomeadamente o seu setor de atividade, dimensão e cultura organizacional [De Haes and Van Grembergen 2015].

4 Modelo Conceptual

A secção anterior apresenta a associação entre os conceitos de Governança Corporativa, Governança das TI e Gestão do Valor das TI e simultaneamente a necessidade de adaptar os mecanismos de GCTI (Estruturas, Processos e Mecanismos Relacionais) ao contexto organizacional onde se pretende implementar.

A definição de um novo modelo de Governança das IES nacionais (secção 3.1) conduziu a novas estruturas organizacionais que devem estar alinhadas com as TI para atingir os objetivos estratégicos pretendidos pela instituição. Num contexto de transformação, a Governança das TI requer uma abordagem holística a toda a instituição (TI, negócio, pessoas, comportamentos) atuando nos cinco domínios da Governança das TI e cuja implementação prática é suportada pelas práticas ou mecanismos de GCTI (Figura 2).

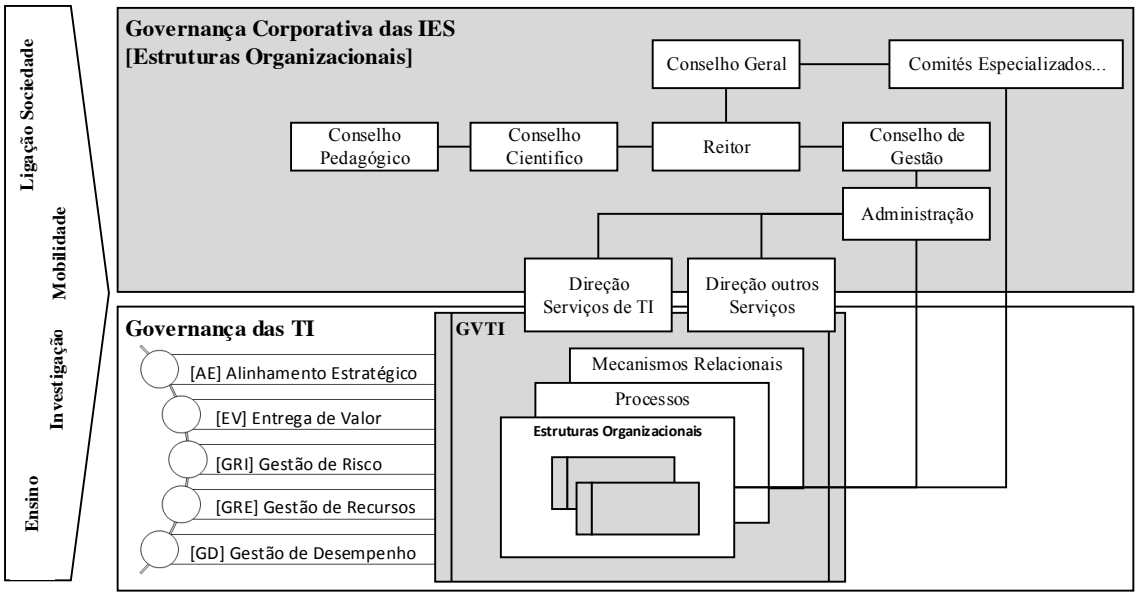


Figura 2 – Modelo Conceptual dos componentes de GCTI em IES Nacionais

Com base na proposta de [CNE 2013] que refere a utilidade e necessidade da existência de estruturas de suporte especializadas para auxiliar os Conselhos Gerais das Instituições de Ensino Superior, o modelo desenvolvido reflete a integração de Estruturas Organizacionais no domínio da Governança e Gestão do Valor das TI no atual modelo de Governança das IES (RJIES).

Os três tipos de mecanismos de GTI são essenciais e devem ser combinados por forma a criar uma abordagem holística que promova o sucesso da implementação prática da GCTI nas IES. No entanto sendo o objetivo deste artigo a identificação das Estruturas Organizacionais de GVTI, apenas estas são destacadas (a tracejado) no modelo apresentado.

5 Resultados da Revisão da Literatura

A Tabela 1 resume as principais Estruturas Organizacionais de Governança e de Gestão de Valor das TI encontradas na literatura. Para identificar as Estruturas Organizacionais mais divulgadas, foram escolhidos quatro trabalhos que enumeram um conjunto de Estruturas Organizacionais de suporte às atividades de Governança e Gestão do Valor das TI.

Com o objetivo de tornar a análise abrangente optou-se por comparar as estruturas propostas para Governança das TI (GTI) sob dois pontos de vista distintos, isto é, um académico (A) [De Haes and Van Grembergen 2015] e um do ponto de vista da comunidade profissional (B) [ISACA 2012]. Quanto às Estruturas Organizacionais de apoio à Gestão do Valor (GV), optou-se por dois trabalhos académicos (C) [Maes et al. 2015] e (D) [Pereira and Ferreira 2015], o primeiro apresenta a visão académica, enquanto o segundo identifica as Estruturas Organizacionais de GV, que refletem a integração da Framework Val IT 2.0 [ITGI 2008] no COBIT 5 [ISACA 2012], dando a visão da comunidade profissional pretendida. A Tabela 1 resume as práticas ou mecanismos de Estruturas Organizacionais que foram identificados nos quatro trabalhos analisados.

As duas primeiras Estruturas Organizacionais (Tabela 1), Comité de Estratégia de TI e Comité Diretivo de TI, refletem respetivamente os dois níveis de tomada de decisão que tipicamente estão presentes numa organização: i) nível estratégico e ii) nível de gestão. O primeiro nível composto por um conselho de administração assessorado pelo Comité de Estratégia das TI e ao nível de gestão uma direção executiva apoiada por um Comité Diretivo de TI.

A influência dos mecanismos de GCTI no domínio da GVTI é confirmada pelo número de Estruturas Organizacionais comuns. Verifica-se na Tabela 1, que em sete Estruturas Organizacionais identificadas apenas uma não é comum.

Sem prejuízo da relevância de todas as Estruturas Organizacionais identificadas, é objetivo deste trabalho propor as que se consideram prioritárias no processo de implementação prática da GVTI nas Universidades Portuguesas. Para tal definiram-se os seguintes critérios de escolha: i) Estruturas Organizacionais referenciadas simultaneamente no âmbito da Governança (GTI) e no âmbito da Gestão do Valor (GV); ii) Estruturas Organizacionais propostas pela comunidade académica e simultaneamente por *Frameworks* desenvolvidas pela comunidade profissional.

Fundamentado nos critérios definidos, as Estruturas Organizacionais para Gestão do Valor das TI propostas são o Comité Estratégico das TI (ao nível do Conselho Geral) e o Gabinete de Gestão do Valor. Nas subsecções seguintes aprofundamos a integração destas na atual estrutura de governo das universidades com base no modelo conceptual da Figura 2.

Estrutura Organizacional	Descrição	GTI		GV	
		A	B	C	D
Comité Estratégico de TI	Composto por elementos seniores nomeados pelo conselho de administração para garantir que este é informado sobre as questões e decisões de TI.	☑	☑	☑	☑
Comité Diretivo de TI (Decisão de Investimentos, Projetos e Programas)	Composto por especialistas de TI e do negócio é responsável pela orientação de programas, projetos de TI, e de portfólios de investimento. Avaliam, priorizam e escolhem os casos de negócio, garantindo a criação de valor e gestão de riscos. É responsável por estimular a gestão de topo e o negócio/liderança das TI.	☑	☑	☑	
Gabinete de Governança das TI	Responsável na organização por promover e gerir os processos de governança das TI.	☑			
Comité Diretivo de Arquitetura	Composto por elementos do negócio e das TI que fornecem orientação dos assuntos e decisões acerca da arquitetura corporativa da organização e pela definição de políticas e normas.	☑	☑		☑
Gabinete de Segurança/Conformidade e Risco	É responsável pela gestão da segurança, conformidade e risco, que possam causar impacto nas TI.	☑	☑		☑
Gabinete de Gestão do Valor	Atua na gestão de portfólios de investimentos e serviços, avaliando e aconselhando acerca das oportunidades de investimento. Apoia os comités de TI na avaliação de casos de negócio, acompanha os investimentos individuais em TI, informando sobre o seu progresso e criação de valor gerado pelos investimentos e serviços.	☑	☑	☑	☑
Gabinete de Programas e Projetos	Órgão responsável pelo apoio aos gestores de programas e projetos, avaliam e reportam informações sobre os projetos que constituem os programas. Definem e disponibilizam as metodologias, normas e ferramentas adotadas pela organização para a gestão dos programas e projetos.		☑	☑	☑

Tabela 1 – Estruturas Organizacionais de Governança e Gestão do Valor das TI.

5.1 Comité Estratégico de TI

Para proporcionar uma estratégia tecnológica que vá de encontro às necessidades da instituição e que apoie na Gestão do Valor das TI é proposta a criação de um Comité Estratégico de TI. Este Comité deve ser composto por elementos nomeados pelo Conselho Geral e se possível alguns independentes, que informem este órgão acerca das questões de TI que afetam a Universidade. Deve ser presidido por um elemento do Conselho Geral e não pelo CIO ou Diretor de Serviços das TI [ISACA 2012]. Os membros devem trazer uma visão das diversas vertentes da instituição onde as TI podem acrescentar valor, nomeadamente no ensino, investigação e ligação à sociedade.

A criação de um Comité Estratégico de TI deve atender aos seguintes aspetos: i) participação de um conselheiro independente com experiencia tanto na área do negócio como nas TI; ii) Os membros executivos, particularmente o CIO, ou gestores de TI devem evitar a discussão das questões de um ponto de vista meramente técnico, antes devem ter capacidade para discutir os temas em termos que sejam compreendidos por elementos não especializados em TI centrando-se nas oportunidades, problemas, riscos e planos; iii) Garantir que os temas relacionados com as TI estão presentes de uma forma regular na agenda do Conselho Geral [Martínez and Largo 2011].

Em síntese o Comité de Estratégia das TI deve ser o promotor e o catalisador da utilização de boas práticas de Governança e Gestão do Valor das TI por parte do Conselho Geral e ajudar os executivos a gerir as TI de acordo com as estratégias definidas para a instituição.

5.2 Gabinete de Gestão do Valor

Se por um lado o Comité de Estratégia das TI, atua a um nível estratégico, o Gabinete de Gestão do Valor atua no nível executivo, nomeadamente na gestão do portfólio de investimentos e serviços de TI, avaliando tecnicamente as oportunidades de investimento alinhadas com a estratégia de TI definida pela IES. Esta estrutura avalia e define processos para a criação e manutenção de casos de negócio, monitoriza os investimentos individuais de TI aprovados e mantem informados os decisores, nomeadamente o Comité de Estratégia e a Administração acerca do progresso dos investimentos e do valor gerado.

O Gabinete de Gestão do Valor apoia os diretores e gestores operacionais de programas, portfólios e projetos com informação e propostas de investimento em TI e reporta à Administração, o que permite uma visão global dos projetos previstos e em execução, por parte dos diversos diretores (TI e negócio). Com esta visão partilhada pretende-se minimizar os riscos e garantir os recursos, particularmente humanos e financeiros, que garantam as necessidades.

Para [Thorp 2003], esta estrutura deve ser complementada por um Gabinete de Decisão de Investimentos, com um papel e responsabilidades semelhante ao Comité Diretivo de TI (Tabela 1). No âmbito do modelo proposto, considera-se o Gabinete de Decisão de Investimentos uma estrutura redundante cujas funções e responsabilidades podem ser assumidas por outras estruturas já existentes nas instituições, desde a Administração até ao Conselho Geral dependendo da dimensão do investimento a realizar.

Esta estrutura não deve ser confundida com uma eventual estrutura de Gabinete de Programas e Projetos (Gestão de Projeto) já existente. O Gabinete de Gestão do Valor tem no essencial um foco no valor dos programas (Casos de Negócio) e não apenas nos custos do projeto.

6 Contribuições e Trabalho Futuro

As organizações apenas conseguem otimizar a criação de valor dos investimentos realizados em TI, através de práticas ou mecanismos de GVTI. Mesmo tendo consciência desta situação, continuam a debater-se com dificuldades na sua implementação. Na nossa opinião a principal contribuição deste estudo é identificar e elucidar os responsáveis das Universidades Nacionais acerca das Estruturas Organizacionais para GVTI a implementar.

A principal literatura no âmbito da Governança das TI, identifica Mecanismos de Governança e Gestão do valor das TI a adotar na implementação prática de um modelo de Governança, sem no entanto se focar concretamente no caso das Instituições de Ensino Superior. Com o modelo conceptual proposto neste artigo procurou-se revelar a flexibilidade existente na atual estrutura de governo das IES Portuguesas, para incorporar os conceitos de GCTI e GVTI, algo que dada a aparente "rigidez" do modelo de governo proposto no RJIES não é evidente.

Naturalmente esta investigação apresenta algumas limitações. O foco dado apenas ao mecanismo de Estruturas Organizacionais, não invalida a necessidade de identificar quais as práticas de GVTI relacionadas com os restantes mecanismos de GTI, isto é, Processos e Mecanismos Relacionais e a forma como devem ser integrados no atual modelo de Governo das IES.

Com base no modelo conceptual apresentado, aponta-se como trabalho futuro, a realização de um estudo exploratório para identificar as atuais práticas de Governança e Gestão do Valor das TI nas Universidades Portuguesas, confrontar os resultados do estudo com as *Frameworks* e modelos de GVTI existentes, e assim detalhar os diversos componentes que compõem o modelo apresentado neste artigo.

7 Referências

ANAO, 2003. Public Sector Governance - Better Practice Guide: Framework, Processes and Practices.

BARTENS, Y. ET AL., 2014. A Visualization Approach for Reducing the Perceived Complexity of COBIT 5. In *Lecture Notes in Computer Science*. pp. 403–407.

VOM BROCKE, J. ET AL., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In *17th European Conference on Information Systems*. pp. 2206–2217.

CAO, G., 2010. A four-dimensional view of IT business value. *Systems Research and Behavioral Science*, 27(3), pp.267–284.

CHAN, Y.E., SABHERWAL, R. AND THATCHER, J.B., 2006. Antecedents and Outcomes of Strategic IS Alignment : An Empirical Investigation. *IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT*, 53(1), pp.27–47.

CNE, 2013. Autonomia e Governança das Instituições Públicas de Ensino Superior.

COLTMAN, T., TALLON, P., SHARMA, R. AND QUEIROZ, M., 2015. Strategic IT alignment: twenty-five years on. *Journal of Information Technology*, 30(2), pp.91–100.

COOK, R., 2007. How To Spot a Failing Project. *CIO Magazine*, pp.1–5.

COOPER, H.M., 1988. Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1), pp.104–126.

DU, S., KEIL, M., MATHIASSEN, L., YIDE SHEN AND TIWANA, A., 2006. The Role of Perceived Control, Attention-Shaping, and Expertise in IT Project Risk Assessment. In *39th Annual Hawaii International Conference on System Sciences (HICSS'06)*. IEEE, p. 192c–192c.

GEROW, J.E., THATCHER, J.B. AND GROVER, V., 2014. Six types of IT-business strategic alignment: an investigation of the constructs and their measurement. *European Journal of Information Systems*, (2014), pp.1–27.

GEWALD, H. AND HELBIG, K., 2006. A Governance Model for Managing Outsourcing Partnerships: A View from Practice. In *39th Annual Hawaii International Conference on System Sciences (HICSS'06)*. IEEE, p. 194c–194c.

GRAJEK, S., 2014. Top Ten IT Issues, 2014 Be The Change You See. *Educause Review*, March/Apri, pp.1–32.

VAN GREMBERGEN, W. AND DE HAES, S., 2005. Measuring and Improving IT Governance Through the Balanced Scorecard. *Information Systems Control Journal*, 2, pp.1–8.

DE HAES, S. AND VAN GREMBERGEN, W., 2009. An Exploratory Study into IT Governance Implementations and its Impact on Business/IT Alignment. *Information Systems Management*, 26(2), pp.123–137.

DE HAES, S. AND VAN GREMBERGEN, W., 2015. *Enterprise Governance of Information Technology: Achieving Alignment and Value, Featuring COBIT 5* Second Edi., Boston, MA: Springer US.

DE HAES, S., VAN GREMBERGEN, W. AND DEBRECENY, R.S., 2013. COBIT 5 and Enterprise Governance of Information Technology: Building Blocks and Research Opportunities. *Journal of Information Systems*, 27(1), pp.307–324.

ISACA, 2012. *COBIT 5: A Business Framework for the Goverance and Managemnt of Enterprise IT*.

ISACA, 2009. *The Risk IT Framework*.

ITGI, 2003. *Board Briefing on IT Governance*, IT Governance Institute.

ITGI, 2008. *Enterprise Value: Governance of IT Investments, The Val IT Framework 2.0*, IT Governance Institute.

ITGI, 2005. *Optimising Value Creation From IT Investments*, IT Governance Institute.

JOHNSON, B. AND HENCKE, D., 2008. Not fit for purpose: £2bn cost of government's IT blunders. *The Guardian*.

KAPLAN, R.S. AND NORTON, D.P., 1992. The Balanced Scorecard: Measures That drive Performance. *Harvard Business Review*, (January-February), pp.71–79.

KEYES-PEARCE, S. V, 2005. *IT Value Management in Leading Firms: The Fit Between Theory and Practice*. The University of Sydney.

KOHLI, R. AND GROVER, V., 2008. Business Value of IT : An Essay on Expanding Research Directions to Keep up with the Times. *Journal of the Association fo Information Systems*, 9(1), pp.23–39.

LUFTMAN, J., PAPP, R. AND BRIER, T., 1999. Enablers and Inhibitors of Business-IT Alignment. *Communications of the Association for Information Systems*, 1(Article 11).

LUNARDI, G.L., BECKER, J.L., MAÇADA, A.C.G. AND DOLCI, P.C., 2014. The impact of adopting IT governance on financial performance: An empirical analysis among Brazilian firms. *International Journal of Accounting Information Systems*, 15(1), pp.66–81.

MACHADO, I.M.D.S., 2014. *Um Modelo De Medição E Gestão De Desempenho Para As Universidades Públicas*. Universidade de Aveiro.

MAES, K., GREMBERGEN, W. VAN AND HAES, S. DE, 2014. Identifying Multiple Dimensions of a Business Case : A Systematic Literature Review. *Electronic Journal Information Systems Evaluation*, 17(1), pp.47–59.

MAES, K., DE HAES, S. AND VAN GREMBERGEN, W., 2015. Developing a Value Management Capability: A Literature Study and Exploratory Case Study. *Information Systems Management*, 32(2), pp.82–104.

MARTÍNEZ, A.F. AND LARGO, F.L., 2011. *Gobierno de las TI para Universidades* CRUE - Conferencia de Rectores de las Universidades Españolas, ed.,

MOELLER, R.R., 2013. *Executive ' s Guide to IT Governance - Improving Systems Processes with Service Management, COBIT, and ITIL*, New Jersey: ohn Wiley & Sons, Inc., Hoboken.

OECD, 2004. OECD Principles of Corporate Governance.

OH, W. AND PINSONNEAULT, A., 2007. On the Assessment of the Strategic Value of Information Technologies: Conceptual and Analytical Approaches. *MIS Quarterly*, 31(2), pp.239–265.

OLIVEIRA, A.C., PEIXOTO, P. AND SILVA, S., 2014. *O papel dos Conselhos Gerais no Governo das Universidades Públicas Portuguesas: a lei e a prática*, Imprensa da Universidade de Coimbra.

PEDROSA, J., SANTOS, H.C., MANO, M. AND GASPAR, T., 2014. Governança da Educação Superior: contextos e processos de mudança em Portugal. *Revista Forges*, 1(1), pp.87–107.

PEPPARD, J., 2010. Unlocking the Performance of the Chief Information Officer (CIO). *California Management Review*, 52(4), pp.73–99.

PEREIRA, C. AND FERREIRA, C., 2014. Governança das Tecnologias de Informação: Um Modelo de Gestão do Valor das TI para as Universidades Públicas em Portugal. In *CAPSI2014 - 14ª Conferência da Associação Portuguesa de Sistemas de Informação*. pp. 356–361.

PEREIRA, C. AND FERREIRA, C., 2015. Identificação de Práticas e Recursos de Gestão do Valor das TI no COBIT 5. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (15), pp.17–33.

RIVARD, S., RAYMOND, L. AND VERREAULT, D., 2006. Resource-based view and competitive strategy: An integrated model of the contribution of information technology to firm performance. *The Journal of Strategic Information Systems*, 15(1), pp.29–50.

SANTIAGO, P., TREMBLAY, K., BASRI, E. AND ARNAL, E., 2009. *Tertiary Education for the Knowledge Society*, Tempus Public Foundation.

SHLEIFER, A. AND VISHNY, R.W., 1997. A Survey of Corporate Governance. *The Journal of Finance*, LII(2), pp.737–783.

THORP, J., 2003. *The Information Paradox, Realizing the Bussiness Benefits of Information Technology* Revised Ed., Fujitsu Consulting.

WEBB, P., POLLARD, C. AND RIDLEY, G., 2006. Attempting to Define IT Governance: Wisdom or Folly? In *39th Annual Hawaii International Conference on System Sciences*. IEEE, p. 194a–194a.

WEBSTER, J. AND WATSON, R.T., 2002. Analyzing The Past To Preapre For The Future: Writing a Literature Review. *MIS Quarterly*, 26(2), pp.xiii–xxiii.

WEILL, P. AND ROSS, W.J., 2009. *IT Savvy: What Top Executives Must Know to Go from Pain to Gain*, Harvard Business Press.

WILKIN, C., CAMPBELL, J., MOORE, S. AND VAN GREMBERGEN, W., 2013. Co-Creating Value from IT in a Contracted Public Sector Service Environment: Perspectives on COBIT and Val IT. *Journal of Information Systems*, 27(1), pp.283–306.

WILKIN, C.L. AND CHENHALL, R.H., 2010. A Review of IT Governance: A Taxonomy to Inform Accounting Information Systems. *Journal of Information Systems*, 24(2), pp.107–146.

YAYLA, A.A. AND HU, Q., 2012. The impact of IT-business strategic alignment on firm performance in a developing country setting: exploring moderating roles of environmental uncertainty and strategic orientation. *European Journal of Information Systems*, 21(4), pp.373–387.

YOUNG, R.C. AND JORDAN, E., 2002. IT Governance and Risk Management: an integrated multi-stakeholder framework. In *Annual Meeting of Asian-Pacific Decision Sciences Institute*. Bangkok.

ZHANG, S. AND FEVER, H. LE, 2013. An Examination of the Practicability of COBIT Framework and the Proposal of a COBIT-BSC Model. *Journal of Economics, Business and Management*, 1(4), pp.391–395.

GeoMarketing: Expansion and Development

Giovanna Monteiro [1], Ana Mafalda Barata[2], Carlos J. Costa[3]

1) ISCTE-Instituto Universitário de Lisboa, Portugal

gimonteiro18@gmail.com, gcmoa1@iscte.pt

2) ISCTE-Instituto Universitário de Lisboa, Portugal
mafaldafb@outlook.com, amfba@iscte.pt

3) ISCTE-Instituto Universitário de Lisboa, Portugal
carlos.costa@iscte.pt,

Abstract

Here, we investigate Customer Relationship Management (CRM) approach and how we can relate it to geographic information in order to improve marketing strategies and support marketer's decisions. Supported in this research, we clarify the geomarketing concept and we propose a process. This process is then applied in the context of CRM, KDD (Knowledge Discovery in Databases) and Data Mining. This process is applied in a specific case.

Keywords: Marketing, Geographic Information Systems, CRM

1. Introduction

Due to the growth of competition, organizations started to research new strategies to keep and strengthen their position in market. The need of understand which methods are more efficient to improve the organization market value, led organizations to set some important questions: Which are new market trends? Who are my customers? What do they need? Where are my competitors? What is their position in market? How can I expand my business?

Since the mid-60s, geographic component is being recognized as a relevant tool that can give answers to the questions above. It led to the creation of Geomarketing concept, which alies the most valuable marketing tools to spatial information.

Geographic Information Systems (GIS) are computer systems that bring together several features of the territory and environment, such as data field surveys, existing maps, remote sensing products, etc. They are able to store, capture, correct processed data, and perform data analysis procedures with spatial information. Therefore, Spatial Data are characterized by any kind of information or data that have location in space and are related to a coordinate system on earth.

GIS have been used more and more by decision makers to solve complex spatial problems. Although not all of GIS provide the capabilities needed for decision support, some of them have been improving in order to be representative in the new branch of decision support – Spatial Decision Support Systems [Desham, 1991]. These systems provide a framework for integration database management systems with analytical models, graphical display, and reporting capabilities.

Addressing geomarketing subject, GIS is a powerful market tool that covers organizations marketing intelligence. In this way, organizations are vertically integrating GIS and decision support system in their strategies and operations since it include analytical tools for spatial, geographical, and statistical analyses.

2. From Knowledge Management to Geomarketing

Customer Relationship Management (CRM) can be described as "the business processes and supporting technologies that support the key activities of targeting, acquiring, retaining, understanding and collaborating with customers" [Band, 2013].

CRM processes and CRM systems can lead an organization to a state of full information share and synchronism between marketing, sales, customer service, and support departments, reducing costs and improving efficiency. [Chen, I., & Popovich, 2003]

Since Knowledge Management (KM) is a multidisciplinary field of study that covers almost all of business activities, it's hard to describe what KM is. In its general business perspective, KM is "a collaborative and integrated approach to the creation, capture, organization, access and use of an enterprise's intellectual assets" [Grey, 1996].

In fact, KM is a set of strategies, tools and techniques that mixed together provide companies some abilities to (re)act in a faster and a more assertive way, such as learn from their history, stablishing good practices, and make decisions based on past situations and background information.

Integrating KM capabilities in CRM is "the utilization of knowledge for, from and about customers in order to enhance the customer-relating capability of organizations" [Dous et al., 2005].

KM plays a key role in CRM success once it represents the bottom of the pyramid where all CRM process frameworks are founded [Croteau & Li, 2013]. KM is responsible for collecting, storing and distributing relevant knowledge that will feed CRM processes. Furthermore, when

an organization tries to know its customers, their needs and their complaints, it can work for them and focus the business in the customer [Croteau & Li, 2013].

Knowledge Discovery in Databases (KDD) is “the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data.” [Fayyad et al., 1996]. By other words, KDD is the process to discover useful knowledge in data sets.

Despite having had its first breakthroughs in the late 80, KDD's researches have experienced a surge in the last decade as organizations gain a major concern about what data could tell about their business and how could they extract knowledge and value from it.

KDD became a biggest process than just data analysis with technology's help. The new wave of KDD refers to it as the overall process of discovering useful knowledge from data and it covers a set of techniques: 1. Selection; 2. Preprocessing; 3. Transformation 4. Data Mining; and 5. Interpretation/Evaluation.

After defining and understanding the application domain and the goals of the end-user, the first step of KDD process flow is to create a target data set – Selection. It is at Selection stage that analysts define a subset of variables or data samples on which discovery will be performed.

The next stage – Preprocessing – has the goal to clean the data identified previously in order to remove noise and outliers and prevent missing data fields.

At Transformation step is where analysts find useful features to represent data. After that, it is Data Mining the activity responsible for searching for patterns using the tasks and the algorithm that better answer to the relevant prior knowledge once the Interpretation of the results has been concluded [Han, Kamber, & Pei, 2006].

Data Mining is the fifth and the analysis step of KDD. It is the process of analyzing data from different perspectives and summarizing it into useful information. The main goal of data mining is to establish relationships among organizations internal factors, such as separate products and services transactions, and external ones, such as economic conjuncture, social and demographic indicators, etc. Due to that, Data Mining purpose it to set a certain level of abstraction and to present relevant information to support decision making.

Data Mining can be separated in two categories: Descriptive and Predictive data mining. Descriptive analytics describes the data set in a summary way. It is the simplest class of analytics and aims to condense big data into summary one. At a predictive level, the data goes through a variety of techniques, such as statistical, data mining, machine learning and modelling, to study historical and recent patterns, and to help analysts to make predictions about the future.

Data Mining has three main tasks: dig, discover and share. In its most simplistic way, data mining system get the data from multiple sources, processed it from various applications and feed a data storage, such as a database or a data warehouse. Then, the data mining tools can be used by all organizations players to discover trends and patterns and to provide them relevant information.

The typical data mining engine relies on a three stages flow [Linoff & Berry, 2011]: 1. Preparing data for mining; 2. Exploratory data analysis; and 3. Applying a model to the new data.

Despites exploratory data analysis is considered a data mining technique as its most primitive form, since it relies on a visual and interactive way of get information from data, it is commonly used as the first step to implement other data mining tasks.

One of the most important steps in data mining process e to choose the technique data most fits the answer that analysts want to get. There are four principal tasks [Linoff & Berry, 2011]: Classification and prediction; Estimation; Binary response; Finding clusters and patterns.

For classification and prediction, the algorithms that best fit are: Decision trees; Logistic regression; Neural networks.

If the goal is to estimate based on data then the models that should be used to that task are: Linear regression; Neural networks.

A binary response is when the dependent variable take only the values zero and one. For that purpose, the models that best fit are: Logistic regression; Decision tree.

At last, if the main purpose of the mining is to find clusters and patterns in data, then the data mining tool should implement any clustering algorithm.

“Marketing is all methods and facilities that an organization has to promote (...) behaviors to achieve their own goals.” [Lindon et al., 1997]

Marketing is the discipline of putting the right product in the right place, at the right price, at the right time and it’s based on four elements – The 4Ps or Marketing mix [McCarthy, 1960]: Product (or Service), Place, Price and Promotion.

Marketing can be separated in three strands: Research Marketing, Strategic Marketing, and Operational Marketing.

Research Marketing includes market research, competitor’s position, and monitoring the effectiveness of marketing campaigns [Bradley, 2013]. Strategic Marketing includes target segments definition, communication strategies, distribution channels, and associate products

and services conception [Cravens & Piercy 2008]. Operational Marketing includes advertising and promotion campaigns, post sales services, merchandising, and stock sale

In one hand, research marketing aims to get to know market position and to study the effectiveness of marketing strategies and operations of the organization[Bradley, 2013] [Cravens & Piercy 2008]. In other hand, strategic and operational marketing focus their efforts in getting into the market and gain customer interest. Strategic marketing sets the directions for achieving an objective and forms the basis of operational marketing action. Operational marketing focus the marketer's action to specific tasks, timetables and responsible personnel within the strategies, i.e. implementing strategic decisions [Bradley, 2013] [Cravens & Piercy 2008].

What if organizations were able to bring spatial location to their knowledge? And if they were able to know where their stores are located, where their competitors are, and where they can meet their customers? That is the main purpose of Geomarketing: bring together marketing and geography [Cliquet, 2002].

Geomarketing is a set of techniques that manipulate and geo-code data in order to introduce space into marketing decisions. It can be seen not as a short-term marketing strategy nor a campaign but a tool at marketer's disposal. Hereupon, Geomarketing is the use of location knowledge to frame marketing efforts, physically analyze questions relevant to marketing, sales and service issues, and use digital mapping to organize and display data for review and decision-making.

3. Geomarketing concept and process

As presented in the following table, Geomarketing also provides solutions to Marketing mix components.

Marketing mix component	**Applicability**
Product	• alignment of product supply to the features of a market segment or area of influence of a particular point of sale
Price	• product price adjustment depending on the sales strategy for a given area or segment
Place	• setting sales targets and maximizing growth opportunity based on the level of attractiveness of each segment;

	• investment optimization in retail outlets through research of optimal location due to the geographical distribution of customers and competitors; • optimization of distribution routes
Promotion	• costs reduction and increasing efficiency of an advertising campaign through publicity adapted to a given area or segment

Table 1 . Geomarketing mix and Applicability

The geomarketing planning process has the following phases:

1 – Identification of the problem;

2 - Project Planning;

3 – Collecting and Organizating Data;

4 – Exploring Data Analysis;

5 – Statistical Data Analysis.

4. Cases

To reveal the potential of geomarketing and how it works are presented below some scenarios and practical cases. The systems and technologies used to implement it are: ArcGIS for Desktop and ArcGIS Online.

Some of the data used to do the simulations were obtained from: Instituto Nacional de Estatísticas (INE, Statistics Portugal) and Instituto Geográfico do Exército (IGeoE, Army Geographical Institute Portugal).

ArcGIS® is a system produced by Esri (formerly, Environmental Systems Research Institute, Inc.) that provides an infrastructure for making maps and geographic information available throughout an organization, across a community, and openly on the Web.

Location Analytics is a product distributed by Esri Inc. and it proposes to add where to business data and systems. This product complements and extends the leading business analytics products, such as Cognos, MicroStrategy, SharePoint, and others, and provides a set of analytical tools and statistical models to improve spatial analysis of organizations and their business, asserting itself as a potent spatial decision support system.

To conceive this practical case, it was simulated an online store whose owners wanted to expand their business and to find a place for a physical store in Lisbon Metropolitan Area, Portugal. In five years of online commerce, they were able to store data about 5000 customers based on their profiles and orders, such as their residence, age, genre, and amount of purchases. They found themselves with a lot of disorganized data that didn't represent value to business and didn't help them to understand their position in market.

Since they had geographic information about their customer's distribution, they decided to organize all the data collected, to understand how the sales were distributed and how could they expand their business and be competitive in the market.

The first step of planning process is to identify the problem, i.e. define what must be achieved. Due to that, it is fundamental to set a group of questions that should be answered during the process. In line with scenario presented, the questions identified were: Where is a concentration of "important" or "less important" customers? What does the distribution of my customers look like? Where are my competitors located?

Answering to those questions, decision makers should be able to find out the solution to the problem proposed – where should the physical store be locate?

That's at project planning step where key variables, technologies, and methodology are defined. To the presented problem were identified the following key variables: (1) Total amount of purchases per customer : numeric; (2) Customer's age : numeric ; (3) Customer's geographic location : x, y coordinates; (4) Competitor's location : x, y coordinates; (5) Public transport's location : x, y coordinates.

The technologies used were: SQL Server 2008 R2, ArcGIS for Desktop, ArcGIS Online, Web AppBuilder for ArcGIS and Microsoft Excel.

The methodology followed matches a traditional KDD process: 1. Selection; 2. Preprocessing; 3. Transformation; 4. Data Mining; and 5. Interpretation and Evaluation

Data can be collected from various sources and stored in multiple data storages, such as databases and datasheets. In this scenario, data was collected from ordering system, which provided information about customer's (residence, age and genre), OpenDataLx Web Services from Câmara Municipal de Lisboa, which provided information about public transportation in Lisbon Metropolitan Area, Portugal, and Official Administrative Map of Portugal, distributed by IGeoE Portugal. Data about competitors was collected doing some research and identifying critical points in Lisbon Metropolitan Area.

After collection, data was georeferenced and stored in ArcGIS geodatabase2. To facilitate exploratory analysis, data was distributed in layers which allows spatial visualization of data, and the understanding of elements distribution on map. The exploratory data analysis provides the means to summarize data sets using visual methods. After data collection and organization, exploratory data analysis is used by analysts to explore data, find missing values, formulate hypothesis to experiments, detect outliers, and extract key variables to use in forthcoming statistical models.

For the presented scenario, exploratory data analysis was conducted in order to visualize: distribution of competitors and their areas of influence, using buffer generation with 750: meter radius.

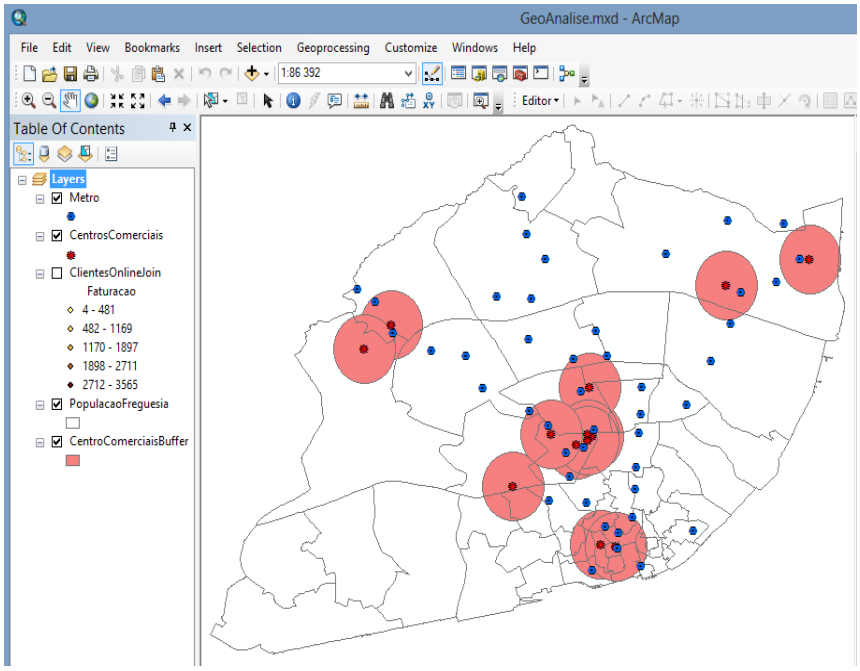


Figure 1 – Competitors (buffer 750 meters) and Public Transportation

The main concentration of customers and amount of purchases.

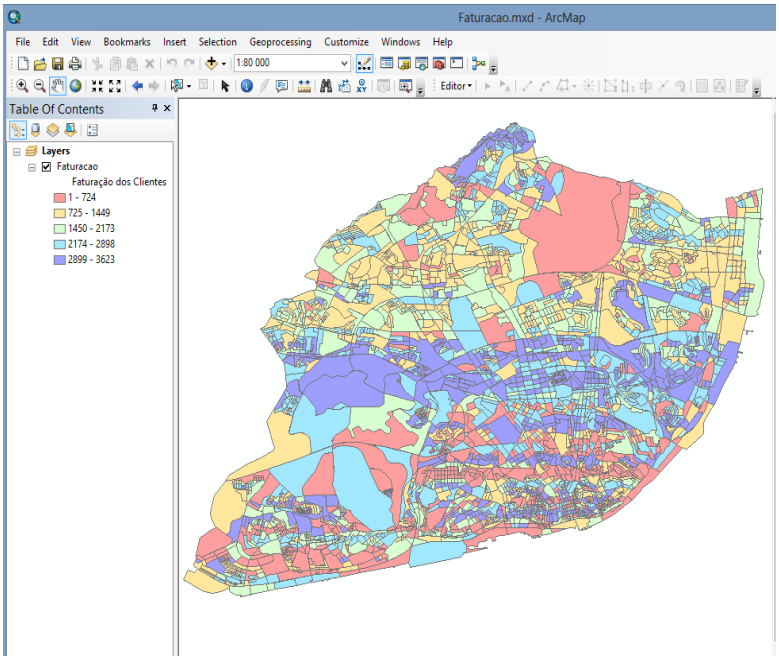


Figure 2 – Sum of total purchases (5 class breaks)

Patterns in customers and competitors locations.

At first, data was organized by administrative regions – Lisbon Metropolitan Area civil parishes – in order to group customers and other points of interest per area. After that, it was calculated sum, mean and standard deviation of amount of purchases.

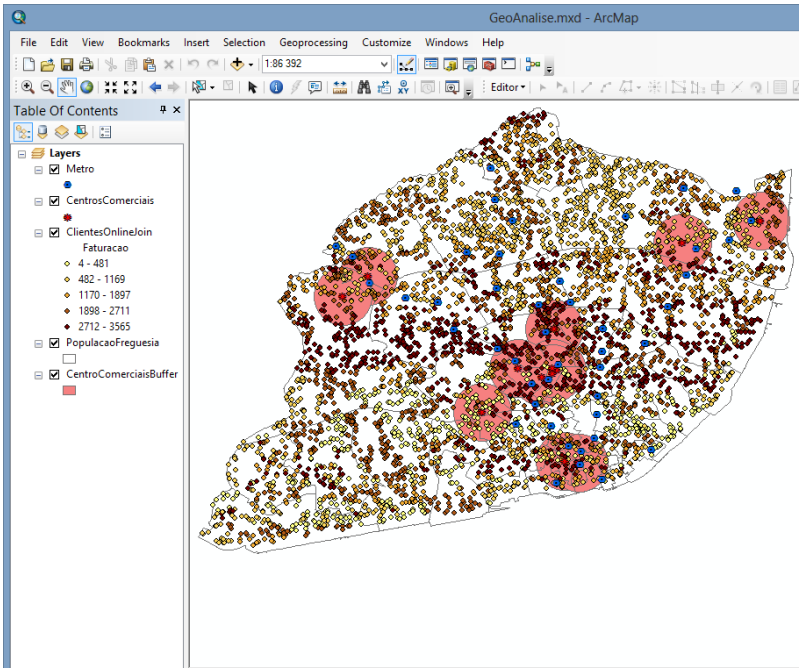


Figure 3 - Customers and Competitors spatial distribution

With data classified by region and amount of purchases, it was applied to data a point interpolation method using geometric break with 5 classes. Point interpolation performs an

interpolation on randomly distributed point values and returns regularly distributed point values, which can be used to predict new data points. For this scenario, interpolation method was performed to predict future turnover. By output, it returns a raster map.

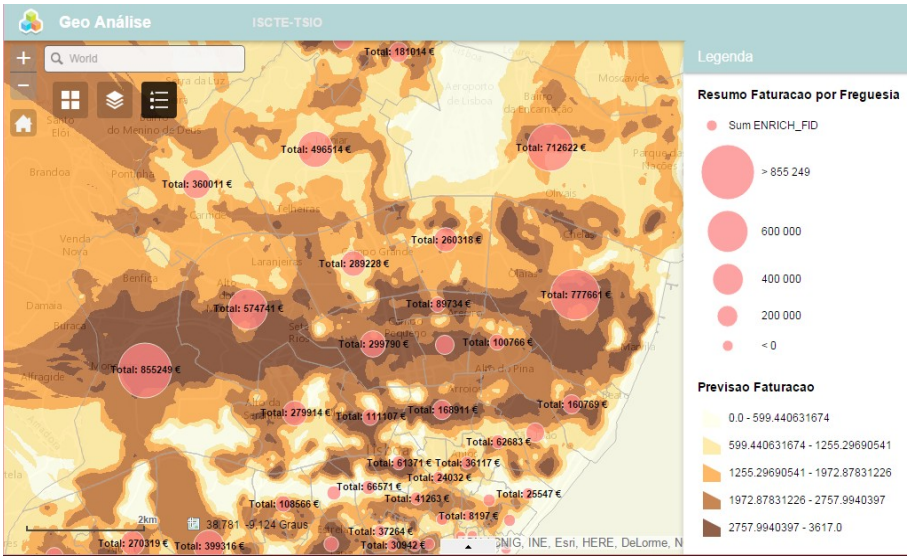


Figure 4 -Predicted turnover and amount of sales per administrative region

After that, I was perfomed a Hot Spot analysis which is a spatial analytic tool that uses Getis-Ord Gi* statistic to identify significant spatial clusters of high values (hot spots) and low values (cold spots) and to aggregate incident data – amount of sales.

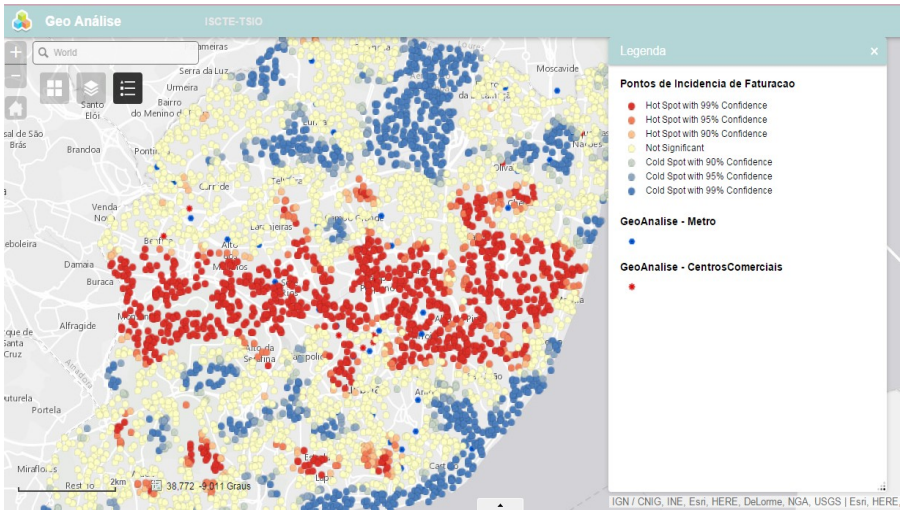


Figure 5 - Turnover HotSpot analysis output

5. Conclusions

After researching about CRM, Knowledge Discovery, Data Mining, and Marketing it became obvious the relevance of data analysis for organizations.

Organizations need to avaliate and understand the market, identify their competitors and know what their customers need. Due to that, organizations apply statistical methods to transform and analyze data in order to get valuable outputs that support future decisions.

Aligning geographic components with the most valuable Marketing techniques and CRM frameworks, allows organizations to know their position in market and understand the spatial distribution of their business.

6. References

Band, W. . *CRM in the Age of the Customer*. KMWorld. 2013

Bradley, N. *Marketing Research. Tools and Techniques*. Oxford University Press, Oxford, 2013

Chen, I., & Popovich. K. "Understanding customer relationship management (CRM) People, process and technology." *Business process management journal* 9.5, 2003, : 672-688.

Cliquet, G. . *Geomarketing: Methods and Strategies in Spatial Marketing*. ISTE Ltd. 2002

Cravens, D. & Nigel F.. *Strategic marketing*. McGraw-Hill Irwin, 2008.

Croteau, A., & Li, P. . Critical Success Factors of CRM Technological Initiatives. *Canadian Journal of Administrative Sciences*, 2003 21-34.

Desham, P. . “Spatial Decision Support Systems”. In D. Maguire, M. Goodchild, & D. Rhind, *Geographical Information Systems*, Volume 1 : Principles Longman 1991 pp. 403-412

Dous, M., Salomann, H., Kolbe, L., & Brenner, W.. “Knowledge Management Capabilities in CRM: Making Knowledge For, From and About Customers Work”. *Eleventh Americas Conference on Information Systems*. Omaha, NE, USA. 2005

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R.. *Advances in Knowledge Discovery and Data Mining*. AAAI. MIT Press. 1996

Grey, D. . What Is Knowledge Management? *The Knowledge Management Forum*. 1996

Han, J., Kamber, M., & Pei, J. . *Data Mining: Concepts and Techniques*. San Francisco, CA: Elsevier Inc. 2006

Lindon, D., Lendrevie, J., Lévy, J., Dionísio, P., & Rodrigues, J. V.. Mercator, *Théorie et pratique du marketing*. Paris, France: Editions Dalloz. 1997

Linoff, G. S., & Berry, M. J.. *Data Mining Techniques: for Marketing, Sales and Customer Relationship Management*. Indianapolis, IN: Wiley Publishing, Inc. 2011

McCarthy, J E. *Basic Marketing. A Managerial Approach*. Homewood, IL: Richard D. Irwin, 1960.

Gestão do Conhecimento nas Organizações
CAPSI/2015

Maria da Conceição de Mello Sampayo de Freitas [1], José Tribolet [2]

1) Instituto Superior Técnico, Universidade de Lisboa, Portugal

mc.freitas@sapo.pt

2) Instituto Superior Técnico, Universidade de Lisboa, Portugal
jose.tribolet@tecnico.ulisboa.pt

Resumo

Estimular novos conhecimentos nas organizações é atualmente um dos principais fatores de sucesso, pois permite a adequação, adaptação e a mudança das organizações num ambiente caracterizado pela globalização e cooperação de organizações à escala mundial suportadas pelas redes informáticas que têm vindo a evoluir a uma velocidade nunca antes permitida pela tecnologia disponível, não se prevendo a sua estagnação.

Hoje em dia uma empresa de sucesso deve ter os seus empregados focados na qualidade de serviço e no cliente tendo como objetivo as relações de longo prazo com os clientes, de forma a gerar novos conhecimentos.

As organizações que promovem o desenvolvimento dos empregados como indivíduos sociais potenciam o desenvolvimento do conhecimento e da competência nas organizações, criando vantagem competitiva através da incorporação de ativos de conhecimento nos seus processos de negócio.

Palavras chave: *Gestão do Conhecimento, Organização, Inovação*

1. Introdução

Com o crescimento da economia global e o declínio da economia industrial baseada em linhas de produção e controlo hierárquico, cada vez mais as organizações dão relevância ao conhecimento e competência para poderem competir e cooperar mundialmente [Hustad, Eli 2004].

Há mais de um século Taylor caracterizou o trabalho como manual, sem necessidade de trabalho intelectual, isto é, reduzindo-o a uma divisão clara de tarefas repetitivas e com uma clara separação entre o pensar e o fazer, desvalorizando o conhecimento humano e as capacidades sociais, enquanto as atuais teorias valorizam os colaboradores como um grupo social que nas organizações pode ajudar a melhorar significativamente a eficácia e eficiência da organização [Dietz, Jan LG, and Jan AP Hoogervorst 2011].

As atuais teorias consideram o envolvimento e participação dos colaboradores como núcleo crítico para uma mudança bem sucedida na organização tendo também em conta que a caracterização do trabalho na maior parte das organizações evoluiu de trabalho físico para trabalho intelectual que inclui criar, processar, integrar e aplicar conhecimento [Dietz, Jan LG, and Jan AP Hoogervorst 2011].

Numa sociedade onde o recurso primário de produção é o conhecimento que é criado e partilhado para aumentar a prosperidade e bem-estar das pessoas em substituição do capital ou mão-de-obra, é denominada uma sociedade do conhecimento [North, Klaus, Gita Kumta 2014].

Na gestão das organizações identificam-se três grandes objetivos ("Intelectual manageability", "organizational concinnity" e "social devotion" por forma a permitir às organizações enfrentarem os desafios com que são atualmente confrontadas [Dietz, Jan LG, and Jan AP Hoogervorst 2011].

- Intelectual manageability – Visão geral da organização e das suas mudanças;
- Organizational concinnity – Otimização e implementação de alterações com sucesso nas organizações operando como um todo unificado e integrado, contemplando todos os aspetos que se considerem relevantes;
- Social devotion – Importância do envolvimento e da participação dos colaboradores na produtividade da organização, qualidade do produto e serviço, orientação ao cliente, aprendizagem e inovação (e subsequente alteração da organização), bem como no desafio das dinâmicas da organização, e na complexidade de lidar com a incerteza dos desenvolvimentos emergentes da organização [Dietz, Jan LG, and Jan AP Hoogervorst 2011].

A aprendizagem individual dos trabalhadores não ocorre na maioria das vezes de forma planeada, mas é uma condição para a inovação, adaptação e mudança das organizações, daí emergindo o carácter de desenvolvimento de muitas organizações [Dietz, Jan LG, and Jan AP Hoogervorst 2011].

Um ponto fundamental na gestão do conhecimento é reconhecer ou redescobrir ativos de conhecimento na organização que não estão a ser utilizados com todo o seu potencial [Ndlela, Lorna Thembisile, and A. S. A. Du Toit 2001].

Neste documento iremos descrever as organizações como sistemas sociais com servidores de carbono e sílica, onde o conhecimento é preponderante na inovação e criação de novos serviços, produtos e estratégias importantes para as organizações. Abordaremos os fatores que potenciam o seu desenvolvimento nas organizações.

2. Servidores de carbono e de sílica

As organizações devem ser vistas como sistemas sociais, nos quais os seres humanos desempenham o papel de indivíduos sociais, com autoridade apropriada e correspondente responsabilização, podendo a organização também ser vista como uma rede de recursos ativos de dois tipos: servidores baseados em carbono (humanos) e servidores de sílica (computadores) [Tribolet, José, Pedro Sousa, and Artur Caetano 2014].

Neste contexto, podemos então definir uma organização o como uma única WEB semântica de servidores de sílica e carbono envolvidos dinamicamente em tempo real, não separável em duas independentes, uma constituída por humanos e outra por computadores, apenas existe uma única rede [Tribolet, José, Pedro Sousa, and Artur Caetano 2014].

As organizações devem ser vistas como sistemas dinâmicos, em constante evolução, adaptação e mudança, o que se encaixa na perfeição no modelo de Controlo de Sistemas Dinâmicos [Tribolet, José, Pedro Sousa, and Artur Caetano 2014]. Na verdade, temos de admitir que o fato das organizações existirem no mundo natural, então as leis da natureza aplicam-se-lhes.

O sistema da organização é composto por subsistemas elementares, que são os nós básicos da rede de atores independentes (atores baseados em carbono e em sílica) e que colaboram entre si bem como produzem comportamento autónomo alterando o estado global do sistema. Os servidores de carbono podem desempenhar várias atividades em simultâneo que não são facilmente separáveis das que são executadas pelos servidores de sílica [Tribolet, José, Pedro Sousa, and Artur Caetano 2014].

Cada uma das variáveis dos sistemas dinâmicos da organização pode ser classificada como:

- observável ou não observável;
- controlável ou não controlável.

Um sistema cujas variáveis não são observáveis não pode ser controlado, assim como, um sistema cujas variáveis não são controláveis também não pode ser controlado.

A teoria de controlo de sistemas também se aplica às organizações, sendo consideradas 2 hipóteses:

Hipótese 1: O servidor de Carbono (humano) tem o seu próprio Sistema Operativo, Canais de I/O, Memória e Processadores, e nele só pode executar aplicações que foram feitas para a sua configuração específica de hardware/software.

Hipótese 2: O servidor humano utiliza o modelo de Controlo de Kalman's.

Como em qualquer sistema dinâmico, o controlo implica feedback. A partir das observações das variáveis "observáveis", faz-se a atualização do modelo organizacional em execução e define-se a atuação controlada das variáveis controláveis.

É imprescindível desenhar as organizações de forma a poderem ser observáveis e controláveis o que justifica o recurso à Engenharia Empresarial.

O grau de previsibilidade das atividades de um servidor de sílica é maior do que o das de um servidor de carbono, no entanto, ainda devemos ter em linha de conta com outros fatores como erros e falhas e ainda a possibilidade de o sistema estar desalinhado com o negócio [Tribolet, José, Pedro Sousa, and Artur Caetano 2014].

3. O conhecimento nas organizações

Devido ao constante aumento da competitividade a nível mundial, as grandes organizações preocupam-se com a eficiência e eficácia para sobreviver no mundo dos negócios e dependem da qualidade do conhecimento que aplicam nos seus processos chave de negócio. Como resultado disto sujeitam-se a alterações drásticas para alcançar mais flexibilidade e devem produzir ativos de conhecimento criando vantagem competitiva e assegurando o negócio sustentado [Hustad, Eli 2004].

Até muito recentemente a modernização de processos de negócio implicava um controlo mais sofisticado devido ao aumento do conhecimento nas organizações e de tecnologia mais evoluída, no entanto, um maior nível de conhecimento pode levar a maior imprevisibilidade, incerteza e menos controlo. Por outro lado, o aumento dos mecanismos de controlo pode afetar os padrões de inovação na organização e destruir as ideias criativas emergentes [Hustad, Eli 2004].

A otimização dos processos de negócio na gestão tradicional baseia-se no pressuposto de que o investimento é recuperado através de um repetitivo modelo de negócio durante um período de tempo suficientemente grande [Sanz, J. L. C. 2013]. Atualmente, a mudança contínua e acelerada dos mercados leva a que o período de vida de um modelo de negócio seja cada vez mais curto, impossibilitando nalguns casos a recuperação do investimento através daquele processo.

De acordo com o princípio de Taylor deve-se, sempre que possível, aplicar a automatização dos processos de negócio e consequentemente dispensar o papel dos servidores de carbono na criação de valor nos processos de negócio [Sanz, J. L. C. 2013].

Atualmente, é muito valorizada a fidelidade de um cliente para com uma organização, pois está diretamente relacionada com a interação entre ambos (número de ações iniciadas, quantidade e tipo de informação trocada relativa a um serviço e ou duração do mesmo [Chew, Eng K 2014].

O grau de satisfação de um cliente é normalmente medido pelas suas interações através de um canal específico. Na maioria das organizações cada canal relaciona-se com o cliente como se fosse um compartimento estanque com estratégia, objetivos, processos de negócio, dados e tecnologia próprios, interagindo com o cliente de forma alheada e, por vezes, até desalinhada de uma estratégia global da organização, não promovendo assim a fidelização do cliente [Sanz, J. L. C. 2013].

A fidelização de um cliente a uma organização pode enquadrar-se num compromisso emocional, físico ou mesmo espiritual, dependendo diretamente da sua motivação para com o ambiente que lhe é oferecido pela organização [Chew, Eng K 2014].

O conceito de cocriação diz respeito a interação entre o cliente e os servidores de carbono da organização, de forma a sugerir inovação ao serviço prestado, isto é, permite à organização colher ideias, potenciando assim melhorias nos serviços disponibilizados [Chew, Eng K 2014].

Os clientes podem assumir três papéis distintos (passivo, informador ativo e criador bidirecional) na inovação, desenvolvimento e exploração de um serviço [Chew, Eng K 2014].

A tendência atual é disponibilizar aos clientes serviços cada vez mais personalizados, isto é, ir ao encontro das necessidades individuais de cada cliente, no entanto no atual estado da arte, as organizações não conseguem atingir este propósito, devido a limitações regulatórias, de escalabilidade e rentabilidade [Sanz, J. L. C. 2013].

Distinguem-se quatro caraterísticas nos serviços prestados a um cliente: intangíveis, heterógenos (os serviços possuem variáveis de input e resultados de desempenho não standards), inseparáveis (a produção e o consumo do serviço ocorrem simultaneamente) e perecíveis (os serviços não são duráveis nem guardáveis) [Chew, Eng K 2014].

O serviço é um sistema complexo adaptativo de servidores de carbono e de sílica que trabalham juntos para criar valor [Chew, Eng K 2014].

Os servidores de sílica podem contribuir para o serviço de quatro formas distintas [Chew, Eng K 2014]:

- Assistem o serviço – apenas os servidores de carbono têm contato com os servidores de sílica;

- Facilitam o serviço - clientes e servidores de carbono têm contato com os servidores de sílica;
- Medeiam o serviço - clientes e servidores de carbono estão geograficamente distantes e a sua interação é realizada pelos servidores de sílica;
- Geram o serviço – os clientes dispõem de self-service sem contato com os servidores de carbono.

As organizações podem utilizar o conhecimento extraído da relação histórica com o cliente para, de forma proactiva, melhor irem ao encontro da sua expetativa, disponibilizando serviços ajustados a cada caso [Setzer, Thomas 2014].

3.1. Conceito

A combinação do conhecimento existente e do novo conduz à inovação, pelo que a utilização do conhecimento existente tem um papel essencial na inovação contínua [Jaanus, Jörgen, and Tobias Ley 2013].

Existem 5 etapas consecutivas através das quais se dá o processo de maturação do conhecimento: expressar ideias, distribuir em comunidades, formalizar, aprendizagem ad-hoc e padronização [Jaanus, Jörgen, and Tobias Ley 2013].

O conhecimento que reside em documentos, relatórios, desenhos, plantas, modelos, standards, bases de dados e texto é considerado conhecimento explícito, pelo que pode ser compactado, codificado e transferido facilmente e consequentemente comunicado, partilhado e expressado facilmente em linguagem formal [Singh, A. 2010], pelo que os sistemas de gestão de conhecimento focam-se neste conhecimento explícito que pode ser muito limitado no que diz respeito ao que pode oferecer aos utilizadores [Reichling, Tim, Kai Schubert, and Volker Wulf 2005].

O conhecimento que refere a experiência académica, o saber-fazer, crenças, valores, perceções, e julgamentos ou opiniões é considerado conhecimento tácito. Este conhecimento é desenvolvido num dado contexto e considerado pessoal para cada indivíduo em particular, sendo específico de um servidor de carbono a sua existência ou disponibilidade muitas vezes é desconhecida, bem como, muito mais difícil de transferir, partilhar, expressar e comunicar [Singh, A. 2010].

A separação total entre o conhecimento tácito e explícito é impossível uma vez que o conhecimento tácito é uma componente necessária de todo o conhecimento, isto é "Know-how"

é uma capacidade para colocar os "know-what" em prática [Hustad, Eli 2004], dependendo do hardware e software dos servidores de carbono bem como do contexto em que estão inseridos.

Para permitir a utilização e partilha mais eficiente da informação na organização a gestão do conhecimento agrega pessoas, processos e tecnologias para tornar disponível o conhecimento certo para o processador certo (ex: servidor de carbono) e ao mesmo tempo com as apresentações adequadas para o preço certo [Singh, A. 2010].

A necessidade de organizar o conhecimento guardado e angariar conhecimento de uma grande quantidade de dados não estruturados levou as organizações a implementar ferramentas para gestão do conhecimento, que recolhem, organizam e partilham, analisam e disseminam o conhecimento. O valor das organizações depende da capacidade da organização de criar e gerir o conhecimento [Singh, A. 2010], que está intimamente ligado com os servidores de carbono e sua interligação, e que podem interagir com servidores de sílica.

A interação diária entre os servidores de carbono é realizada através da comunicação verbal e não verbal (gestos e modo do corpo são utilizados para transmitir uma mensagem sozinha e/ou em conjunção com a fala) [Hayes, Cory J., Charles R. Crowell 2013].

A organização é composta por comunicação, ação e informação onde os atores são gnomos vermelhos, verdes e azuis que respetivamente tratam de transações de negócio, infológicas e datalógicas, aqui também é referido os gnomos vermelhos (servidores de carbono), os únicos capazes de alterar o estado de produção no mundo, sendo que os gnomos verdes e vermelhos podem ser equiparados a servidores de sílica. Os servidores de carbono são indispensáveis numa organização pois são legalmente responsáveis pelas ações dos servidores de sílica [Dietz, Jan L.G. 2010]

3.2.Utilização e Partilha

O conhecimento difere da informação, pois é sempre pessoal e reside nas pessoas. "Saber" significa não apenas compreender ou acreditar, mas utilizar ou aplicar o conhecimento, pelo que o conhecimento da organização constrói-se e desenvolve-se dependendo da interação das pessoas, da cultura, comunicação e colaboração [Hustad, Eli 2004].

"O processo de transferência de conhecimento pode ocorrer a diferentes níveis; entre indivíduos, de indivíduos para fontes explícitas, de indivíduos para grupos, entre grupos, e de grupos para a organização" [Hustad, Eli 2004].

Com a globalização das organizações criam-se novas e diferentes perspetivas do poder, orientações, identidade e redes e os processos de negócio atravessam fronteiras [Caetano, Artur,

João Pombinho, and José Tribolet 2007]. Os processos de negócio são sujeitos a padronização, homogeneização, concentração e coordenação numa base mundial.

Apesar do acesso a tecnologias de comunicação avançadas é difícil ou quase impossível criar um sentimento de comunidade local entre pessoas que trabalham a grandes distâncias, para além das barreiras linguísticas e diferentes abordagens na tomada de decisão nos diferentes países [Hustad, Eli 2004]. "A transferência de conhecimento pode perder o significado intencional quando não é relacionado com o seu contexto ou origem" [Henderik A. Proper and Marc M. Landkhorst 2014].

Para que uma organização seja considerada uma entidade coletiva tem de ter um modelo semântico onde os seus servidores de carbono têm consciência de si próprios e do contexto onde estão inseridos [Henderik A. Proper and Marc M. Landkhorst 2014].

3.2. Atualização

Inovação é uma nova configuração do conhecimento cuja aplicação origina a criação ou melhoramentos de produtos, serviços, processos ou modelos de negócio [North, Klaus, Gita Kumta 2014].

A utilização do conhecimento permite um número infinito de opções para gerar novas e melhores formas de obter produtos e realizar serviços, ou otimizar processos ou modelos de negócio alinhados com a realidade, sendo muito mais difícil de imitar pela concorrência a capacidade de inovar instalada numa organização através do criar, combinar, transferir ou guardar conhecimento que é referida como capacidade de aprendizagem de uma organização [North, Klaus, Gita Kumta 2014].

As ferramentas de Gestão do Conhecimento permitem desenvolver a capacidade de aprendizagem de uma organização onde indivíduos, equipas e organizações bem como redes, regiões e nações criam, partilham e aplicam conhecimento de forma sistemática aumentando a eficiência e eficácia das operações através da inovação [North, Klaus, Gita Kumta 2014].

Uma organização para o ser, tem de ter um qualitativo clima organizacional, um senso coletivo de identidade e deve influenciar a aprendizagem e partilha do conhecimento por parte dos seus membros através das interações e transferência de conhecimento no grupo, no entanto, o processo de aprendizagem individual e coletivo não pode ser completamente dominado e controlado [North, Klaus, Gita Kumta 2014].

Para as organizações acompanharem as mudanças cada vez mais rápidas dos mercados com elevada velocidade de inovação e ciclos de vida de produto curtos, necessitam de serem mais ágeis e flexíveis na criação e alteração de processos de negócio, onde flexibilidade se refere à capacidade geral de reagir a mudanças e a agilidade ao tempo de resposta para implementar as alterações [Gong, Ywei and Marijn Janssen 2010].

4. Fatores que potenciam o desenvolvimento do conhecimento nas organizações

A criação de conhecimento numa organização corresponde ao papel de agregar, integrar e explorar numa forma ótima a criação de conhecimento que reside no plano individual [Jaanus, Jörgen, and Tobias Ley 2013].

No modelo de maturação de conhecimento existem forças de estabilização (exemplo: rotinas e processamento de informação estruturada e não estruturada) e forças dinâmicas que potenciam a inovação, pelo que, todas as organizações procuram um equilíbrio entre o desenvolvimento da inovação e a estabilização [Jaanus, Jörgen, and Tobias Ley 2013].

As organizações estão cheias de ferramentas IT que têm sido introduzidas nas organizações para formalizar e institucionalizar processos e práticas de manusear conhecimento, bem como a aprendizagem organizacional e gestão da inovação. Mas o conhecimento não é um estado que pode ser acumulado, preservado ou guardado, mas que é desenvolvido em múltiplas formas em redes pessoais e artefactos, pelo que um bom sistema de gestão do conhecimento está centrado nas pessoas, nas suas necessidades, preferências e na sua estrutura de conhecimento, em vez de forçar as pessoas a adotar sistemas padronizados [Jaanus, Jörgen, and Tobias Ley 2013].

Os servidores de carbono têm simultaneamente um papel na capacidade operacional, isto é, na execução periódica e repetitiva do dia-a-dia, bem como na capacidade de transformação, que tem a ver com a capacidade de inovar, enquanto os servidores de sílica apenas participam na capacidade operacional das organizações [Henderik A. Proper and Marc M. Landkhorst 2014].

Apenas para os servidores de carbono podemos falar no conceito de competência que se traduz em manifestações de conhecimento conseguidas através do desempenho de uma ação num contexto específico, isto é a ação certa no tempo certo [Caetano, Artur, João Pombinho, and José Tribolet 2007]

Na gestão da informação e conhecimento deve-se ter em conta três áreas de atuação [North, Klaus, Gita Kumta 2014]:

- Gestão estratégica do conhecimento que deve responder ao saber necessário para alcançar os objetivos da organização;
- Conhecimento operacional que trata da interligação entre a informação e o conhecimento de saber fazer, permitindo a conversão do conhecimento tácito em conhecimento explícito;
- A gestão de dados e informação de modo a guardar e distribuir informação que permita a criação e transferência de conhecimento.

5. Conclusão – Investir no conhecimento

As organizações para sobreviverem nos mercados cada vez mais competitivos e à escala mundial necessitam de diferenciar produtos e serviços personalizando-os, isto é, valorizando-os com conhecimento e satisfazendo as expectativas dos seus clientes. O conhecimento alocado a produtos serviços se não puder ser facilmente copiado pela concorrência, permitirá à organização obter vantagens competitivas a longo prazo e um melhor posicionamento no mercado.

As mudanças nos mercados ocorrem a velocidades cada vez maiores obrigando as organizações a se adaptarem e a inovarem, por outro lado, o desenvolvimento da tecnologia permite atualmente armazenar grandes quantidades de informação bem como a colaboração e troca de conhecimento entre pessoas localizadas geograficamente a grandes distâncias, o aproveitamento pelas organizações destas vantagens promovendo a capacidade de aprendizagem da organização bem como estabelecimento de mecanismos de eliminação de conhecimento irrelevante são fundamentais.

O desenvolvimento da tecnologia não substituí na íntegra a comunicação não verbal de proximidade entre os servidores de carbono que favorece um sentimento de comunidade local e que pode conduzir à perca do contexto e do caráter intencional da transferência de conhecimento.

A competência é a capacidade das pessoas em aplicar conhecimento para realizar a ação certa no tempo certo, sendo que o servidor de carbono só age se estiver motivado e esta ação depende do seu hardware, software, cultura e contexto onde está inserido.

4. Referências

Caetano, Artur, João Pombinho, and José Tribolet, "Representing organizational competencies." Proceedings of the 2007 ACM symposium on applied computing. ACM, 2007.

Chew, Eng K., Service innovation for the digital world. volume 9, German Informatics Society, 2014, 70-89.

Dietz, Jan L.G., Red garden gnomes don't exist. In RED GARDEN GNOMES DON'T EXIST IN SEARCH OF THE ESSENCE OF ORGANIZATION, Quantes / Artoos Communicatiegroep (www.quantes.nl), 2010.

Dietz, Jan LG, and Jan AP Hoogervorst. "Enterprise engineering manifesto." Advances in Enterprise Engineering I. LNBIP 10 (2011).

Gong, Ywei and Marijn Janssen, Measuring Process Flexibility and Agility, 2010

Hayes, Cory J., Charles R. Crowell, and Laurel D. Riek. "Automatic processing of irrelevant co-speech gestures with human but not robot actors." Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction. IEEE Press, 2013.

Henderik A. Proper and Marc M. Landkhorst. Enterprise architecture, volume 9, German Informatics Society, 2014, 5-21

Hustad, Eli. "Knowledge networking in global organizations: the transfer of knowledge." Proceedings of the 2004 SIGMIS conference on Computer personnel research: Careers, culture, and ethics in a networked environment. ACM, 2004.

Jaanus, Jörgen, and Tobias Ley. "Aligning Knowledge Development between Innovation-Driven Context and Knowledge Organization Systems." *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies*. ACM, 2013.

Ndlela, Lorna Thembisile, and A. S. A. Du Toit. "Establishing a knowledge management programme for competitive advantage in an enterprise." International journal of information management 21.2 (2001): 151-165.

North, Klaus, Gita Kumta Knowledge Management Value Creation Through Organizational Learning, Springer, 2014

Reichling, Tim, Kai Schubert, and Volker Wulf. "Matching human actors based on their texts: design and evaluation of an instance of the ExpertFinding framework." Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work. ACM, 2005.

Sanz, J. L. C. "Enabling customer experience and front-office transformation through business process engineering." Vienna: IEEE Conference on Business Informatics. 2013.

Setzer, Thomas, Data-driven decisions in service engineering and management. volume 9, volume 9, German Informatics Society, 2014, 106-117.

Singh, A. "Knowledge based expert systems in organization of higher learning." Proceedings of the International Conference and Workshop on Emerging Trends in Technology. ACM, 2010.

Tribolet, José, Pedro Sousa, and Artur Caetano. The role of enterprise governance and cartography in enterprise engineering, volume 9, German Informatics Society, 2014, 38-49.

On the evolution of hyperlinking

Artur Marques [1]

1) Escola Superior de Gestão e Tecnologia, Instituto Politécnico de Santarém, Portugal

artur.marques@esg.ipsantarem.pt

Abstract

Across time the hyperlink has supported distinct classes of applications and studies. This paper is one perspective on the evolution of the hyperlinking concept, context and behaviors.

In a travel across a spectrum of hyperlinking applications and practices, an effort is made to contrast the status quo with its related, broader, conceptual roots. The paper also bridges to some theorized and prototyped hyperlink variations, namely "stigmergic hyperlinks", in an effort to highlight that the ubiquitousness of some objects and their usage patterns might obfuscate opportunities to (re)think them.

In trying to contribute an answer to "what has the common hyperlink (such an apparently simple object) done to society, and what has society done to it?" the paper elects situations that have become more and more embedded into the everyday life, which does not help in spotting their roots and the chance for alternatives - maybe the travel can help.

Keywords: hyperlink, hyperlinking, stigmergic hyperlink, www, search

1. Introduction

Vannevar Bush was an electrical engineer, appointed MIT's Vice President in 1932, and then head of the OSRD (U.S. Office of Scientific Research and Development) during World War II. OSRD was the office which supported scientific and military projects such as the atomic bomb / "Manhattan Project" and the radar. In 1945 Bush was already asking what significant problems would researchers face after the war, and answering the question with the challenge of how to promote cross-awareness of knowledge and intellectual production in different scientific areas, because the then existing "methods of transmitting and reviewing the results of research" were, in his perspective, "generations old" and "totally inadequate" [Bush 1945, Bush 1996]. Bush considered that conventional indexes/citations embody such a distance to the originals/cited works that they are creativity-unfriendly. Since the mind "operates by association", a natural solution would be to create links between recorded ideas in context.

Bush's plan of a "memex" machine and its ability to link any points between microfilmed documents is the hyperlink precursor which inspired Theodore Nelson, Douglas Engelbart, and others, even if using a much different memory support and an association representation that

most current users couldn't begin to imagine because so many of us take the contemporary hyperlink for granted, as a dead simple interface to something that sometimes is even no longer perceived as a problem. Still the original association challenge persists and the current WWW hyperlink usage is happening on a massive and industrialized scale from which emerge hyperlinking motivations and hyperlinking use cases that strongly deviate from plain citation and navigation, let alone knowledge association.

The common hyperlink is not a one-size-fits-all device, has limitations, can disappoint even its creators in certain ways and, in certain scenarios, can fall victim to its own success.

One challenge in researching, extending, changing or just thinking about the hyperlink is how deceptively simple "it" appears to be. However, when looking past the surface layer, one should gradually be able to sense its depth and great importance in today's world, as envisioned by scientists such as Vannevar Bush from the mid-1940s and now voiced in publications as generalist as the New York Times Magazine, all suggesting the concept and the object as one of all times' greatest inventions.

Besides its everyday star role on the Web, its place in computer science in general and hypertext in particular, the hyperlink as an entity and as an interface mechanism doesn't rest confined to such fields. Hyperlinks have become a pervasive tool that "changed the fundamental dynamics of human communication" [Hespos 2008] be it in information sciences, political sciences, sociology or media studies, etc., usually exposing the same familiar attributes and functionality, which makes them easy to use but also somewhat conditioning because of their impregnated usage pattern that makes it hard to think of them differently, "out of the box", in particular when required to do so without disruptive consequences for the users. "Hyperlinks are more than technical artefacts" [Maeyer 2013] supporting a gamut of situations in a way which could eventually be different if other attributes and behaviors were also available.

Some hyperlinking behaviors may now cause frustration - e.g. broken links and invasive ads publishing - or even social damage, as in phishing attacks. Solutions such as "Stigmergic hyperlinks" [Marques and Figueiredo 2010], or "stighs", can be of immediate help in providing relief to some of the "bad" outcomes of the common hyperlink: in particular, they constitute a fully automatic response to broken links in pages. While many solutions exist for identifying broken hyperlinks and other structural issues [Ingham, Caughey et al. 1997, Haslhofer and

Popitsch 2009, Morishima, Nakamizo et al. 2009, Kovilakath and Kumar 2012], very few exist for helping in fixing the problem. The relatively scarce responses to identify *and* solve the broken links usually follow a recommender system approach, which depends on users to rank alternative substitute resources, before committing to a fix. Such approach is applicable to all situations where indeed there is a replacement to be provided and human labor available to be invested. The fully autonomous solutions, which identify and solve broken references without human assistance, are much rarer and limited to specific domains.

However, it is important to understand that such solutions are not intended as replacements of the common hyperlink! They are only adequate for authors with particular needs, e.g. wanting to have "publish and forget" autonomous systems, wanting to have web metrics at the hyperlink-level or wanting to provide different contextual "searches". The hyperlinking behavior here discussed is mostly based on the common hyperlink.

The hyperlink's ubiquitousness would only come with HTML and the WWW. Such is the penetration of the hyperlink as structured in HTML that it became hard to conceive it in a different dress, with other attributes and other interactions. Despite everyday problems - e.g. how easy it is to "phish", the dependency on browsers for travelling back to sources, the unverified destination pointer and the dead link disruption [Leuf 2006] - the dominant hyperlink solution remains mostly unchanged and unchallenged, for reasons that include its simplicity.

At a lower level, the everyday hyperlink does serve for citation and navigation, but on a higher level those purposes are being leveraged for serving other intentions not immediately inferable from the root design, namely "gatekeeping" sites, "computing" contents and "measuring" or "tracking" users. These functions are byproducts of the underlying business, technology and marketing forces. The hyperlink is literally the entity, by which some of the world's biggest organizations live or die, from search-engine based businesses to the contents industry, to the entertainment world in general: failing to link or to be linked, regardless of other core merits, can push one into dire straits.

This paper first delves into some motivations for hyperlinking, across time; then it approaches the subject of "Web Metrics" for the understanding of the hyperlinking behavior. It concludes with one perspective on the consequences of the current hyperlinking practices.

2. Motivations for Hyperlinking

Why do authors create hyperlinks? What are their possible motivations?

This is an active research question, with partial answers that vary across the timeline: the main reasons for hyperlinking in the mid-1990s may be different from the ones during the blogs' boom of the mid-2000s, and are probably distinct from what drives authors and publishers today.

In [Ciszek and Fu 2005] the authors share the perspective that hyperlinks "have evolved beyond being citation knowledge tools to become social linking mechanisms". The statement captures that in the early Web many hyperlinks were plain citations-as-references, from a document to other documents or to parts of itself. If hyperlinks are authored with a citation functionality in mind, then some identified motivations [Kim 2000] are: to guide the reader to background information, to provide an example, to give supportive evidence, to provide a definition, for historical background, or to compare with an alternative. Also implicit in the statement is that other factors, beyond those identified in academic referencing, would determine hyperlink authoring.

[Thelwall 2003] argues that "academic hyperlink creation" and "citations in journals" can however be "very different phenomena" from a scholar's professional point-of-view, in the sense that the hyperlinks present in university Web pages can exist for purposes other than assisting the reader, namely connecting to other institutions, employers, partners and personal bookmarks. The author presents a classification which distinguishes "ownership links" (for acknowledging authorship or co-authorship), "social links" (those with a social reinforcement role), "general navigation", and "gratuitous" links (those serving no communication function at all). This taxonomy is limited to links in academic Web pages.

In [Miles-Board, Carr et al. 2002] the authors identify the most frequent "linking practices" in "genuine associative-linked pages", after narrowing an initial sample of 770,992 random Web pages to 576 meeting certain criteria of link density and link distribution, in order to avoid dubious or non-associative content. They conclude that "reference links", which tie nouns such as names of people, institutions, products, places, etc., to corresponding resources are the most

frequent, occurring in 84% of the filtered "well-linked pages"; followed by "deep links" which associate keywords or concepts to deeper explanations (37%); followed by "glossary links" which connect alphabetical arranged terms (35%); followed by "citation links" as in endnotes (32%); followed by "structural links" which relate parts of content, for example performing navigation between sections, or pointing to figures (17%). It should be noted that these conclusions are specific to what the authors call the "well-linked pages" and not to all pages in general; if no filters were in place, then "structural links" would probably be the most observed, because while publishers can make the editorial control option to rarely have explicit external links, as some newspapers did [Barnhurst 2002, Dimitrova, Connolly-Ahern et al. 2003, Tremayne 2005], there is no escape from navigational structures, if the site is to be usable.

Expectations of reciprocity, as in B expecting a link from A because B links to him, is one emergent social behavior around the hyperlink [Adamic 2008, Adamic 2009]. "Emergent" because nothing in hyperlink authoring implies such consequence; still there was an époque of "guestbook" proliferation, pages of "favorites ", and "circles" of sites on a common theme, and there are modern equivalents in social networks.

Besides citation, hyperlinks always supported navigation. Navigation became less innocent as the Web grew in its number of consumers and producers. At first, navigation enabled nothing but a guiding structure to "consume" a site's available contents and "teleport" the audience to other cited content providers. Then, the growing number of Web content publishers increased the awareness that visitors might leave and never return, and that one could be left "talking" alone. For authors who were just enjoying the technology and whose drive didn't depend on an audience nothing changed, but for all the others who wanted readers and/or viewers, it became clear that it would be necessary to manage processes of acquisition and retention of visitors.

Hyperlink authoring would become less spontaneous; authors started having second thoughts before publishing external links: what if the destination was badly perceived? Would the negative perception harm the publisher somehow? What if the destination was favorably perceived? Could it mean traffic lost for a rival? What if the external resource is moved or deleted? These questions point that when audience matters, there is value in traffic volume and traffic quality.

Value means money: faster than many established business models could adapt, the Web was hosting old and new businesses. With money at play - a powerful motivator - accounting leapfrogs in relevance. Accounting for the Web means "Web metrics" which can be considered as the measurement of Web properties.

In leveraged applications, the hyperlink can translate to a sale in commerce, to a subscription, to a donation, to an advertisement or to valuable information about visitors. These activities correspond to all the possible revenue streams in any type of business [Krishnamurthy 2003], thus they constitute hyperlinking with a monetization motivation. Some revenue streams which exist since the dawn of money but could not scale up to support big operations, namely donation-based and freemium models, can now deliver in the WWW-based hyperlinking society.

The monetization possibilities of the hyperlink are the cause of its industrialization, happening in a "marketplace of attention" [Webster 2008]. This is a dichotomized marketplace: as a whole, in providing ever more contents, Web publishers are a force of attention fragmentation; however people have a tendency to socialize and consume with and what they are already familiar with, which is a factor of attention concentration/polarization. Combining the bias for attention polarization with the winner-takes-all form that the public WWW is assuming in a number of situations (in e-commerce, in Web search services, in blogging, etc.) contributes for information "enclaves", pushing in the opposite direction of fragmentation.

The marketplace of attention happens in a context of "convergence", "abundance" and "scarcity". Convergence because all content, new and old, is being made available in the same global network. "Abundance" because the volumes of original content and new content being imported from other media grows by numbers which can only be abstracted. "Scarcity" because the number of hours in a day hasn't changed and, in potential, "a wealth of information creates a poverty of attention" [Simon 1971]. These long established ideas of the challenges posed by an abundance of information have never been more pertinent, because "the real scarce resource will always be human attention and attracting that attention will be the necessary precondition of social change" [Cowen 2007].

This is why any filters between content producers and content consumers have become so important: the market is so competitive and the human attention so relatively inelastic that, from a producer's perspective, missing a chance for visibility by being filtered-out, or getting a chance for visibility by being filtered-in, can be decisive. On the same token but from the consumption side, failing to filter-in preferred information or to filter-out less desirable alternatives, can translate to a bad attention investment - the scarcest of the investable resources in the new economy. Therefore, the power of search engines such as Google, who perform as filtering devices in the marketplace of attention, is just immense and the public space can benefit from other devices and other search possibilities.

In automatically eliminating ill-favored links, systems of stigmergic hyperlinks function as filtering devices in the market of attention. They are valuable in assisting communities of users who consistently express preferences in the same polarizing direction. On the other hand, if the consumers/users don't really constitute a community, by having fragmented interests; or the producers/publishers aren't being able to attract a stable audience, then fragmentation takes place and the effects of running this and other equivalents of recommender systems can be erratic.

3. Web metrics

In [Dhyani, Ng et al. 2002] the authors classify the metrics for the Web in six categories: "Web graph properties", for quantifying structural properties of networks; "Web page significance", for formalizing the concepts of "quality" and "relevance"; "usage characterization" for measuring user behavior; "Web page similarity" for quantifying the relatedness of pages; "Web page search and retrieval", for comparing different search and retrieval services; and "information theoretic", for capturing properties of information needs, production and consumption.

When using a hyperlink, users manifest behaviors which can be measured with tools that fit the "usage characterization" category; on the long run, it should be possible to better understand their information preferences with some "information theoretic" reasoning. For most hyperlinking use-cases, these are the most relevant categories, but it is almost consequential to put other types of measurements to work: for example, when a system of stigmergic hyperlinks is to replace one link whose vitality nulled, it computes its replacement from the network of "survivors". There is a chance for increased "Web page similarity" across the linked

destinations, strictly based on the nodes' proximity, yet there was no content analysis and the network graph emerges from basic reward/decay rules, in bio-inspired stigmergic fashion. Identifying the survivors and their hyperlinks is a graph based measurement.

Behind the scenes, most clicks now spark computations based on data ranging from the basic variables provided by every client-server environment (local time, IP address, referrer, etc.) to the complex, specific to each application's purpose.

Measuring hyperlinking in the "Deep Web" is much harder [Bergman 2001, He, Patel et al. 2007]. The retrieval of deep webs where different contents can have the same URL is a challenge that cannot be approached with horizontal solutions dependent on a 1:1 correspondence between address and content, requiring very specific approaches or a radically new general search engine which can index content not only by node address, but also by trails of user interactions which open the content: I call it the trails-based search engine ("TB search engine"). There is no such solution yet and the consciousness of its need won't reach critical mass while the information needs of most people are being satisfied with the current services, who deliver in quantity and with a quality that can't be contrasted with alternatives that don't exist. However, some technical client-side factors for hidden data, such as content induced Javascript, have begun to be handled - see the "CrawlJax" tool by [Mesbah, Deursen et al. 2012]. Others are working on the front-end of search, trying to think tomorrow's Web with more engaging "interactive visualizations" [Dork, Williamson et al. 2012].

4. Some consequences of the status quo

Such is today's hyperlinked society that an entire new industry was created to support the contemporary hyperlink needs and wants, and research new ones. At the core of this industry, Google is probably the strongest example of a hyperlink-based business. Most of Google's revenue comes from ad selling, with ads being delivered via sponsored search results presented as hyperlinks, or via the AdWords + AdSense programs. AdWords enables do-it-yourself Web campaigns, while AdSense enables participants to be part of an ad space network. One fundamental difference between selling ads online and selling ads on more mature media like newspapers, TV and radio, is that the Web is currently the only medium with a prevalent two-stage business model for advertisements, charging the advertiser not for the mere advertisement display, but only for subsequent interactions with it, in the form of clicks. In other media, the business model charges per audience quantity and not per interaction. Not only there is this

significant difference, but the available Internet technologies provide real time, finer granularity, measuring tools that have the potential to support more detailed, better informed, more transparent campaigns.

Google, online ad networks in general, and sponsored web search results are examples of relatively new businesses. Orbiting the search engine players yet on the same marketplace for attention, a myriad of other organizations sell services which were unimaginable a few years ago: there are markets for domain names, infomediaries in buying/selling, appraisal services for all types of content for reasons which include placing higher in search engines' rankings, security services, and many other neo offerings, including new educational curricula tailored for the hyperlink-based economy. In all the diversity there is one common goal: a "better" hyperlink, with "better" having many subjective meanings, e.g. "shorter, easier to remember", "better placed, better redacted, worthy of higher rank", "immune to link rot", or "less prone to phishing".

The point is that while the hyperlink is always at the center, there are many different, sometimes contradictory, reasons for its publication. The industry is gaming the hyperlink: for example, when a Google search returns some results, it is doing its algorithmic best to deliver value to the user; but, at the same time, some firm in the Search Engine Optimization (SEO) business, will try to push its customer's site higher and higher, for certain search expressions, regardless of alternatives, playing every artificial hyperlink trick in the book, if necessary. There is a huge disproportion between the scale and objectivity of these tasks: Google's is massive and spread, trying to serve millions, while the SEO player's is singular and focused, with a single client to please. The resources available to each are even more disproportionate; in the end this is a cat and mouse situation with two very different motivations for hyperlink publishing. Every time such extreme disproportions appear, as in this particular hyperlink use case, one has to ask: is there room in the middle?

As elaborated in [Marques 2014], there might indeed be room in the middle. One idea is to reinforce the original association purpose of the hyperlink by enabling custom in-context search solutions. However, in the current "search environment", heavily absorbed by a handful of large scale solution providers, it is very hard for these topical solutions to find a place. The need for vertical solutions is sometimes not perceived at all because of the amount of already available

indexed content - we can only pay attention to so much, and there is a reasonable chance that in the already charted WWW there is more than enough to distract us from unfulfillment. The industry knows our human limits and comfort zones: for example, the default 10 search results per page that most search engines will display is an option which finds support in TV research suggesting that, no matter how many content channels might be available, the majority of viewers will only manage and consume 10 to 15 of them [Jeffres 2001], at most.

And who may be those authoring custom search/hyperlinking solutions? Following Benkler's [Benkler 2006] view of "ideal-type information production strategies", they'll probably approach, but not necessarily match, the "Joe Einsteins" or the "Romantic Maximizers", depending if their main driver is "fame" or "fortune", respectively. Benkler thinks of "Joe Einsteins" as those willing to "give away information for free in return for status, benefits to reputation, and value of the innovation to themselves". The "romantic maximizers" are authors who sell their creations, for example to publishers. In both cases issues of intellectual property can make the situation less clear because Joe's "free" solution might infringe unanticipated publisher rights which can force him to have to charge to be able to pay, or desist; and the romantic's product might violate copyleft licenses which force him to open his code and to have to adapt to a different revenue model, for example support-based.

It is worth noting that "open source software" and "free software" are legal concepts, corresponding to software with a license in which the copyright holder provides at least the rights to use, modify, copy and distribute the software, to anyone, for any purpose. Some licenses "attach" to derivative works (copyleft licenses) so that future versions and/or branches must reciprocate the conditions the author first received. The information economy is still relatively young and many questions of intellectual property haven't been ruled in court at all; some decisions will influence information products and services [Lindberg 2008].

If the most fundamental object in hypertextuality could exhibit a customizable behavior, then hyperlink authors and publishers would have reasons to think beyond the citation & navigation seminal purposes, beyond the current tracking and ad selling functions, and could even rethink what "search" is: search does not have to be literally about finding something at a destination resource; it can be also about identifying alternatives to the resource, or providing richer visualizations. For example, imagine yourself reading some article that, at a point, links to somewhere with an anchor like "<movie name here> is frightening"; in today's WWW the usage pattern creates the expectation that clicking the anchor will bring you to some site about the film

and that "searching" it is to look for some expression there, but the search behavior could handle the string differently, e.g. as a request to suggest/compare with other frightening movies using some fright-scale, or to produce some representation of classifiers associated with the movie, highlighting the searched one; or to find community discussions in that context. So, there are alternatives to centralized horizontal solutions, there are alternatives to literal textual searches [Dork, Williamson et al. 2012], and there are behavioral possibilities besides the vanilla hyperlink's features.

Hyperlinking is just beginning, but has already become so embedded in our lives that having an awareness of how we got here and being able to escape established patterns should be relevant.

5. References

Adamic, L. (2008). The social hyperlink. The Hyperlinked Society.

Adamic, L. A. (2009). The social hyperlink. Proceedings of the 20th ACM conference on Hypertext and hypermedia. Torino, Italy, ACM**:** 1-2.

Barnhurst, K. G. (2002). "News Geography & Monopoly: the form of reports on US newspaper internet sites." Journalism Studies **3**(4): 477-489.

Benkler, Y. (2006). The Wealth of Networks: How Social Production Transforms Markets and Freedom, Yale University Press.

Bergman, M. K. (2001). "The deep web: Surfacing hidden value." Journal of Electronic Publishing **7**(1): 07-01.

Bush, V. (1945). As We May Think. The Atlantic.

Bush, V. (1996). "As we may think." interactions **3**(2): 35-46.

Ciszek, T. and X. Fu (2005). "An annotation paradigm: The social hyperlink." Proceedings of the American Society for Information Science and Technology **42**(1): n/a-n/a.

Cowen, T. (2007). "Richard A. Lanham, The Economics of Attention: Style and Substance in the Age of Information." Journal of Cultural Economics **31**(2): 167-169.

Dhyani, D., W. K. Ng and S. S. Bhowmick (2002). "A survey of Web metrics." ACM Comput. Surv. **34**(4): 469-503.

Dimitrova, D. V., C. Connolly-Ahern, A. P. Williams, L. L. Kaid and A. Reid (2003). "Hyperlinking as Gatekeeping: online newspaper coverage of the execution of an American terrorist." Journalism Studies **4**(3): 401-414.

Dork, M., C. Williamson and S. Carpendale (2012). "Navigating tomorrow's web: From searching and browsing to visual exploration." ACM Trans. Web **6**(3): 1-28.

Haslhofer, B. and N. Popitsch (2009). DSNotify - Detecting and Fixing Broken Links in Linked Data Sets. Proceedings of the 2009 20th International Workshop on Database and Expert Systems Application, IEEE Computer Society**:** 89-93.

He, B., M. Patel, Z. Zhang and K. C.-C. Chang (2007). "Accessing the deep web: a survey." Communications of the ACM **50**(5): 94-101.

Hespos, T. (2008). How Hyperlinks Ought to Change the Advertising Business. The Hyperlinked Society: Questioning Connections in the Digital Age. J. T. L. Tsui, The University of Michigan Press**:** 137-144.

Ingham, D. B., S. J. Caughey and M. C. Little (1997). "Fixing the" broken-link" problem: The W3Objects approach." TECHNICAL REPORT SERIES-UNIVERSITY OF NEWCASTLE UPON TYNE COMPUTING SCIENCE.

Jeffres, K. A. N. D. J. A. L. W. (2001). "Reconceptualizing Channel Repertoire in the Urban Cable Environment." Journal of Broadcasting & Electronic Media **45**(3): 464-482.

Kim, H. J. (2000). "Motivations for hyperlinking in scholarly electronic articles: a qualitative study." Journal of the American Society for Information Science **51**(10): 887-899.

Kovilakath, V. P. and S. D. M. Kumar (2012). Semantic broken link detection using structured tagging scheme. Proceedings of the International Conference on Advances in Computing, Communications and Informatics. Chennai, India, ACM**:** 16-20.

Krishnamurthy, S. (2003). E-Commerce Management : Text and Cases / Slatanic - the disastrous move from free to paid content, Thomson / South-Western College Pub.

Leuf, B. (2006). Enhancing the Web. The Semantic Web - Crafting Infrastructure for Agency, Wiley**:** 3-30.

Lindberg, V. (2008). Intellectual Property and Open Source, O'Reilly Media.

Maeyer, J. D. (2013). "Towards a hyperlinked society: A critical review of link studies." New Media & Society **15**(5): 737-751.

Marques, A. (2014). Stigmergic Hyperlinks' contributes to Web search. CAPSI. Santarém, Portugal.

Marques, A. and J. d. Figueiredo (2010). "Stigmergic Hyperlink: a new social web object." IJISSC - International Journal of Information Systems and Social Change **2**(4): 31-43.

Mesbah, A., A. v. Deursen and S. Lenselink (2012). "Crawling Ajax-Based Web Applications through Dynamic Analysis of User Interface State Changes." ACM Trans. Web **6**(1): 1-30.

Miles-Board, T., L. Carr and W. Hall (2002). Looking for linking: associative links on the Web. Proceedings of the thirteenth ACM conference on Hypertext and hypermedia. College Park, Maryland, USA, ACM**:** 76-77.

Morishima, A., A. Nakamizo, T. Iida, S. Sugimoto and H. Kitagawa (2009). Bringing your dead links back to life: a comprehensive approach and lessons learned. Proceedings of the 20th ACM conference on Hypertext and hypermedia. Torino, Italy, ACM**:** 15-24.

Simon, H. A. (1971). Designing Organizations for an Information-Rich World. Computers, communications, and the public interest, Johns Hopkins University Press.

Thelwall, M. (2003). "What is this link doing here? Beginning a fine-grained process of identifying reasons for academic hyperlink creation." Information Research **8**(3).

Tremayne, M. (2005). "News Websites as Gated Cybercommunities." Convergence: The International Journal of Research into New Media Technologies **11**(3): 28-39.

Webster, J. G. (2008). Structuring a Marketplace of Attention. The Hyperlinked Society. J. T. L. Tsui, University of Michigan Press**:** 23-38.

Tecnologias Informáticas em Organizações de Saúde: o que ainda deve ser repensado

Luís Costa [1], Alberto Carneiro [2]

1) ISCTE-IUL, Portugal

lmpcost@gmail.com

2) Universidade Autónoma de Lisboa & Universidade Europeia, Portugal
albert.carneiro@netcabo.pt

Resumo

Os Sistemas de Informação (SI) em organizações de saúde têm um elevado valor na vida dos cidadãos e nas políticas de saúde. A introdução das Tecnologias de Informação e Comunicação (TIC) como tecnologias de suporte deve considerar preferências dos diversos atores envolvidos e necessidades de informação dos diferentes serviços e possibilitar a partilha de informação entre estes, para garantir a integração dos vários sistemas. Perante a complexidade de problemas e insuficiências que esses sistemas ainda revelam, é necessário repensar as potencialidades das TIC e procurar novas soluções de modo coordenado e integrado. Este artigo pretende ser uma reflexão sobre os problemas ainda existentes no que respeita à contribuição das tecnologias informáticas para a implementação de estratégias que possibilitem a obtenção de melhores níveis de prestação de cuidados de saúde para todos e de forma sustentada. São apresentadas várias propostas que têm por finalidade aperfeiçoar linhas de desenvolvimento dos SI em organizações de saúde.

Palavras chave: TIC, organizações de saúde, sistema de informação

1. Introdução

É surpreendente o impacto da informática e das suas diversas tecnologias na prática da medicina. As técnicas não invasivas de produção de imagem, como a ultrassonografia, a medicina nuclear, a tomografia axial computorizada (TAC) e a ressonância magnética, alteraram muito os processos de diagnóstico médico. Os novos equipamentos digitais para o processamento de sinais biológicos e para a monitorização de doentes, como a videolaparoscopia e os analisadores automáticos de eletrocardiogramas, de fluxos sanguíneos e gasosos, globais e regionais, realizam a oferta de informações vitais que ajudam os médicos de diversas especialidades, quer na eficácia das terapias, quer no apoio à investigação.

Na prática médica diária, as tecnologias informáticas ajudam na recolha, no registo e na análise de dados, na definição de terapêuticas com medicação, nas relações com os utentes e com outros médicos. Os Sistemas de Informação (SI) em saúde podem apoiar a monitorização do processo de assistência à saúde e têm contribuído para a melhoria da qualidade da assistência ao utente na

medida em que auxiliam o processo de diagnóstico ou a prescrição de medicamentos e respectivas posologias, possibilitando a inclusão de indicações clínicas para o acompanhamento da assistência, de avisos sobre efeitos secundários dos princípios ativos, das possíveis interações e de protocolos clínicos.

A definição de SI em cuidados de saúde é ambígua. No entanto, é usualmente referida aos sistemas de Tecnologias de Informação e Comunicação (TIC) infraestruturas que são usados na prestação de cuidados médicos. Hospitais, a indústria da saúde, a indústria farmacêutica, as organizações de cuidados de saúde e os respectivos fornecedores estão entre as entidades importantes ligadas às infraestruturas de TIC nos SI no domínio da saúde [Srivastav et al. 2011; Mantzana et al. 2007; Gans 1997]. Estes SI relatam o papel dos doentes, dos profissionais médicos e não-médicos no manuseamento da informação relativa aos cuidados de saúde [Tähkäpää 2007].

Pacientes e médicos estão entre as entidades envolvidas na troca de informações relativas aos cuidados de saúde [Mantzana et al. 2007]. Por isso e em termos gerais, os cuidados de saúde podem ser definidos como actos que os profissionais de saúde realizam para diagnosticar, tratar, e prevenir doenças nos seres humanos.

No sector dos cuidados de saúde, a utilização das TI nas suas actividades quotidianas simplifica e melhora o fornecimento e a qualidade dos serviços prestados. De acordo com um estudo da União Europeia sobre a dinâmica dos sistemas de saúde e as novas tecnologias informáticas [EU 2011], as aplicações das TIC nos sistemas dos cuidados de saúde são classificadas nos seguintes grupos:

- Aplicações para o direcionamento dos serviços e para a gestão das patologias, isto é, prevenção da doença, diagnóstico, reabilitação, terapia e telemedicina.
- Aplicações de suporte a actividades de gestão geral, logística e fornecimento de bens e serviços relacionados com a saúde. Exemplos: registos electrónicos do doente (*Electronic Patient Record* - EPR), registos electrónicos de saúde (*Electronic Health Record* - EHR), e arquivo de imagens e sistemas de comunicação (*Picture Archiving and Communication Systems* - PACS).
- Aplicações utilizadas o fornecimento de infraestruturas de conhecimentos de saúde, educação médica, investigação em ciências médicas e ensaios clínicos.
- Aplicações para a infraestrutura das TI, as operações dos diversos sistemas e o enquadramento dos problemas de segurança e de recuperação de dados.

Alguns prestadores de serviços de saúde já adotaram a computação em nuvem (*cloud computing*) e algumas empresas de TI oferecem várias aplicações que também podem ser utilizadas nestes serviços. Por exemplo, a Microsoft healthVaults é uma plataforma web para partilha de informações entre os diversos intervenientes dos sistemas de cuidados de saúde, tal como doentes, médicos e outros técnicos de saúde [Microsoft 2013].

Também a Oracle oferece uma solução integrada de cuidados de saúde para sustentar a gestão do desempenho, as melhores práticas e a melhor utilização dos recursos existentes e a Amazon tem ao dispor um serviço de várias soluções (Amazon web service - AWS), o qual pode ser integrado nos cuidados médicos.

Embora a área de computação médica tenha contribuído significativamente para a melhoria da atenção dada à saúde, ainda há tremendos desafios para os médicos. Por exemplo, frequentemente imagina-se que os benefícios de um sistema de registo médico sem papel são facilmente percebidos, uma vez que toda a informação disponível em papel é digitalizada. Um estudo feito por Shortliffe lança uma sombra sobre este otimismo. Em verdade, as 168 visitas sucessivas feitas a uma clínica de Medicina Interna basearam um estudo que revelou que, apesar de ter todas as histórias clínicas de seus doentes, com resultados de laboratório e de radiologia, em 81% das visitas alguns dados julgados importantes pelo médico não estavam disponíveis [Shortliffe 1976].

Por outras palavras, estes dados não tinham sido introduzidos no processo conducente à obtenção da informação que julgamos necessária hoje. Este é um problema que provavelmente já teria sido resolvido através de digitalização dos recursos de informação baseados em papel, tornando-os disponíveis para os médicos. É um problema de informação e não um problema de informática.

Situações semelhantes foram relatadas pelo *Workgroup for Electronic Data Interchange (WEDI)*, o qual se referiu ao facto de que 50% dos registos médicos baseados em papel hoje, ou estão perdidos completamente ou contêm dados incompletos [WEDI 1996]. Um relatório do *Institute of Medicine* dos EUA refere que cerca de 30 % das prescrições de tratamento médico nunca são documentadas [IOM 1999]. Estes são problemas de aquisição de informação e requerem um conhecimento íntimo do conteúdo de informação, do ambiente no qual ela é recolhida e a tecnologia que é usada para obtê-la, de maneira a obter uma solução viável.

O domínio da informática médica não é novo e tem o potencial de beneficiar o atendimento ao utente. Os seus benefícios diretos serão derivados de ser possível aumentar as capacidades e a acção dos médicos e de outros profissionais de saúde, através de um melhor acesso ao conhecimento médico e informação.

2. Sistemas de Informação nas Organizações de Saúde

A finalidade dos cuidados de saúde é melhorar o estado físico e mental e também o bem-estar social das pessoas [Koskinen 2010]. Alguns autores afirmam mesmo que "*healthcare organization is any settings where healthcare related services are provided*" [Wager et al. 2005]. Deste modo, o ambiente dos cuidados de saúde tem de ser compreendido como o conjunto de condições nas quais são realizadas actividades tais como diagnósticos, tratamentos, descoberta de medicamentos e prevenção de doenças. Neste ambiente incluem-se hospitais, clínicas, laboratórios médicos e farmácias e todos os participantes (atores) na prestação dos cuidados de saúde.

Os principais participantes do ambiente dos cuidados de saúde são os utentes (doentes e familiares que intervêm nos vários processos), pessoal médico, enfermeiros, paramédico, outros técnicos de saúde e pessoal administrativo. Ampliando esta abordagem, pode-se dizer que os participantes compreendem tanto indivíduos como organizações que afetam e são afetados pelos SI no domínio da saúde [Mantzana et al. 2007; Siau 2003]. No que se refere aos indivíduos estes podem ter atividades profissionais em diferentes organizações incluindo hospitais, clínicas, indústria farmacêutica, indústria de equipamentos hospitalares e de instrumentos de medida, farmácias, companhias de seguros e outros tipos de empresas cuja oferta sirva os diversos organismos onde se efetue a prestação de serviços de saúde.

Aumentando a frequência e os níveis do acesso à informação generalizada e ao conhecimento específico, as TIC tinham de atingir a entidade que procura nas organizações de saúde os cuidados necessários ao seu bem-estar: o doente, melhor dito, o utente. Sendo o ator principal em muitos processos de cuidados de saúde, também tem de ser repensada a sua posição nos SI o que está associado ao conceito de "paciente informado" [Kivits 2004]. Modernamente, muitos utentes procuram um melhor conhecimento do seu estado de saúde usando as disponibilidades das TIC, transformando-se em "cidadãos informados" ("well-informed citizen") capazes de contribuir para cuidados de saúde prestado de modo interativo [Henderson 2006].

3. Informação na Gestão Hospitalar

Desde o início dos anos 1980, o SI apareceu como uma ferramenta ao serviço de uma melhor gestão hospitalar. A informatização do SI hospitalar progrediu desde este período, situando-se inicialmente nos aspectos administrativos e contabilísticos e depois na própria actividade médica. Geralmente, distingue-se dois grandes tipos de Sistemas de Informação Hospitalar (SIH):

- Os sistemas de gestão administrativa: dizem respeito às actividades que suportam o processo de aceitação e entrada do doente, mas não participam directamente na produção da prestação de cuidados de saúde. As informações criadas e geridas por este primeiro tipo de sistemas respeitam por exemplo a serviços como a gestão de pessoal, operações contabilísticas e financeiras, gestão de refeições e a gestão de roupas, desde lençóis a fardamentos;
- Os sistemas relacionados com o próprio doente (todos os assuntos de natureza administrativa, médica e paramédica): estes sistemas incluem todas as informações necessários à realização dos actos do processo de produção dos cuidados de saúde prestados ao doente e à circulação dos fluxos físicos que esta produção exige.

No domínio da Saúde, os SI têm sido indicados como valiosas ferramentas para o diagnóstico de situações de saúde, tendo em vista intervenções mais aproximadas das necessidades da população. Visto que um hospital é um conjunto de recursos humanos, físicos (infraestrutura, tecnologias), recursos financeiros e normas para o desenvolvimento e a prestação de serviços de saúde, o hospital é também uma entidade que gera uma enorme quantidade de dados que são parte essencial do seu SI.

Muitas vezes, há um grande aparato no que se refere aos dados e à sua diversidade, mas é necessário enfatizar a análise da informação, a produção de novos conhecimentos possibilitando a tomada de decisões mais adaptadas às necessidades organizacionais. Além disso, é conveniente estabelecer mecanismos que garantam o registo, o seguimento e a avaliação dos dados obtidos, a fim de facilitar o seu processamento e posterior transformação em informações úteis para a dinâmica da organização [Londoño 2003]. Por isto, o SI é um meio de apoio à gestão institucional no melhoramento de eficácia, eficiência e da competitividade, contribuindo também fundamentar a tomada de decisões em todos níveis hospitalares.

Os SI têm um papel notório na certificação da informação necessária à organização que a suporta. No sector da saúde, é fundamental que os SI assegurem a informação útil e necessária às inúmeras funções dos profissionais de saúde, nos mais diversos níveis da instituição. Para que tal seja alcançado é indispensável definir o SI, como uma conjunto de estruturas sólidas, capazes de reunir, armazenar, processar e fornecer informação relevante, de modo a torná-la acessível e útil para todos, mas de acordo com uma diferenciação de acessos. Deste modo, a optimização do fluxo de informação numa organização de saúde será a solução adequada para a melhoria da qualidade de cuidados prestados.

Actualmente o hospital é uma instituição organizada nos mínimos detalhes, com a finalidade de apresentar soluções para as necessidades e problemas dos clientes/utentes. No seu contacto com

o hospital, um cliente/utente origina uma elevada quantidade de dados administrativos (admissão, dados estatísticos, contas a pagar...) e assistenciais (tratamentos médicos, de enfermagem, exames de diagnóstico, testes terapêuticos...). A crescente quantidade e a complexidade destes dados originam a necessidade da gestão dos mesmos com uma maior presença das tecnologias informáticas. Dada a importância de um SI em contexto hospitalar, importa que os seus objectivos sejam:

- Disponibilizar as informações dispersas relativas aos utilizadores em tempo real de modo integrado para garantir a eficiência do processo de tomada de decisões;
- Impedir que os funcionários do hospital executem tarefas administrativas repetitivas;
- Dotar o hospital de uma ferramenta de medição e controlo de desempenho e de gestão de custos da atividade hospitalar [Londoño 2003].

A fim de modernizar as infraestruturas informáticas da saúde, a Administração Central do Sistema de Saúde (ACSS) definiu um Plano de Transformação dos Sistemas de Informação Integrados da Saúde (PTSIIS), com o objectivo de sintetizar uma visão estratégica para os SI que decorre das principais orientações definidas para a saúde no Plano Nacional de Saúde (PNS), no Programa de Governo e em diretrizes da União Europeia.

Este modelo procura delinear o caminho a percorrer pelos SI da saúde para que estes possam assegurar uma contribuição na concretização das estratégias para obter mais e melhor saúde para todos e de forma sustentada. De acordo com a Administração Central do Sistema de Saúde, este modelo futuro "(...) procura assegurar que os Sistemas de Informação Integrados da Saúde (SIIS) suportem a preconizada abordagem centrada no utente, no ciclo de vida e na gestão integrada da doença. Pretende assegurar que os SIIS possam contribuir de forma decisiva para a materialização das estratégias e para a gestão da mudança, nomeadamente através de uma mudança centrada no cidadão/utente e na capacitação do sistema de saúde para a inovação. Procura-se garantir que os SIIS potenciem a obtenção de uma visão integrada do sistema de saúde, a melhoria do acesso aos cuidados de saúde primários, hospitalares e continuados e a melhoria da qualidade da prestação de serviços de saúde. Por último, pretende-se assegurar que os SIIS possam suportar as estratégias para garantir a monitorização da execução das orientações estratégicas definidas PNS, no Programa de Governo e nos compromissos assumidos ao nível da União Europeia" [ACSS 2007].

4. Problemas, dificuldades e insuficiências

Os SI nas Organizações de Saúde e a utilização das tecnologias informáticas que os suportam apresentam uma complexidade de problemas e de insuficiências que uma abordagem técnica tem de denunciar. Devido à natureza desta comunicação são seguidamente indicados apenas alguns:

- Os dados existentes nos serviços de saúde são de diversas naturezas e dimensões e os processos de registo e o respetivo armazenamento tem de ser geridos por métodos muito diferentes, o que exige uma difícil organização. Por exemplo, alguns registos podem ser arquivados por tema, enquanto outros exigem uma denominação própria ou números de identificação;
- As informações têm de ser classificadas de acordo com critérios oficiais ou universalmente instituídos a fim de possibilitar a sua clara compreensão;
- Em cada situação há informações consideradas como críticas para efeitos de decisão e de avaliação, as quais podem estar deficientemente armazenadas, no que se refere à segurança do sistema e à eventual recuperação em caso de perdas ocasionais ou por desastre;
- A gestão dos SI e das tecnologias associadas pode apresentar dificuldades devidas a deficiências de organização e à ausência de conexões estratégicas com as principais actividades das organizações de saúde;
- Inexistência de processos de controlo interno e de auditorias periódicas à qualidade de desempenho de todos os componentes de cada SI;
- Por vezes, existem graves inadequações entre as necessidades tecnológicas dos SI e as possibilidades orçamentais.

5. Repensar as TI nas Organizações de Saúde: Breves Propostas

Exigindo sempre uma grande abertura às mudanças de aperfeiçoamento, é importante que a evolução dos SI em saúde considere quais serão as principais linhas de desenvolvimento e quais as consequências futuras dessa evolução [Haux 2010]. A introdução das TI nos sistemas dos cuidados de saúde não pode limitar-se a uma abordagem *top-down* para que não sejam esquecidos os pontos de vista, as preferências dos diversos atores do ambiente das organizações de saúde [Coiera 2009]. Assim sendo, julga-se necessário propor o repensar de domínios como dados, atores, modernização e integração, tendo em atenção os seguintes aspectos:

DADOS

1. Os dados relativos à saúde devem ser tratados predominantemente pelas tecnologias informáticas (processamento, distribuição e armazenamento) e não limitadas ao suporte em papel [Chaudhry et. al. 2006];
2. É indispensável considerar a adequada colheita e processamento de dados de natureza imagética por serem a base de muitas práticas nos domínios da saúde [Blobel, et al. 2004];
3. No domínio da saúde, certos dados têm de ser considerados não só para a prestação de cuidados adequados, mas também no que interessa às políticas de saúde e para repensar a eficiência de práticas clínicas [Anderson 2004];

ATORES

4. Todos os atores podem ter acesso às informações presentes no sistema, mas de forma diferenciada de acordo com o seu papel na prestação dos cuidados de saúde [Kickbusch & Lister 2006]. Esta linha de evolução considera a segurança e a proteção da privacidade;
5. Os utentes têm de ser membros dos SI na medida em que são atores e, por isso, fontes de dados e também consumidores críticos das informações produzidas;

MODERNIZAÇÃO

6. Revisão periódica e constante atualização dos SI no que respeita também aos avanços das tecnologias informáticas e respectivos benefícios [Mahaney & Lederer 2003];
7. A gestão dos SI tem de estar sempre aberta à possível adoção dos avanços tecnológicos que possam participar e/ou contribuir numa maior eficácia da gestão da informação;

INTEGRAÇÃO

8. Muitos tratamentos das patologias têm como suporte tecnologias que, embora diferentes, podem ser unificadas integradamente por tecnologias informáticas, o que tem de acompanhar a modernização de organizações de saúde tanto públicas como privadas;
9. O aumento dos custos dos meios e dos processos de prestação de cuidados de saúde tem levado algumas organizações a terem de repensar as suas práticas de funcionamento, criando disponibilidade para a re-engenharia de processos integrados que possa

conduzir ao uso mais eficiente e mais rendibilizados dos recursos técnicos, humanos e financeiros;

10. É indispensável ter uma perspectiva integrada do Sistema Nacional de Saúde [Ramsay et al. 2009] e esta integração tem de dar preferência a modelos organizativos que considerem as relações entre as organizações em termos ou regionais ou nacionais [Leutz 2005].

6. Conclusões

É um facto que no desenvolvimento alcançado pela sociedade atual, os avanços científicos e técnicos desempenham um papel decisivo e o uso generalizado das TIC dão origem ao que é conhecido como Sociedade da Informação e Comunicação (SIC). A sociedade da informação em geral e, em especial, as novas TIC têm um impacto significativo em todos os níveis do mundo socioeconómico, em particular no que se refere ao dos cuidados de saúde.

A modernização dos organismos que integram o sistema de saúde tem de incluir a utilização intensiva de TI enquanto investimento estratégico, a fim de possibilitar a racionalização dos recursos informacionais disponíveis e de incrementar a eficiência dos serviços e a qualidade da prestação de cuidados de saúde. Tendo as tecnologias informáticas como suporte, os SI têm de oferecer respostas adequadas às necessidades de informação dos diferentes serviços e também possibilitar a partilha de informação entre estes, de modo a garantir a integração.

Embora o domínio da saúde seja uma das áreas onde existe uma intensa inovação, a utilização estruturante e eficiente dos SI manifesta ainda a necessidade de maiores atenções. A gestão integrada de SI em ambiente hospitalar já foi e continua a ser uma área de risco e de desafio, devido por um lado à necessidade de criar infraestruturas de suporte às aplicações a implementar e, por outro, pela definição dos modos e dos suportes tecnológicos com que estas têm de ser adaptadas aos processos dos diversos serviços organizacionais.

7. Referências

ACSS – Administração Central do Sistema de Saúde, *Definição do Plano de Transformação dos Sistemas de Informação Integrados da Saúde (PTSIIS)*. (2007), [Em linha]. Lisboa, 2007. [Consult.05 maio 2014]. Disponível em *http://www.acss.min-saude.pt/Portals/0/Images/ConteudoGenerico/PTSIISsumarioexecutivo1.pdf*

Anderson, W. L. - Some challenges and issues in managing, and preserving access to, long-lived collections of digital scientific and technical data, *Data Science Journal, Vol3, p*. 191-201, 2004

Blobel, B.; Gell, G.; Hildebrand, C. Engelbrecht, R. (eds.), Contribution of Medical Informatics to Health – Integrated Clinical data and Knowledge to Support Primary, Secondary and Tertiary Home Care, *Proceedings of the European Federation for Medical Informatics*, Special Topics Conference, June 13-16, 2004, Munich, Germany, IOS Press

Chaudhry, B.; Wang, J.; Wu, S.; Maglione, M.; Mojica, W.; Roth, E.; Morton, S.; Shekelle, P. - Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care. *Annals of Internal Medicine,* 144(10), p. 742-752, 2006

Coiera, E., "Building a National Health IT System from the middle out", *Journal of American Medical Information Association*, 16, 3, (2009), 271–273

EU - European Union. *Dynamic health systems and new technologies: e-Health solutions at local and regional levels*. 2011, ISBN: 978-92-895-0536-9. Retrieved April 15, 2013 from http://cor.europa.eu/en/documentation/studies/Documents/c24aa096-55a7-4e43-be10-d938dfad6251.pdf

Gans, B. M., "Strategic impact of the changing healthcare landscape", In *Proceedings of the 19th International Conference of the IEEE, 1997*, Chicago, IL, USA. Vol. 6, (1997), 2670 – 2671

Haux, R., "Medical informatics: Past, present, future", *International Journal of Medical Informatics*, 75, 3-4, (2010), 268–281. (http://science.icmcc.org/2010/07/08/medical-informatics-past-present-future/)

Henderson, A., "Boundaries Around the 'Well-Informed' Patient: The Contribution of Schutz to Inform Nurses' Interactions", *Journal of Clinical Nursing* 15, 1, (2006), 4-14

IOM - Institute of Medicine - *To Err Is Human: Building a Safer Health System (*Report). Washington, DC: National Academy Press, 1999

Kickbusch, I. and Lister, G. (Eds.) (2006), *European Perspectives on Global Health – A Policy Glossary*, EFC: Brussels.

Kivits, Joëlle, "Researching the "informed patient": The Case of Online Health Information Seekers", *Information, Communication & Society*, U.K., Routledge, 2004.

Koskinen, J., "Phenomenological view of health and patient empowerment with personal health record", In *Proceedings of the Well-being in the Information Society (WIS) conference*. University of Turku, Finland, (2010), p. 1-15

Leutz, W. "Reflections on integrating medical and social care: Five laws revisited", *Journal of Integrated Care*, **13** (5), 3-12, 2005

Londoño, G. M., "Generalidades sobre Administração Hospitalar*",* In: Londoño, G.; Morera, R. & Laverde, G. *Administração hospitalar*. 2ª ed. Rio de Janeiro: Guanabara Koogan, 2003

Mahaney, R.; Lederer, A. Information systems project management: an agency theory interpretation. Journal of Systems and Software, 68(1), p. 1-9, 2003

Mantzana, V.; Themistocleous, M.; Irani, Z.; Morabito, V., "Identifying healthcare actors involved in the adoption of information systems" *European Journal of Information System*, 16, (2007) 91-102. Available at: http://www.palgrave-journals.com/ejis/journal/v16/n1/full/3000660a.html. Last accessed 25th mar 2013

Microsoft - **Cloud services for health**. [Em Linha]. 2013. [Consult. 03 fev. 2014]. Disponível em: http://www.microsoft.com/health/en-us/initiatives/pages/cloud-services-for-health.aspx

Ramsay, A.; Fulop, N.; Edwards, N. The Evidence Base for Vertical Integration in Health Care. *Journal of Integrated Care* **, 17** (2), 3-12, 2009

Shortliffe, E. H. *Computer-Based Medical Consultations: MYCIN*. New York: Elsevier/North Holland, 1976

Siau, K., "Healthcare informatics", *Journal of IEEE Transactions on information technology in biomedicine*, 7 (1), (2003), 1–7

Srivastava, P.; Yadav, R.; Razdan, P., "Cloud Computing in Indian Healthcare Sector", In *Proceedings of ASCNT 2011, CDAC, Noida, India*, (2011), 1-8

Tähkäpää, J., "Managing the Information Systems Resource in Healthcare", *Publications of Turku School of Economics*, 2007, Turku, Finland. (PhD Thesis)

Wager, K.; Lee, F.; Glaser, J., *Managing healthcare information systems: A practical approach for healthcare executives*. San Francisco: Jossey-Bass, (2005).

WEDI - Workgroup for Electronic Data Interchange, *Report to Secretary of U.S. Department of Health and Human Services*, July 1996. Available from: URL: http://www.wedi.org/. Accessed April 16, 2013

ARTIGOS

DOCTORAL CONSORTIUM

ARQUITETURAS INFORMACIONAIS PARA A ORGANIZAÇÃO DA INFORMAÇÃO EM REDES COLABORATIVAS EMPRESARIAS

Solange Francisca Mazzaroto (estudante de doutoramento) [1], António Manuel Lucas Soares (orientador) [2]

1) Faculdade de Letras, Universidade do Porto, e INESC TEC, Portugal

sfm@inesctec.pt

2) Faculdade de Engenharia, Universidade do Porto, e INESC TEC, Portugal

als@fe.up.pt

Resumo

Esta investigação está inserida no âmbito das áreas de gestão da informação e de gestão do conhecimento e procura abordar o tema das redes colaborativas, com foco nas organizações empresariais.

Essas redes são estabelecidas a partir de relações entre as organizações. São inúmeros os benefícios alcançados com a prática da colaboração, no entanto, verifica-se que as plataformas digitais, utilizadas para a partilha da informação, são limitadas no que diz respeito à organização dessa informação, o que dificulta a sua recuperação.

Além disso, nas redes colaborativas, podem existir várias plataformas, por isso, verifica-se a necessidade da integração entre as informações disponibilizadas, ou seja, é necessário que haja uma interoperabilidade entre essas plataformas.

De uma forma geral, as necessidades de gestão da informação, nos contextos colaborativos, são heterogêneas e dinâmicas, o que faz com que as arquiteturas informacionais utilizadas desatualizem-se rapidamente. Assim, o problema identificado é que a informação partilhada nas redes colaborativas empresariais não é organizada de forma que possa ser recuperada a longo prazo.

Em virtude deste cenário, este estudo propõe desenvolver uma arquitetura informacional, com propriedades de flexibilidade e adaptabilidade aos objetivos das redes colaborativas empresariais, de modo que a informação partilhada seja melhor organizada, o que contribuirá com a sua posterior recuperação.

Palavras chave: Redes Colaborativas Empresariais. Arquiteturas Informacionais. Organização da Informação.

1. Situação da Investigação

A investigadora está matriculada, desde setembro de 2014, no Programa de Doutoramento em Informação e Comunicação em Plataformas Digitais (ICPD), ministrado em convênio entre o Departamento de Comunicação e Arte (DECA), da Universidade de Aveiro (UA), e a Faculdade de Letras (FLUP), da Universidade do Porto (UP).

Além da realização do curso de doutoramento, a investigadora também está enquadrada ao INESC TEC, no Centro de Engenharia e Sistemas Empresariais (CESE), na área de Redes Colaborativas, no Porto, local onde será desenvolvida a investigação.

O projeto de tese está inserido na área de Ciência da Informação, que apresenta natureza multidisciplinar. Assim, os temas abordados serão: gestão da informação, gestão do conhecimento, organização da informação e redes colaborativas empresariais. A previsão para a submissão da tese é julho de 2017.

Para a investigadora, participar do *Doctoral Consortium* (DC) representa uma excelente oportunidade para se beneficiar da experiência e do conhecimento da comunidade científica. A partir das recomendações recebidas, será possível ajustar o projeto de tese, já que o seu desenvolvimento ainda está na fase inicial.

2. Contexto e Motivação

O projeto de tese tem com foco principal a organização da informação em redes colaborativas empresariais. Para partilharem informações entre si, essas redes utilizam plataformas digitais, que nada mais são do que sistemas de informação, onde, para além da partilha da informação, também são utilizados para a sua organização, o seu armazenamento e a sua recuperação.

Assim, a maior motivação da investigadora em desenvolver uma tese sobre este assunto é devido à sua formação básica em Ciência da Informação, área científica que tem como objeto de estudo a informação. Entre os temas de competência desta área, está a elaboração e o desenvolvimento de técnicas e de ferramentas para representarem e organizarem a informação, de modo que a sua recuperação possa ser realizada de forma imediata e a longo prazo [Lima 2003].

No decorrer da sua vida acadêmica, a investigadora tem se preocupado com o tema das redes colaborativas empresariais. Por este motivo, verificou que é necessário que seja criada uma solução que possa contribuir com a melhoria no desempenho dessas redes, a partir do desenvolvimento de formas de organização da informação, que irá auxiliar na recuperação das mesmas, e, assim, tornem mais eficazes os processos de trabalho colaborativo.

3. Fundamentos e Trabalho Relacionado

O tema central da investigação são as redes colaborativas empresariais, que possuem várias outras denomimações e definições. Para Popp et al. [2013] essas redes são definidas como três ou mais organizações que estabelecem relações, em direção a um objetivo comum.

Uma rede colaborativa é constituída por uma variedade de entidades (organizações e pessoas, por exemplo) que são, em grande parte, autônomas; estão geograficamente distribuídas; e são heterogêneas em termos de: ambiente operacional, cultura, capital social e objetivos. Essas entidades colaboram para melhorar a realização comum de objetivos compatíveis, e cujas interações são suportados por uma rede de computadores [Camarinha-Matos et al. 2005].

Essa definição é corroborada por Carneiro et al. [2007], que definem uma rede colaborativa como uma rede de relações, entre indivíduos ou grupos de indivíduos, que se baseia, fundamentalmente, em processos colaborativos mediados por tecnologias de informação. Uma rede de organizações é considerada uma rede colaborativa se as relações de colaboração são uma parte importante das relações inter-organizacionais, existentes entre essas organizações.

As redes inter-organizacionais podem ser entendidas como arranjos institucionais que permitem uma organização eficiente das atividades econômicas, a partir da coordenação de ligações sistemáticas, estabelecidas entre empresas interdependentes. Uma das ligações que caracterizam quaisquer tipos de redes é a partilha de informações [Lages et al. 2013].

Da mesma forma que as redes colaborativas, não existe um consenso para o termo partilha de informações. Para Lages et al. [2013] partilha de informações é definido como o processo de comunicação da informação entre os membros da rede. Este processo é compreendido para ocorrer num contexto social, o que significa que as necessidades informacionais dos indivíduos não são meramente cognitivas, mas diretamente relacionadas aos gupos sociais e de trabalho, o qual eles pertecem.

As plataformas utilizadas para partilha de informações, no entanto, possuem arquiteturas informacionais que demonstram-se ineficientes para a organização dessas informações, o que dificulta a sua posterior recuperação.

O termo arquitetura refere-se ao projeto de ambientes habitáveis. Já a arquitetura da informação está preocupada com a criação de ambientes para entendimento. Assim, como arquitetura possibilita ambientes para habitação, pelas formas de organização e espaços, a arquitetura da informação permite ambientes para a compreensão, a partir da organização de nós e *links*[1] [Arango 2011].

No ambiente empresarial, a arquitetura de informação é utilizada como um termo coletivo para descrever os vários componentes da infra-estrutura global de informação, que usam o modelo de

[1] Os nós são elementos de conteúdo (textos, imagens, vídeos, etc.), que comunicam, em conjunto, um conceito ou ideia. Já um link é a relação entre dois ou mais nós. Um exemplo é o *hiperlink*, em que um nó (uma palavra ou grupo de palavras) refere-se a um segundo nó (uma página *web*, ou uma de suas partes) [Arango 2011].

negócio, e os componentes do seu processo, e entregam sistemas de informação que os suportam e os disponibilizam [Resmini et al. 2011].

Para desenvolver arquiteturas informacionais flexíveis e adaptáveis aos interesses das redes colaborativas empresariais, a partir de uma análise sobre as técnicas e ferramentas disponíveis atualmente, a investigação pretende apronfundar-se nos estudos sobre as ontologias e os dados ligados (*linked data*), que apresentam-se com alternativas que podem contribuir com a solução do problema de investigação.

No âmbito corporativo, a ontologia é utilizadas para descrever as entidades comuns dos seus sistemas e as relações entre elas. Como resultado, é possível tratar todos os conhecimentos e as competências disponíveis como uma distribuída base de conhecimento [Smirnov et al. 2008].

A ontologia é uma tecnologia emergente para a implementação de sistemas orientados para o conhecimento. Fornece uma semântica rica para criar estruturas conceituais e descrições lógicas em um domínio de interesse [Chi 2010].

Em geral, as ontologias fornecerm uma compreensão comum e partilhada sobre uma área de conhecimento, que pode ser comunicada entre as pessoas, a partir de heterogêneas e distribuídas

fontes de conhecimento e sistemas de aplicação. O seu papel prático, no modelo de partilha de conhecimento é:

- explicar o conhecimento tácito necessário para as tarefas da vida real; e
- preencher uma lacuna entre as representações heterogêneas de conhecimento explícito [Smirnov et al. 2008]

Com relação aos dados ligados, verifica-se na literatura que o termo *linked data* surgiu em 2006, como uma alternativa para tentar resolver o problema da quantidade de dados disponíveis na *web*, ou seja, de textos ou arquivos, dos mais variados formatos, dos quais muitos só podem ser interpretados por seres humanos. Isso impedia que aplicações (máquinas) conseguissem extrair as informações reais, contidas nesses documentos [Cabral et al. 2012].

O *linked data* possibilita a ligação entre diferentes fontes de informação, formando a *web* de dados. Isso torna possível que as aplicações genéricas operem sobre um conjunto de dados mais completo [Cabral et al. 2012].

De uma forma geral, o *linked data* é definido como um conjunto de melhores práticas para a publicação e a conexão de dados estruturados na *web* [Heath et al. 2009].

4. Definição da hipótese ou do problema de investigação

Conforme abordado nos itens anteriores, as empresas, de uma forma geral, estão implementando processos colaborativos, baseados na utilização de plataformas digitais. Essas plataformas, no entanto, são limitadas em termos de organização e recuperação da informação. Este fato pode ser explicado a partir da arquitetura informacional utilizada, já que foi desenhada para resolver uma solução imediata, ou seja, a partilha momentânea de informações.

Além dessa problemática, também verifica-se que as necessidades de gestão da informação, em contextos colaborativos são heterogêneas e dinâmicas, ou seja, os modelos e arquiteturas usados desatualizam-se rapidamente. Assim, a partir contexto exposto, o problema a ser solucionado por esta investigação é que a informação partilhada nas redes colaborativas empresariais não é organizada de forma que possa ser recuperada a longo prazo.

A questão de investigação formulada para o desenvolvimento deste estudo é: como desenvolver arquiteturas informacionais, para a organização e a recuperação da informação, em uma rede colaborativa empresarial, que tenham propriedades de flexibilidade e adaptabilidade às necessidades das organizações?

A hipótese sugerida para responder a essa questão é que a partir da utilização de ontologias e dados ligados (*linked data*) é possível desenvolver uma arquitetura informacional, com propriedades de flexibilidade e adaptabilidade aos objetivos das redes colaborativas empresariais, que possam facilitar a recuperação da informação a longo prazo.

5. Objetivo, Metodologia e Métodos de Investigação

O objetivo desta investigação é propor medidas para melhorar o desempenho das redes colaborativas empresariais, a partir do desenvolvimento de formas de organização e de recuperação da informação, que tornem mais eficazes os processos de trabalho colaborativo.

Como enquadramento metodológico, este estudo será pautado pelo paradigma do *Design Science Research*. O objetivo deste paradigma é abordar a investigação com a finalidade de desenvolver artefatos que, uma vez testados, serão capazes de fornecer soluções inovadoras para necessidades específicas do negócio.Também pretende desenvolver novos conhecimentos de modo a conceber soluções para problemas de campo específicos [van Aken 2004].

Na figura 1, a seguir, foi realizado o enquadramento desta investigação, de acordo com os pressupostos do *Design Science Research*.

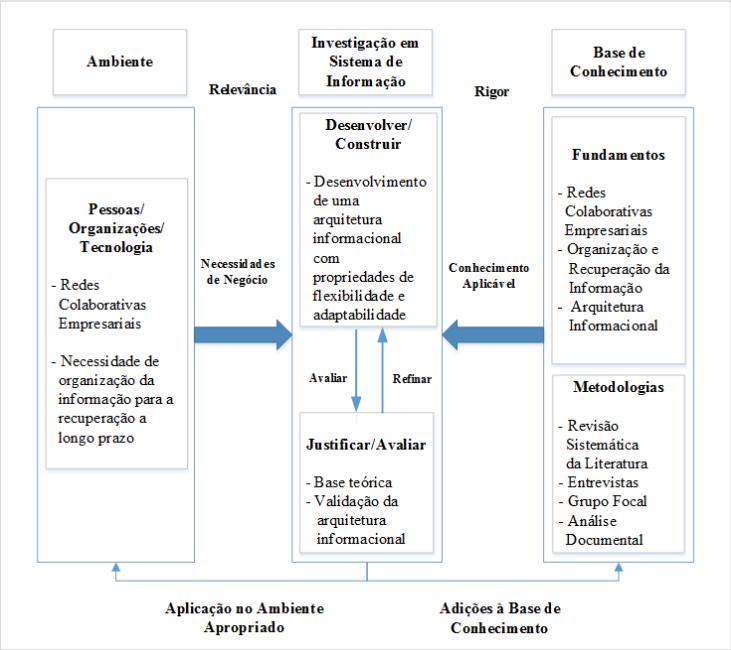


Figura 1: Information Systems Research Framework
Fonte: Adaptado de Hevner et al. [2004]

Esta investigação será de natureza qualitativa e o método de investigação utilizado será o estudo de caso, que tem como objetivo investigar determinada situação que poderá representar outras situações semelhantes [Severino 2007 & Gil 2010]. Quanto a fase empírica da investigação, na figura 2 está descrita a sua operacionalização.

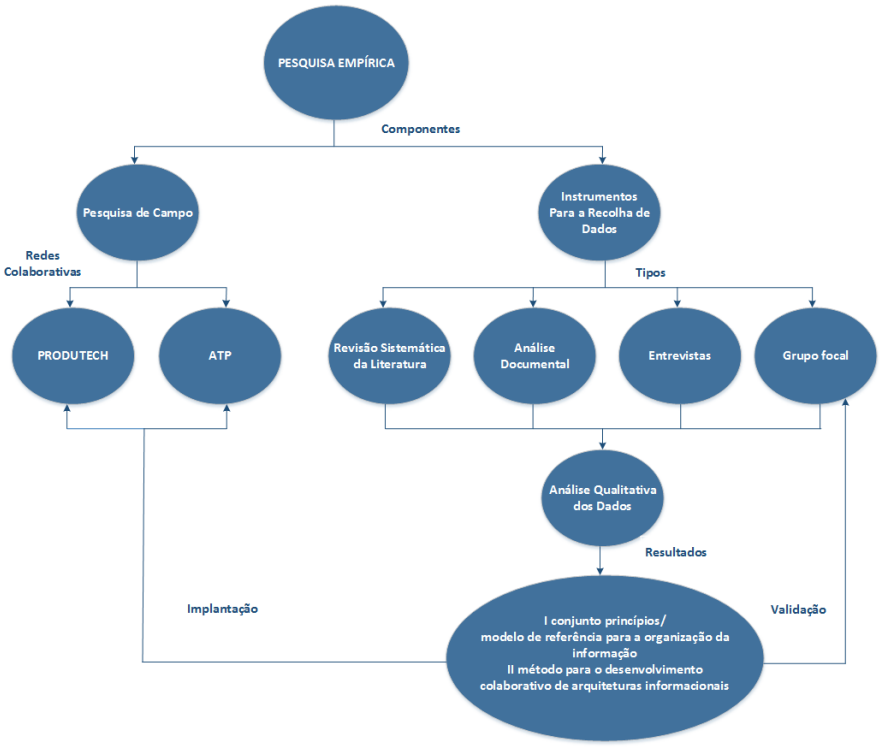


Figura 2 - Operacionalização da Investigação
Fonte: Autor

6. Estado da Tese

Para o desenvolvimento da tese, foi elaborado o cronograma operacional, descrito no quadro 1, seguir.

ATIVIDADE/TEMPO (trimestre)	2015				2016				2017			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
1ª FASE - DESENVOLVIMENTO DO PROJETO												
Elaboração do Projeto de tese												
Qualificação do Projeto de tese												
2ª FASE - DESENVOLVIMENTO DA TESE												
Levantamento Bibliográfico/Revisão Sistemática												
Definição dos Instrumentos para a Coleta de Dados												
Coleta de Dados												
Análise de Dados												
Elaboração do Desenho da Arquitetura Informacional												
Validação da Arquitetura Informacional												
Escrita da Tese												
Entrega e Defesa da Tese												

Quadro 1 – Cronograma Operacional
Fonte: Autora

A partir da participação no DC, a investigadora pretende obter o apoio científico dos seus pares, por meio de sugestãoes e recomendações que possam auxiliar no desenvolvimento do estudo.

7. Conclusão e Contributos Esperados

O projeto de tese, a ser desenvolvido, tem como objetivo contribuir com a melhoria no desempenho das redes colaborativas empresariais, a partir do desenvolvimento de formas de organização e de recuperação da informação, que tornem mais eficazes os processos de trabalho colaborativo.

Dessa forma, espera-se desenvolver o desenho de uma arquitetura informacional, com propriedades de flexibilidade e adaptabilidade aos objetivos das redes colaborativas empresariais, para que a informação partilhada seja organizada de modo estruturado e, com isso, a sua recuperação possa ser realizada a longo prazo.

Para alcançar esse resultado, serão necessários:

a) Um conjunto de ontologias para a organização genérica da informação nas redes colaborativas empresariais; e

b) Um método para o desenvolvimento colaborativo de uma arquitetura informacional, baseada em dados ligados (*linked data*), utilizando o conjunto de ontologias.

O desenvolvimento do desenho de uma arquitetura informacional, além de melhorar o desempenho das redes colaborativas empresariais, também irá contribuir muito com a

comunidade de investigadores em Sistemas de Informação, já que irá desenvolver uma solução informacional para um problema real e atual.

8. Referências

Arango, J. Architectures. *Journal of Information Architecture*, v. 3, n. 1, 2011.

Cabral, S. P. Beduschi, N. B. Zancanaro, A. Todesco, J. L. Aplicando linked data na publicação de dados do ENEM. ONTOBRAS-MOST, v. 938. *Proceedings*... Recife: 2012, p. 176-181. http://ceur-ws.org/Vol-938/ontobrasmost2012_paper16.pdf.

Camarinha-Matos, L. M. Afsarmanesh, H. Collaborative networks: A new scientific discipline. *Journal of Intelligent Manufacturing*, v. 16, p. 439–452, 2005.

Carneiro, L. Soares, A. L. Patrício, R. Alves, A. *Redes colaborativas de elevado desempenho: no norte de Portugal*. INESC PORTO, 2007.

Chi, Y. Rule-based ontological knowledge base for monitoring partners across supply networks. *Expert Systems with Applications*, v. 37, p. 1400-1407, 2010.

Gil, A. C. *Métodos e técnicas de pesquisa social*. 6. ed. Atlas, 2010.

Heath, T. Hepp, M. Bizer, C. (eds.). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, Special Issue on Linked Data, 2009.

Hevner, A. March, S. Park, J. Ram, S. Design Science in Information Systems Research. *MIS Quarterly*, 2004.

Lage, E. Alturas, B. Information Sharing among Innovative SME: An Exploratory Study within the Portuguese SME Innovation Network. In: *Smart Innovation, Systems and Technologies*, v. 18, ed. Howlett, Robert J.; Gabrys, Bogdan; Musial-Gabrys; Katarzyna & Roach, Jim, p.159-168, By Springer-Verlag, ISBN: 978-3-642-34218-9, DOI: 10.1007/978-3-642-34219-6_18.

Lima, G. A. B. Interfaces entre a ciência da informação e a ciência cognitiva. *Ci. Inf.*, Brasília, v. 32, n. 1, p. 77-87, jan./abr. 2003.

Popp, J. MacKean, G. Casebeer, A. Milward, H. B. Lindstrom, R. *Inter-organizational networks: a critical review of the literature to inform practice*, 2013.

Resmini, Andrea, Rosati. A brief history of information architecture. Journal of Information Architecture, v. 3, n. 2, 2011.

Severino, A. J. *Metodologia do trabalho científico*. 23. ed. Cortez, 2007.

Smirnov, A. Shilov, N. Kashevnik, A. Developing a knowledge management platform for automotive build-to-order production network. *Human Systems Management*, v. 27, p. 15-30, 2008.

Van Aken, J. Management Research Based on the Paradigm of the Design Sciences: the Quest for Field-Tested and …. *Journal of Management Studies*, 2004.

Information, knowledge and collaborative networks in internationalisation processes of SMEs

Eric Costa [1], António Lucas Soares [1], Jorge Pinho de Sousa [1]

1) INESC TEC and Faculty of Engineering, University of Porto, Portugal

eric.m.costa@inesctec.pt, als@fe.up.pt, jsousa@inescporto.pt

Abstract

Small and medium enterprises (SMEs) still face challenges and obstacles in internationalisation processes, mainly due to the poor management of information, which constrain the creation and accumulation of knowledge for making decisions. This doctoral project aims to study the internationalisation processes of SMEs in collaborative networks contexts and to design state-of-the-art information and knowledge management models to significantly improve collaborative decision-making in those processes. The research will study two Portuguese industrial enterprise associations (IEAs) and their respective associates (SMEs), as a collaborative network context, following two complementary approaches: one exploratory-oriented (survey and case studies) and the other design-oriented (reference models development). The case studies will uncover in detail the relationships between information, collaboration and knowledge creation in internationalisation, focusing two types of industries: IT/electronics and textile industries. From the results of the case-studies information management models and decision support systems will be developed.

Keywords: information management; knowledge management; collaborative networks; internationalisation; small and medium enterprises

1. Research Situation

The student is enrolled since February 2015 in the Doctoral Program in Industrial Engineering and Management (PRODEIG) at the Faculty of Engineering of the University of Porto (FEUP). The estimated time for submitting the thesis is March 2018. At the moment, to conclude the 1st year at FEUP, the student only needs to submit his thesis plan and make a presentation of the research proposal (estimation in October 2015). Since September 2014 the student is also working as a researcher at INESC TEC, whereby all the work performed for the research project that he is involved will be used for his doctoral thesis. The thesis status can be found at Section 6 of this document.

The project presents a multi-disciplinarily nature, focusing on different research areas: business management, information and knowledge management, and collaborative networks. Besides the CAPSI Doctoral Consortium, the student also has presented his work in the Doctoral Congress in Engineering (DCE 2015) at FEUP in June 2015. DCE allowed the student to present and discuss his ideas and research proposal to a community of researchers of engineering, with specific focus on industrial engineering and management. Therefore, the CAPSI Doctoral Consortium can be very useful for the work to be developed by the student, mainly in what concerns the areas of information and knowledge management. As can be analysed in the next sections of this document, the project will include an important part of Information Systems (IS), with the development of information management models for improving collaborative decisions internationalisation of SMEs. So, the student will have the opportunity to benefit from the experience and knowledge of the scientific community and to adjust his work according to

the received recommendations. This thesis will be conducted using a well-known research paradigm in the IS community, the Design Science Research (DSR). Though DSR will be applied in a business management context, i.e. internationalisation process of SMEs, which can demonstrate the importance and applicability of this paradigm to different research areas and which can be a good contribution to the remaining participants of CAPSI.

2. Context and Motivation

There is a growing evidence that internationalisation of SMEs has become a priority to gain competitive advantage [Dutot et al. 2014]. However, SMEs still face major challenges and obstacles during their internationalisation processes. One of the main obstacles is the poor management of the information relevant to the internationalisation processes, which constrain the creation and accumulation of knowledge about those processes [Hsu et al. 2013].

Information and knowledge are seen as key resources for facilitating SMEs international expansion, reducing risk and uncertainty and increasing creativity in decision-making [Child & Hsieh 2014]. From another perspective, it has been found that collaborative networks can represent an important facilitator in the internationalisation of SMEs, particularly by nurturing knowledge sharing [Musteen et al. 2010]. In despite of this, it is surprising the lack of exploratory studies on how state-of-the-art information management models and technologies can be used to significantly improve collaborative decision making in the internationalisation of SMEs. Also, there is no evidence that networking and collaboration are systematically explored in the benefit of more effective internationalisation outcomes.

Therefore, this doctoral project aims to study the internationalisation processes of SMEs in collaborative networks contexts and to design state-of-the-art information and knowledge management models to significantly improve collaborative decision-making in those processes. The research will be focused in the Portuguese industrial enterprise associations (IEAs) addressed as collaborative contexts for internationalisation.

3. Fundamentals and Related Work

It is important to make a clear distinction between information and knowledge. Although they are both critical in the SMEs' internationalisation process, they influence it in distinct ways [Knight & Liesch 2002]. Information is inextricably antecedent to knowledge [Dretske 1981]. In one way, information is defined as structured and understandable data in order to be a useful input to knowledge [Child & Hsieh 2014]. In the other way, knowledge is created from information and is about beliefs and commitment, validated in a person's perception or expectation for taking actions [Nonaka & Takeuchi 1995].

Managing effectively the information in internationalisation decision processes is a critical success factor that has been somehow addressed in the literature, showing that information is crucial for more rational decision-making [Child & Hsieh 2014], to manage international complexity and ambiguity [Hsu et al. 2013], and to reduce risks and uncertainty [Nguyen et al. 2006]. Studies on the importance of the role of information in the creation of internationalisation knowledge has also been made [Casillas et al. 2010]. From another perspective, it has been found that networks can represent an important facilitator in the internationalisation of SMEs, in different ways such as: (i) providing important channels of information and knowledge to decision-makers [Musteen et al. 2010]; (ii) influencing the approach adopted by SMEs' leaders [Child & Hsieh 2014]; and (iii) influencing decisions on foreign market selection and entry mode [Ibeh & Kasem 2011].

This doctoral project will study, explore and develop solutions addressing the following research opportunities, based on future directions for research proposed in previous studies of internationalisation of SMEs:

(o1) there is a lack of comprehensive and systematic studies investigating how SMEs can acquire and use information generated in a collaborative network context to achieve higher levels of internationalisation [Child & Hsieh 2014];

(o2) only limited studies show results on how SMEs can convert information into knowledge for making decisions towards internationalisation [Fletcher & Harris 2012];

(o3) there has been only limited analysis on the development of methods and techniques to improve the decision-making process of internationalisation of SMEs [Schweizer 2012];

(o4) there is a need to explore how different contexts, where SMEs are embedded, would result in different information and knowledge needs to internationalise [Dutot et al. 2014];

This project contributes to the scientific knowledge on business management by addressing how information, knowledge and collaboration influences the internationalisation process of SMEs. Second, it contributes to the field of information and knowledge management with the development of approaches and techniques to support information and knowledge management for internationalisation decisions in collaborative networks.

The context for collaboration to be studied are Portuguese IEAs and their role in internationalisation processes of SMEs, which, to the best of the authors' knowledge, has not been addressed in detail in previous research. For example, [Lage & Alturas 2013] studied important aspects about information sharing among SMEs in the COTEC Portugal Network, a business association for innovation. However, the work was more focused on understanding how and why information sharing occurs within the innovation network, without addressing internationalisation processes. This project will demonstrate that the IEAs can act in a more active way in processes of internationalisation of their associates, by using a platform to provide advanced brokering, facilitation and coordination services, not possible so far.

The project will extend the work developed by [Knight & Liesch 2002] and later by [Nguyen et al. 2006], with the development of a model for information internalisation in the context of firms acting in collaboration, and not firms in isolation. Additionally, since there has been only limited analysis on the development of ways of improving the effectiveness of making decisions in internationalisation of SMEs, a collaborative decision support system will be designed and adapted for internationalisation purposes, based on some novel multi-criteria decision-making models and on business intelligence.

4. Hypothesis and Research Problem Definition

Many SMEs still face challenges and obstacles during their internationalisation processes, such as lack of capital, lack of adequate information and knowledge about the target markets and process of internationalisation, and lack of adequate state support [Hsu et al. 2013].

The problem definition for this doctoral thesis was defined through a scoping study (exploratory literature review) to make clear the existing basis for the work to be developed and to specify the problem to be addressed. Consequently the main research problem defined was:

(p) There is a of lack comprehensive and systematic studies which investigate how SMEs can access, organize and use the information generated in a collaborative network context and how they can collaboratively convert this information into knowledge to support decision-making in internationalisation processes.

The general hypothesis of this doctoral project is that information models and technologies, providing information, collected from heterogeneous sources, mediated by favourable collaborative contexts, positively influence the decision-making capability in the internationalisation processes of SMEs.

5. Objectives, Methodology and Research Methods

5.1. Research Framework

The research of this thesis will be conducted and its outcomes will be analysed and validated following a design science research (DSR) approach. The design activity has been studied in different disciplines, with particular relevance on Information systems (IS). [Hevner et al. 2004] considered that IS research has produced knowledge considering a complementary research cycle between two distinct paradigms, behavioural science and design science.

The origins of the DSR are based on engineering and the sciences of the artificial and is considered essentially a problem-solving paradigm [Hevner & Chatterjee 2010]. The main contribution of DSR is a new way of thinking about what makes IS research relevant to different kinds of audiences, such as managers, practitioners, and researchers [Hevner & Chatterjee 2010]. Therefore, following the DSR, an approach based on a complementary 3-cycle model will be used for this doctoral project: (1) relevance cycle – to bridge the contextual environment of the research project with the design science activities; (2) rigor cycle – to connect the design science activities with the knowledge base (from scientific foundations, experience, and expertise) comprising the research project; and (3) design cycle – to iterate between the core activities of developing and evaluating the design artefacts and processes of the research project.

5.2. Objectives and Research Questions

Two main objectives are to be achieved: from one side, the authors want to know more about the internationalisation decision processes involving SMEs in a collaborative network context, in particular how information and knowledge support those processes. From the other side, the authors aim to design informational models that foster collaborative decision making and learning in internationalisation decisions for cooperating SMEs. The research questions that will be answered by this doctoral project are:

RQ1: What is the role of collaboration in the effectiveness of the internationalisation process of SMEs? How does it mediate between information and internationalisation knowledge? How does it apply in IEA's as locus of collaborative networks?

RQ2: How can state-of-the-art information and collaboration technologies be used to substantially improve SMEs internationalisation processes in a collaborative network context?

The specific objectives and goals of the project are thus:

(g1) to identify and characterise internationalisation decisions and decision modes in the scope of SMEs' internationalisation strategies developed in collaborative network contexts;

(g2) to design information management models for searching, collecting, integrating, organising and visualizing internationalisation information, from a variety of internet accessible, heterogeneous public and private sources, to be used by SMEs in internationalisation decisions and strategy formulation;

(g3) to design a decision support system for the collaborative development of internationalisation strategies, incorporating multi-criteria decision models about market selection, choice of partners and resources, assessment of risks and negotiation.

5.3. Methodology

This project will consist of tasks consisting of different methodologies and research methods:

1) **Research design and preparation**

This research has two complementary approaches: an exploratory-oriented approach, consisting of a survey and a multiple-case study, and a design-oriented approach, to develop reference models based on results of the exploratory study.

The exploratory-oriented approach of this project is a mixed-method approach: the multiple-case study will follow the survey carried out in Task 2. Most of the questions of the survey will be refined and complementary data will be collected to develop the case studies. The design type will be a multiple-case design composed by multiple units of analysis.

This project has some complexity due to being multi-method, needing to be carefully designed and prepared. The theoretical model, research questions and hypotheses will be refined in this task. Then, a plan articulating the survey, two case studies and the models to be developed will be established.

2) Survey on SMEs' information and collaboration needs for internationalisation

The goal of the survey is to characterise SMEs' information needs for internationalisation and to understand their expectations about the role that collaborative networks can have on the process. The results of this survey will provide the contextual knowledge for the design of the multiple-case study research.

The survey will address the following topics:

- Past experiences of internationalisation processes of SMEs;
- Decision modes and internationalisation strategies applied by SMEs;
- Key persons behind the decision-making process;
- Decisions made by SMEs before and during processes of internationalisation;
- Level of relationship between SMEs acting in collaboration;
- Type of information shared between partners;
- Reasons why SMEs internationalise.

The results of this task will provide inputs to Task 3 and Task 4. The design of the multiple-case study will be influenced by the conclusions drawn in this task.

3) Multiple-case study of internationalisation processes in IEAs

Using two IEAs as case studies, this task aims to describe the existing relationships between information, collaboration and knowledge creation in the internationalisation processes of SMEs and to specify the requirements for an effective transformation of information into internationalisation knowledge, mediated through collaborative processes and tools.

The specific goals are:

- To describe and explain the factors influencing the management of information relevant for internationalisation purposes;
- To characterise comprehensively the decision processes involved in the internationalisation of SMEs and how the IEA influences them;
- To describe and explain the factors influencing the implementation of collaborative processes for internationalisation;
- To synthesise scenarios and specify the requirements for SMEs internationalisation processes in the context of IEAs.

This research will study in detail two Portuguese IEAs using a multiple-case study approach: to represent two different types of industries, i.e. IT and electronics industry and textile industry, respectively. The research will be focused on analysing information and knowledge needs in different contexts where SMEs are embedded: industry type context, institutional context, decision mode context, and internationalisation strategy context. Other focus will be the process of converting information into knowledge, i.e. the information internalisation capacity of SMEs.

Data will be collected in a way that will allow theory building and cross-case analysis. To interpret the results, the analysis will be contrasted with existing theories in the literature.

4) Information, knowledge and collaboration models for decision support in internationalisation processes

This task will develop information, collaboration and knowledge management models to support SMEs' decision-making in internationalisation in the context of IEAs. This task will unroll in close connection with Task 3, having as main inputs from it the scenarios and requirements specification. Departing from a knowledge based model of internationalisation, the specific goals are the following:

- To identify and characterise the information sources that will contribute to market knowledge and experiential (network, cultural, entrepreneur) knowledge;
- To develop an informational model adaptable to the processes of internationalisation in collaboration, mediated by IEAs;
- To design an analytic solution for the adaptation of the informational model according to the type and mode of internationalisation;
- To develop a collaborative decision support model that explore the information model and capitalise internationalisation knowledge.

The model development will jointly occur with the case studies task and, at some points in time during the task period, data will be collected from the members of IEAs and SMEs that will be involved in discussing and giving feedback on intermediate concepts and prototypes of the model. The information management model will allow to search, collect, integrate, organise and visualise internationalisation information, from different internet accessible, heterogeneous public and private sources (e.g. export assistance, export intelligence, international marketing research or the Internet), to be used and adapted in different contexts where SMEs are embedded.

The concept of information internalisation will also be considered in the development of the solution. As information and knowledge are complementary and transformable it is important that SMEs should find ways to transform information into knowledge in order to enhance the internationalisation process [Nguyen et al. 2006]. Information internalisation is the process of absorbing both tacit and explicit information into the organisation and translating it into knowledge, which is then applied to some purpose [Knight & Liesch 2002]. Therefore, this project will develop a model for information internalisation in the context of firms acting in collaboration, exploring the following modes of information internalisation: socialisation, combination, externalisation, and internalisation.

The information management model will be part of the collaborative decision support system that will be designed and adapted for internationalisation purposes, based on some novel multi-criteria decision-making models and on business intelligence. This decision support system will allow SMEs to select markets, choose partners and resources, assess risks and make negotiation with partners. Business intelligence will be included in the solution to facilitate the process of gathering, analysis and distribution of information and also to support the strategic decision-making process [Rouhani et al. 2012].

5) Model application and validation

The goal of this task is to demonstrate that through the model developed in Task 4, IEAs can intervene more actively in the internationalisation of their associates and SMEs can improve their decision-making capacity to internationalise. Other goal of this task is to test and validate the model in different scenarios faced by SMEs. Plans for the evolution and sustainability of the model will also be developed, as well as the dissemination of the scientific and technical results.

The implementation scenarios and the implementation plan of the developed model will be defined in this task. This task will also include the definition of evaluation metrics of the model

to be implemented. After that, the model will be implemented in different industries for validation purposes. During this stage of the project, modifications and improvements will be made according to the results obtained.

6. Thesis Status

The present research status of this thesis is described in Table 1.

Task	Description
Exploratory literature review performed	To define the path and the research area to be followed, as well as some problems and opportunities for further research.
Systematic literature review (SLR) performed	An SLR was performed to specify the research problem, to analyse the state-of-the-art in the scientific area where the student will develop his research work, and to discover new opportunities for research.
Research design and work programme established	With the result of the previous tasks, a research design and the work programme to be developed within the three years of the thesis project were developed and specified. Nevertheless, there is still opportunity for improvements, specifically with the inputs and suggestions that can be obtained with the Doctoral Consortium of CAPSI.

Table 1 – Thesis status with tasks performed

This document presents the results of almost all these tasks. The research work performed resulted in two publications, one accepted for presentation in an international conference (PRO-VE 2015) and the other submitted to an international journal (International Business Review). The future tasks to be performed in the next years are presented in Table 2.

Task	Description	Start date	End date
2	Survey design, conduction and analysis	01-07-2015	31-03-2016
3	Case-study design and conduction, and cross-case analysis and reporting	01-09-2015	30-09-2017
4	Information management and decision support models development	01-04-2016	31-03-2018
5	Model application and validation	01-04-2017	31-03-2018

Table 2 – Future tasks

With the participation in this Doctoral Consortium, it is expected new insights, feedback and contributions from the participants, mainly in what concerns the IS techniques and methodologies to be applied in the project, i.e. Task 4.

7. Conclusion and Expected Contributions

This doctoral thesis is multi-disciplinary in nature, contributing to the scientific knowledge on business management, information and knowledge management, and collaborative networks. The conclusion of the project will deliver the following results:

- A description and explanation of the factors influencing the management of information relevant for internationalisation purposes;
- A comprehensive characterisation of the decision processes involved in the internationalisation of SMEs and how the IEA influences them;
- A description and explanation of the factors influencing the implementation of collaborative processes for internationalisation;
- The identification and characterisation of the information sources that contribute to market knowledge and experiential (network, cultural, entrepreneur) knowledge,
- An informational model adaptable to the processes of internationalisation in collaboration, mediated by IEAs,
- An analytic solution for the adaptation of the informational model according to the type and mode of internationalisation,
- A collaborative decision support model that explore the information model and capitalise internationalisation knowledge.

Acknowledgments

This work is financed by the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project UID/EEA/50014/2013.

8. References

Casillas, J., Acedo, F., & Barbero, J. (2010). Learning, unlearning and internationalisation: Evidence from the pre-export phase. International Journal of Information Management, 30(2), 162–173.

Child, J., & Hsieh, L. (2014). Decision mode, information and network attachment in the internationalization of SMEs: A configurational and contingency analysis. Journal of World Business, 49(4), 598–610.

Dretske, F. (1981). Knowledge and the flow of information. Cambridge, MA: MIT Press.

Dutot, V., Bergeron, F., & Raymond, L. (2014). Information management for the internationalization of SMEs: An exploratory study based on a strategic alignment perspective. International Journal of Information Management, 34(5), 672–681.

Fletcher, M., & Harris, S. (2012). Knowledge acquisition for the internationalization of the smaller firm: Content and sources. International Business Review, 21(4), 631–647.

Hevner, A., & Chatterjee, S. (2010). Design research in information systems: Theory and practice. Springer (Vol. 22).

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75–105.

Hsu, W., Chen, H., & Cheng, C. (2013). Internationalization and firm performance of SMEs: The moderating effects of CEO attributes. Journal of World Business, 48(1), 1–12.

Ibeh, K., & Kasem, L. (2011). The network perspective and the internationalization of small and medium sized software firms from Syria. Industrial Marketing Management, 40(3), 358–367.

Knight, G., & Liesch, P. (2002). Information internalisation in internationalising the firm. Journal of Business Research, 55(12), 981–995.

Lage, E., & Alturas, B. (2013). Information Sharing among Innovative SME: An Exploratory Study within the Portuguese SME Innovation Network. In Innovation through Knowledge Transfer 2012 (pp. 159-168). Springer Berlin Heidelberg.

Musteen, M., Francis, J., & Datta, D. (2010). The influence of international networks on internationalization speed and performance: A study of Czech SMEs. Journal of World Business, 45(3), 197–205.

Nguyen, T., Barrett, N., & Fletcher, R. (2006). Information internalisation and internationalisation: Evidence from Vietnamese firms. International Business Review, 15(6), 682–701.

Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation. New York: Oxford University Press

Rouhani, S., Ghazanfari, M., & Jafari, M. (2012). Evaluation model of business intelligence for enterprise systems using fuzzy TOPSIS. Expert Systems with Applications, 39(3), 3764–3771.

Schweizer, R. (2012). The internationalization process of SMEs: A muddling-through process. Journal of Business Research, 65(6), 745–751.

Information management in supply networks: a framework to support risk management decisions

Dario Messina [1,2], Ana Cristina Barros [2],

António Lucas Soares [1,2]

1) Faculty of Engineering, University of Porto, Portugal

dario.messina@fe.up.pt , als@fe.up.pt

2) *INESC TEC,* Portugal
acbarros@inescporto.pt

Abstract

This thesis aims at contributing with new research at the intersection of the research fields of information management and supply chain risk management. Information is a fundamental element for an effective risk management in supply networks. In spite of this, research on how to use internal and external information for risk management decisions in supply networks is scarce. Therefore, this thesis main goal is to characterize information models used in supply networks and the supply chain risk management process in order to propose an aggregated model of internal and external information for risk management decision support in supply networks. This work is based on the Design Science paradigm, which will be applied through several research methods: systematic literature review to analyse the current state of the art on information management in supply networks; field research to analyse the influence of different information organisation models in the process of supply chain risk management and to develop an information model based (artifact); and action research to implement the artifact for the definition of risk mitigation strategies in aeronautic supply networks.

Keywords: Information Management, Supply Chain Risk Management, Information Model, Supply Network, Design Science.

1. Research Situation

This thesis is an interdisciplinary work at the intersection between information management and supply chain risk management. I started my Ph.D. in September 2014 in the "Engineering and Industrial Management" Doctoral Program, and taking into account the two areas of this work I think they perfectly fit in a modern and competitive industrial context. Regarding the benefits of my participation at CAPSI Doctoral Consortium please refer to the CV.

2. Context and Motivations

During the past decades Original Equipment Manufacturers (OEM's) have been increasingly focusing on their core competences (overall design, architecture, integration, and final assembly and delivery to end customers), though becoming large scale integrators by the coordination of a supply network [Ro et al., 2008]. Consequently, OEM's have today a higher integration with first-tier suppliers, but a looser control over second-tier suppliers, mainly due to lack of contractual relationship and the difficulty of strategic alignment between OEM and second-tier suppliers [Fawcett and Magnan, 2002]. Therefore, first-tier suppliers have assumed a critical

role in the supply chain, since they have to coordinate second-tier suppliers and help OEM's managing supply chain risk. This leads to the development of new supply chain capabilities in terms of supply chain risk management. These changes in the role of supply chain players calls for more research on how supply chains risk are being managed, and how information should be managed among supply chain partners to support risk management related decision-making. Consequently, this thesis aims at addressing this research gap and its research design is outlined in the following sections. Section 3 reviews the literature on information management in supply networks, supply chain integration, and risk management. Section 4 defines the problem underlying this thesis while in section 5 research questions, objectives and research methodologies are presented. Section 6 shows the actual status of this thesis work. The last section provides the main conclusions and points future research work.

3. Fundamentals and Related Work

The literature review of this thesis is structured in accordance with the main topics representing the pillars of this work. We start analysing how information is managed in supply networks, then we describe supply chain integration with a focus particularly on the risk management process, in order to identify the best strategies for supply networks risk mitigation.

Information Management in Supply Networks

A supply network is defined as a collection of interconnected organisations with the primary purpose of procure, use and transform resources in order to provide goods and services [Harland et al., 2004]. An important aspect for the management of supply networks is to establish how the information is acquired, organised and used among members, towards a proper risk management of the supply network [Hung et al., 2011]. Also, a correct management of the information flow is critical for increasing visibility over the network [Barratt and Oke, 2007] in order to support decision-making process and mitigate the risk [Yu and Goh, 2014]. Even though previous research has confirmed the strategic importance of information and several authors were interested in developing information models to proper manage this information flow [Chi, 2010], there is still a need to explore in depth this field in order to analyse how different information models influence the risk management decision process of supply networks.

Supply Chain Integration, and Risk Management

Due to the global and turbulent environment in which firms compete, several researchers have highlighted the need for a closer relationship between manufacturers and their partners as a way to improve supply chain performance [Childerhouse and Towill, 2011]. Even if across industries the level of implementation results differ, supply chain integration seems a good alternative to the verticalisation of operations [Figueiredo et al., 2008]. In order to ensure supply chain integration is necessary that supply chain partners share and collaboratively manage processes [Flynn et al., 2010]. According to the literature, in an integrated supply chain, the information is the key element that firms have to share with their partners [Ragatz et al., 2002]. Within the supply chain, the typical information shared is related with material requirements, demand forecasts and inventory levels (Chandra et al., 2007, Omar et al., 2010, Datta and Christopher, 2011, Hung et al.,2011). To ensure the effectiveness of long-term collaborations and to generate trust among partners, these information exchanges should be held within processes with high level of visibility, which should desirably enable effective communications and exchange of real-time data [Chengalur-Smith et al., 2012].

Companies also value the importance of establishing a risk management process to identify, measure, mitigate, and control risks in their supply network (Elsalmi and Hachicha, 2014). In order to be able to identify these risks, companies are starting to work with their network partners towards creating supply networks visibility (Nooraie and Parast, 2015). By sharing

what sometimes is sensitive and proprietary information with their network partners, companies aim at aligning their common objectives though ensuring the efficient management of the whole chain. Still, there is a need to create visibility over the specific information that will enable companies to identify and act upon the risks and opportunities of their supply network. In fact, in the global and digital context in which firms operate, information assumes a distinctive role, as it supports effective decision-making process. In this thesis we propose a framework for the organization of information among a supply network that enables the identification of strategies for supply network risk mitigation.

4. Research Problem definition

The aeronautic sector has shown a trend in the past decade in providing new solutions related to the management of new product development (Strategy&, 2015 (a), (b)). This sector is characterized by increasing costs in developing new products, and consequently there is the need to transfer the risk downstream the value chain. In this context, supply chain is characterized by a geographical dispersion of partners, and an increasing need of communication and coordination that is possible to achieve sharing vital information [Pereira, 2009] and through the adoption of ICT solutions.

Considering the scarcely research on how to use internal and external information for risk management decisions in supply networks, and the relevance that supply chain risk management and mitigation plays for practitioner (Global supply chain institute report, 2014), this research thesis will characterize the information models in supply network in order to develop and implement an information model, our artifact, that support strategic decision making in supply network in order to assess the risk along the whole chain, and provide a practical solution for decisions related to the configuration of the network. In particular, this work will be based on the specific needs of the aeronautic sector.

5. Objectives, Methodology and Research Methods

This thesis presents three research questions:

RQ1: What models exist for managing strategic information in supply networks?

RQ2: How information organization models influence the process of supply chain risk management?

RQ3: What are the specifics that characterise an information organisation model for supply chain risk management and mitigation?

RQ4: How the developed model support strategic decision-making related to aeronautic network configuration?

The research questions led to formulate the following goals:

- To characterize information management models used in supply networks;
- To describe the risk management process of supply networks and the role that first-tier and second-tier suppliers have in this process;
- To propose a multi-view and aggregated model of internal and external information that allow to bring to the surface the main causes of risk, and to identify the strategies to mitigate it.

5.1. Design Science Research Paradigm

The centrality of design activity has been recognised in several disciplines, and particularly it results even more crucial in the areas of Information Systems. Information systems for its own nature represent the meeting point between Information Technology (IT) and human interface, and for this reason it is characterised by challenging problems related to design activity in order to provide unique and creative solutions [Hevner and Chatterjee, 2010]. Based on the seminal book of Herbert Simon "*Sciences of the Artificial*", a pool of scholars in information systems sets the basis of a new paradigm, called Design Science Research Paradigm.

This thesis work is based on Design Science Research Paradigm, and the schema in fig. 1 represent the structure underlying it. The ultimate goal of design science is to address research in order to develop artifacts that, once tested, are able to provide innovative solutions for specific business needs [Holmstrom *et al.,* 2009].

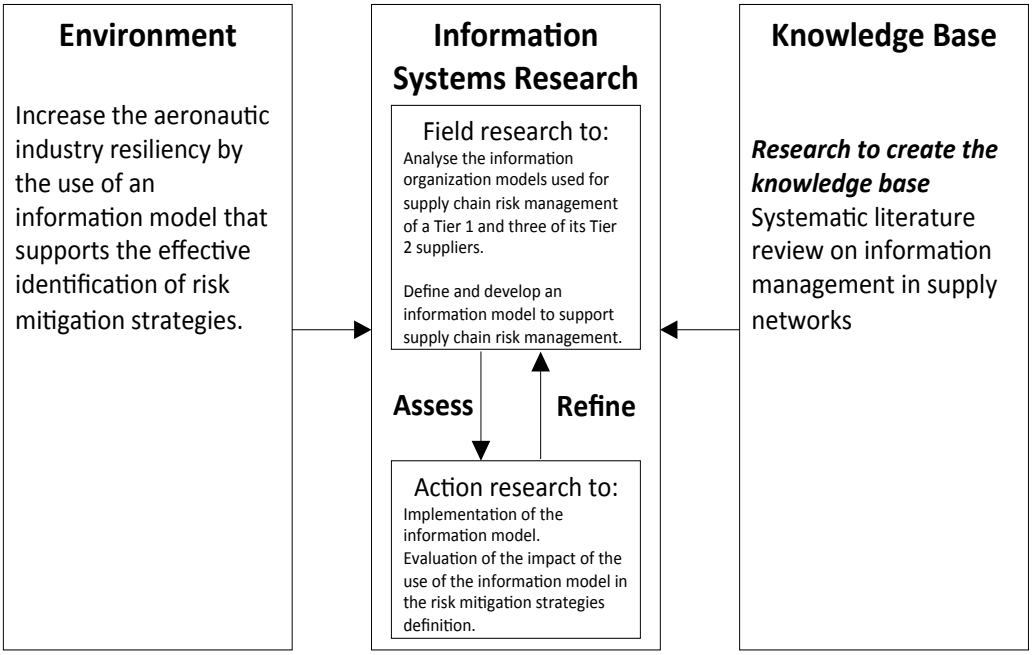


Figure 1 - Design Science Research Framework. Based on Havner et al.[2004]

Starting from the "environment" block we want to identify the needs on the aeronautic sector, then we adopt systematic literature review to built the foundation of the "knowledge base" block, and finally through the adoption of field research and action research, belonging to the "information systems research" block, we want to develop, implement and test an information model to support strategic decision in supply network. The purpose of design science is to develop new knowledge to provide to practitioner in order to design solutions for specific field problems [van Aken, 2005].

Due to the exploratory nature of this research a portfolio of qualitative research methods are used: systematic literature review to analyse the current state of the art on information management in supply networks; field research to analyse the supply chain risk management process of a first-tier and three of its second-tiers in supply networks; and action research to implement and test an information model for the evaluation of the impact of the use of the information model in the risk mitigation strategies in aeronautic supply networks. The different methodologies here are applied in order to fulfil the three different section of the design

research framework presented in fig.1. The following explains briefly the application of each of these methods.

Systematic Literature Review

The systematic literature review is the methodology that allows to build the theoretical foundation, and to ensure rigor in this thesis project. Considering the strategic relevance of information for supply chain management a systematic literature review will be carried out in order to discuss and synthesize the literature on information management in the context of supply networks. This research method is considered appropriate because it allows to identify the most relevant scientific contribution to a field or question [Becheikh et al., 2006], by adopting a replicable, scientific and transparent research process [Tranfield et al., 2003], In order to assure the quality of the review process, search terms will be selected from the literature, a detailed review protocol will be developed, and a team of four researchers will be involved in the paper selection and analysis phases.

Field Research

After describing how information is managed in supply networks, it is important to investigate how supply chain risk management is influenced by the information organisation models adopted, in order to identify the specifics necessary to develop our own information organisation model for supply chain risk management and mitigation. Considering the exploratory nature of this work, field research was adopted. A highly complex supply chains will be selected for this study, namely aircraft maker industry. Data collection will be carried out by means of semi-structured interviews at first-tier and three of its second-tier suppliers belonging to the network. All the interviews will be recorded, transcribed and then coded with the support of specific software (MAXQDA).

Action Research

The results from the two research works above will contribute for the design of an action research study. In particular the action research will act as *trait d'union* between the business needs provided in the field research and the applicable knowledge resulting from the systematic literature review, in order to implement and test the relevance of the developed artifact and the rigor underlying it.

Action research is appropriate to investigate the development of solutions for actual problems [Eriksson and Kovalainen, 2008] where researchers and operations managers collaboratively contributed to the development of solutions [Gummesson, 2010]. In this thesis, action research allows to implement our information model for risk management of supply networks, and to observe how the artifact is used in a manufacturing environment.

6. Thesis Status

A comprehensive analysis of the literature was made both in the field of Information Management and Supply Networks to identify the information management models adopted.

Information management models

At the moment, preliminary results exist only from the Systematic Literature Review. In particular the results of this methodology can be resumed in Table 1:

	STAGES			
Chi (2010)	Problem domain identification	Acquisition	Modelling	Ontological information representation

Haug (2013)	Acquisition	Representation	Evaluation	
Maçada *et al*., (2013)	Definition	Acquisition	Distribution	Usage

Table 1: Information Model in Supply Network

After the analysis of models obtained, the one proposed by Maçada *et al.*, (2013) was considered the most complete in the sample, for its characteristics. Even so, this model doesn't seem exhaustive, we found it lacking of some additional stage between the acquisition and the distribution stages. Analysing other generic information model in the literature, two of them presents some additional stages, namely in Choo (2002) and Detlor (2010). Choo's model is characterised by the following six stages: identification of information needs, acquisition, organisation and storage, development of product and services information, distribution, and usage; also Detlor's model uses six stages, that are: creation, acquisition, organisation, storage, distribution, and usage.

Merging the information from the three models we develop our own characterised by the following six stages: identification of information needs, acquisition, organisation, storage, distribution, and usage. Regarding the other methodologies, we plan to conduct interviews in Spain and Portugal in order to create the base for field research, and so provide all the elements needed to move to the third and last methodology, the action research.

7. Conclusions and Expected contributions

The topic of information management in supply chains has been extensively studied in the past, though most researchers focus on the dyad OEM - first-tier or the dyad first-tier - second-tier.

This paper sheds lights on how the risk management process is currently carried out in a complex multi-tier supply network and the role supply chain partners have in this process. Furthermore, this thesis contributes to the supply chain management field with a multi- view and aggregated model of internal and external information for decision support in supply networks in aeronautic sector. The implementation and pilot use of this model in a real setting will confirm the relevance of this work for practitioners and develop further managerial insights on the implementation of such models in practice.

Acknowledgements

This work is financed by the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project CMUP-ERI/TPE/0011/2013 of the CMU Portugal Program.

8. References

Barratt, M., Oke, A. "Antecedents of supply chain visibility in retail supply chains: a resource-based theory perspective." *Journal of operations management*, *25*(6), (2007): 1217-1233.

Becheikh, N., Landry, R., Amara, N. "Lessons from innovation empirical studies in the manufacturing sector: A systematic review of the literature from 1993–2003." *Technovation*, *26*(5), (2006): 644–664.

Chandra, C., Grabis, J., & Tumanyan, A. "Problem taxonomy: a step towards effective information sharing in supply chain management." *International Journal of Production Research*, *45*(11), (2007), 2507-2544.

Chengalur-Smith, I., Duchessi, P., Gil-Garcia, J. R. "Information sharing and business systems leveraging in supply chains: An empirical investigation of one web-based application." *Information & Management*, *49*(1), (2012): 58-67.

Chi, Y. L. "Rule-based ontological knowledge base for monitoring partners across supply networks." *Expert Systems with Applications*, *37*(2), (2010): 1400-1407.

Childerhouse, P., Towill, D. R. "Arcs of supply chain integration." *International journal of production research*, *49*(24), (2011): 7441-7468.

Choo, C.W. *Information management for the intelligent organization: the art of scanning the environment*. Information Today, Inc., 2002.

Datta, P.P. and Christopher, M.G. "Information sharing and coordination mechanisms for managing uncertainty in supply chains: A simulation study", *International Journal of Production Research*, Vol. 49 No. 3, (2011), pp. 765–803.

Detlor, B. Information management. *International Journal of Information Management, 30*, (2010) 103-108.

Elmsalmi, M., & Hachicha, W. (2014, May). Risk mitigation strategies according to the supply actors' objectives through MACTOR method. In *Advanced Logistics and Transport (ICALT), 2014 International Conference on* (pp. 362-367). IEEE.

Eriksson, P., & Kovalainen, A. *Qualitative methods in business research*. Sage. 2008

Fawcett, S. E., Magnan, G.M. "The rhetoric and reality of supply chain integration." International Journal of Physical Distribution & Logistics Management 32.5 (2002): 339-361.

Figueiredo, P., Silveira, G., Sbragia, R. "Risk sharing partnerships with suppliers: the case of Embraer." *Journal of Technology Management & Innovation*, *3*(1), (2008): 27-37.

Flynn, B. B., Huo, B., Zhao, X. "The impact of supply chain integration on performance: A contingency and configuration approach." *Journal of Operations Management*, *28*(1), (2010): 58–71.

Global supply chain institute report, 2014: http://globalsupplychaininstitute.utk.edu/publications/documents/Risk.pdf

Gummesson, E., "Qualitative Methods in Management Research", Sage, 2000.

Harland, C., Zheng, J., Johnsen, T., Lamming, R. "A Conceptual Model for Researching the Creation and Operation of Supply Networks". *British Journal of Management*, *15*(1), (2004): 1–21.

Hevner, A., Chatterjee, S. Design Science Research in Information Systems Good. In *Design Research in Information Systems* (pp. 9–22). Springer Science+Business Media. 2010

Hevner, A., March, S. T., Park, J., & Ram, S. Design science in information systems research. *MIS quarterly*, *28*(1), (2004): 75-105.

Holmström, Jan, Mikko Ketokivi, and Ari-Pekka Hameri. "Bridging practice and theory: a design science approach." *Decision Sciences* 40.1 (2009): 65-87.

Hung, Wei-Hsi, Chin-Fu Ho, Jau-Jeng Jou, Yi-Ming Tai. "Sharing information strategically in a supply chain: antecedents, content and impact." *International Journal of Logistics: Research and Applications* 14, no. 2 (2011): 111-133.

Maçada, A. C. G., Costa, J. C., Oliveira, M., Curado, C. "Information management and knowledge sharing in supply chains operating in Brazil." *International Journal of Automotive Technology and Management, 13*(1), (2013): 18–35.

Nooraie, S. V., & Parast, M. M. (2015). A multi-objective approach to supply chain risk management: Integrating visibility with supply and demand risk. *International Journal of Production Economics*, *161*, 192-200.

Omar, R., Ramayah, T., Lo, M.-C., Sang, T.Y. and Siron, R. "Information sharing, information quality and usage of information technology (IT) tools in Malaysian organizations", *African Journal of Business Management*, Vol. 4 No. 12, (2010), pp. 2486–2499.

Pereira, J.V. "The new supply chain's frontier: Information management." *International Journal of Information Management* 29.5 (2009): 372-379.

Ragatz, G. L., Handfield, R. B., Petersen, K. J. "Benefits associated with supplier integration into new product development under conditions of technology uncertainty." Journal of Business Research, 55(5), (2002): 389-400.

Ro, Y. K., Liker, J.K., Fixson, S.K. "Evolving models of supplier involvement in design: the deterioration of the Japanese model in US auto." Engineering Management, IEEE Transactions on 55.2 (2008): 359- 377.

Strategy&, 2015 (a): http://www.strategyand.pwc.com/perspectives/2015-aerospace-defense-trends

Strategy&, 2015 (b): http://www.strategyand.pwc.com/perspectives/2015-aviation-trends

Tranfield, D., Denyer, D., Smart, P. "Towards a methodology for developing evidence in formed management knowledge by means of systematic review." British journal of management 14.3 (2003): 207-222.

Van Aken, J. E. "Management research as a design science: Articulating the research products of mode 2 knowledge production in management." *British Journal of Management*, *16*, (2005): 19–36.

Yu, M. C., Goh, M. "A multi-objective approach to supply chain visibility and risk." *European Journal of Operational Research* 233, no. 1 (2014): 125-130.

Governança de TI em universidades públicas: Proposta de um modelo

Isaías Scalabrin Bianchi[1] e Rui Dinis de Sousa[2]

[1]Universidade do Minho, Centro ALGORITMI, Portugal
isaias.bianchi@gmail.com

[2]Universidade do Minho, Centro ALGORITMI, Portugal
rds@dsi.uminho.pt

Resumo

O objetivo deste trabalho de investigação é o desenvolvimento de um modelo de governança de TI em universidades públicas para que mais eficaz e eficientemente cumpram a sua missão. A investigação irá identificar os mecanismos de governança de TI utilizados em diferentes realidades universitárias e avaliar a eficiência e eficácia na sua implementação considerando fatores contingenciais. Recorrer-se-á à administração de *surveys* em contextos sobre os quais já se detém e se tem facilidade em adquirir conhecimento, nomeadamente, universidades federais brasileiras e universidades públicas portuguesas. Design Science Research será utilizada pela sua adequação à resolução de problemas organizacionais complexos que requerem o desenvolvimento de artefactos, neste caso, um modelo para governança de TI em universidades públicas. O modelo proposto gerado será avaliado por especialistas e profissionais da área TI através de *surveys,* entrevistas e *workshops*. Analisar-se-á a eficácia e facilidade de implementação de mecanismos e ferramentas que auxiliem na governança de TI. Espera-se contribuir com um modelo e um guia para a sua implementação nas universidades públicas enquanto se reforça o corpo de conhecimento na governança de TI.

Palavras chave: Governança de TI; Universidades; Modelo; Referenciais; DSR

1. Situação do candidato

O presente trabalho está a ser desenvolvido no seio da comunidade de investigadores em engenharia e design organizacional particularmente focada nas arquiteturas empresariais e na gestão orientada por processos, domínios em investigação no *Information Systems and Technologies for the Transformation of Organizations and Society* (ISTTOS), grupo de investigação na linha *Information Systems and Technologies* do Centro ALGORITMI da Universidade do Minho.

O candidato encontra-se no segundo ano do programa doutoral em Tecnologia e Sistemas de Informação da Universidade do Minho, no qual ingressou em Fevereiro de 2014 e cuja defesa de tese está prevista para janeiro de 2018. Em março último, apresentou um plano da tese. Neste momento, está a aprofundar a revisão de literatura inicial, bem como a metodologia de investigação a seguir e a estabelecer contactos para a seleção de universidades nas quais possa ser administrado um *survey* e realizados estudos de caso; em março do próximo ano, o candidato pretende defender a proposta da tese.

Espera-se que o Doctoral Consortium constitua uma oportunidade para discutir a realidade da governança de TI nas universidades e a sua pertinência para a investigação, enquanto se recolhem contributos para desenvolver a proposta de tese.

2. Contexto e Motivação

Entre vários fatores motivacionais, destaca-se o contributo que se procura para a governança de TI nas universidades, especialmente nas públicas, devido à importância que aquelas representam para o desenvolvimento do país. Pela obrigação e missão que têm na sociedade pela transferência de conhecimento e tecnologia e prestação de serviços de qualidade, as universidades constituem um forte agente do desenvolvimento económico e científico mundial (Sarkar, 2012).

As universidades são organizações complexas devido ao seu modelo de gestão, metas e objetivos sendo que o seu ambiente tecnológico consiste numa variedade de aplicações, diferentes plataformas, sistemas académicos, aplicações na nuvem e uma diversidade heterogénea de tecnologias (Svensson & Hvolby, 2012). A governança desse ambiente complexo para suportar a investigação e ensino, exige a busca de mecanismos de governança de TI adequados, modelos de maturidade e *frameworks* (Bajgoric, 2014; Conger, Winniford, & Erickson-Harris, 2008; Hicks, Pervan, & Perrin, 2012). Os referenciais existentes, genéricos, sem que tenham em conta especificidades do tipo de organização, não se revelam propensos à sua adoção. Não existem, contudo, dúvidas de que a adoção de bons modelos de governança de TI pode trazer inúmeros benefícios às organizações.

Um estudo realizado em duzentas e cinquenta organizações de vinte e três países, mostra que as organizações com bons modelos de governança de TI apresentam um desempenho superior em 20% em relação às suas concorrentes (Weill & Ross, 2004). A investigação de Lunardi et al. (2014) também conclui que as organizações que adotaram mecanismos formais de governança de TI melhoraram o desempenho organizacional em termos de rentabilidade, eficiência e redução de custos.

A governança de TI, desenvolvida com um conjunto de mecanismos que envolvem estrutura, processos e mecanismos de relacionamento, pode auxiliar na gestão desse ambiente complexo (Steven De Haes & Van Grembergen, 2004, 2005, 2009; Weill & Ross, 2004). Propor os mecanismos mais adequados para um determinado contexto é uma tarefa complexa e depende de fatores externos e internos sendo que mecanismos que podem ser eficazes para uma organização, podem não ser eficazes e apropriados para outras organizações (Brown & Grant, 2005; Steven De Haes & Van Grembergen, 2008; Van Grembergen, De Haes, & Guldentops, 2004). Investigar tais mecanismos por forma a dotar as universidades públicas de um modelo que facilite a governança de TI será uma das principais motivações deste trabalho.

À motivação principal acresce uma outra, que decorre da possibilidade de evoluir no conhecimento capitalizando em trabalho já desenvolvido na dissertação de mestrado em administração universitária na qual o investigador desenvolveu e implantou um sistema de informação para gestão de serviços de TI baseado nas práticas *ITIL* para um departamento de TI, num centro de ensino no qual trabalhou. Os benefícios foram diversos na gestão do centro de ensino da universidade após a implementação do gerenciamento de serviços de TI, com impactos nas atividades de ensino e investigação.

O facto de o investigador estar inserido no contexto universitário enquanto funcionário da área de TI há quatro anos, reforça o seu interesse na área da governação de TI. Durante os últimos anos esteve envolvido integralmente em atividades relacionadas com TI em universidades. Visitou e contactou gestores de diversas universidades em vários países para melhor entender a relevância das tecnologias da informação nas universidades. Pode observar que a governança de TI em universidades era ainda artesanal quando comparada com a indústria de TI em que trabalhou mais de dez anos em projetos. A adoção de melhores práticas, planos e referenciais é escassa, em especial, nas universidades federais brasileiras. Nota-se, no entanto, nos últimos três anos, que as universidades começaram a preocupar-se com a profissionalização

da governança de TI. No entanto, as publicações a nível mundial com recomendações sobre as temáticas de governança e gestão de TI para o sector são, ainda, escassas.

A proposta de uma tese de doutoramento nesta área é, pois, também motivada pelas observações já referidas do investigador que lhe suscitam o interesse em compreender mais sobre o assunto por forma a poder contribuir para o conhecimento e resolução de problemas na governação de TI, domínio que continua a despertar o interesse de investigadores.

Um estudo de Flaviano & Lucca (2013) sobre a realidade da governança de TI nas universidades federais brasileiras conclui que apenas 41% das 59 instituições de educação superior federais brasileiras possuem um Planeamento de Tecnologia da Informação. Conclui ainda que as universidades desenvolveram esse planeamento mais por obrigações legais do que pela importância de adotar e manter boas práticas de gestão. Nesse contexto da investigação constatou-se também que apenas três universidades operacionalizaram o seu planeamento de tecnologia de informação e que exercem algum tipo de governança em relação à gestão de tecnologia da informação. Entre as universidades que implementaram, não foi possível identificar ferramentas, técnicas, modelos e fatores chaves de sucesso dessa implementação, bem como o nível de maturidade e práticas referentes à governança de TI que utilizaram. Tal poderá indicar o baixo índice da adoção de mecanismos formais de governança de TI.

Os principais trabalhos sobre governança de TI têm estado limitados à indústria. Partindo desses trabalhos, a investigação noutros sectores, como o ensino superior, onde a discussão destas questões está ausente ou é incipiente, constitui uma oportunidade de investigação e de contribuição para o melhor desempenho de um sector vital para o desenvolvimento das sociedades.

2. Problema

A adoção de um modelo de governança de TI pela organização é importante, mas a sua eficácia dependerá da sua implementação. A definição dos mecanismos de governança a utilizar constitui uma etapa importante, mas será a implementação desses mecanismos como uma solução eficiente que se revelará audaciosa (Brown & Grant, 2005; Steven De Haes & Van Grembergen, 2005).

A literatura apresenta uma variedade de referenciais para auxiliar as organizações a implementarem a governança de TI. Contudo, esses referenciais são complexos, complicados e difíceis de serem interpretados (Bin-Abbas & Bakry, 2014; Rúben Pereira & da Silva, 2011; R. F. d. S. Pereira & Silva, 2010). Um *survey* pelo *IDC* de 225 organizações identifica os seguintes referenciais utilizados: ITIL (27,1%); Six Sigma (23,6%); ISO 20000 (14,7%); COBIT (12,9%); CMM/CMMI (8%); outro (2.2%), sendo o desenvolvimento de modelos próprios internas, aquele que se apresenta com maior destaque: 43,6% (Broussard & Tero, 2007).

Um outro estudo, de Lunardi et al. (2014) em 101 organizações brasileiras que possuem capital aberto na bolsa de valores, identifica os seguintes referenciais de governança de TI que aquelas utilizam: *COBIT* (54%), *ITIL* (44%), *SOX* (36%), *BS7799/ISO17799 (27%), PMI (23%), SLA/SLM (18%), BSC / IT BSC (10%),*sendo mais uma vez expressiva a adoção de modelos próprios (32%). Outros *(27%)* há que são referidos, como *COSO (Committee of Sponsoring Organizations); the IT service catalog; shared domain knowledge; Six Sigma; SOA (Service Oriented Architecture); the IT project-linked compensation practices; BPM (Business Process Management); ISO9000.*

O *ITIL* e o *COBIT* destacam-se como os dois principais referenciais utilizados pelas organizações para implementarem a governança de TI. Ko & Fink (2010), Selig (2008) e Van Grembergen (2007) afirmam que o *ITIL* é um referencial com todas as características necessárias para suportar a governança de TI. Para Iden & Eikebrokk (2014), o *ITIL* é popular e influenciador na implementação da governança de TI. Por outro lado, De Haes, Van Grembergen, & Debreceny (2013) afirmam que as organizações tendem a adotar *COBIT* para a governança de TI. Todavia, ITIL e COBIT são vistos por Bin-Abbas & Bakry (2014), Pereira & da Silva(2010) (2011) (2012a) como complexos, genéricos e difíceis de serem interpretados,

sem que se saiba o que implementar primeiro e o que realmente é necessário para um determinado contexto organizacional**.**

Percebe-se, pois, uma lacuna na literatura em identificar qual o referencial e quais os mecanismos a incluir num modelo, bem como o nível de maturidade organizacional em que devem ser utilizados para atender às necessidades na governança de TI. Os resultados dos *surveys* mostram que as organizações optam ainda por desenvolverem os seus próprios modelos, aproveitando os pontos fortes de vários referenciais. A identificação de cada mecanismo de governança e a eficácia desses exige um certo esforço também.

M. Marrone et al. (2014) afirmam que a adoção de referenciais de governança de TI depende de diferentes fatores e, usando a teoria institucional num estudo, apontam os seguintes fatores: tamanho da organização, país e tipo da organização. Pereira & da Silva (2010) (2012b), em linha com o estudo anterior, referem que a implementação da governança de TI depende de fatores contingenciais como a dimensão, se é pública ou privada, bem como o contexto da organização, ambiente interno e externo.

Os estudos identificados estão limitados à indústria e concentrados principalmente na Austrália e Estados Unidos, em especial, em práticas de gestão de serviços de TI. Na Europa, os principais trabalhos sobre governança de TI, foram realizados na Bélgica e em Portugal, mas focados na indústria financeira De Haes & Van Grembergen (2005) Pereira, Almeida, & Silva (2014), no Brasil Lunardi et al. (2014) com a identificação dos mecanismos de governança de TI no desempenho organizacional em indústrias do tipo financeira. Justifica-se a investigação em outros sectores de atividade e geografias. O presente trabalho desenvolver-se-á no contexto do ensino superior, em particular, europeu e brasileiro, dadas as ligações já estabelecidas pelo investigador.

A literatura apresenta poucos trabalhos sobre modelos de governança em universidades, ainda com escopo limitado e superficiais: Espanha, Fernández e Llorens (2009); Inglaterra. Coen e Kelly (2007); China, Zhen e Xin-yu (2007) e Wan e Chan (2008); Austrália. Bhattacharjya e Chang (2006); Portugal, Ribeiro e Gomes (2009), Malásia, Saleh e Almsafir (2013), sendo o estudo com maior rigor e em número de universidades, o de Kallaya et al. (2015) na Tailândia, porém limitado a normas e leis do país em específico. Não se identificou qualquer estudo com preocupação na construção de um modelo com base em referenciais que contemplem detalhadamente os mecanismos de governança em função de fatores contingenciais e específicos do setor.

Nesta proposta de projeto de tese, entende-se que a governança de TI vai muito para além da simples adoção de um referencial. A escolha dos mecanismos de governança adequados é um primeiro passo e importante. Mas não menos importante e árduo, será a consideração de fatores contingenciais na seleção desses mecanismos que permitam definir um modelo para um determinado contexto que contemplem o uso de ferramentas tecnológicas adequadas, bem como a identificação e *benchmarking* entre universidades da eficácia das práticas.

Pretende-se, assim, que esta investigação contribua com um modelo que contemple os mecanismos necessários para ser implementado numa realidade universitária, algo que não se identificou na revisão de literatura, para que melhor as universidades públicas possam cumprir a sua missão. Formula-se, então, a seguinte **questão de investigação:**

- **Como governar as Tecnologias da Informação nas universidades públicas para que mais eficaz e eficientemente cumpram a sua missão?**

3. Objetivos da Investigação

Por forma a responder à questão de investigação que acaba de ser formulada, definem-se os seguintes objetivos da investigação:

1. **Identificar** na literatura mecanismos que permitam a governança de TI;
2. **Identificar** a eficiência e eficácia na implementação de mecanismos de governança de TI utilizados em diferentes realidades universitárias considerando fatores contingenciais, através da administração de surveys em contextos sobre os quais já se tem conhecimento e se tem facilidade em adquiri-lo, nomeadamente, universidades federais brasileiras e universidades públicas portuguesas;
3. **Avaliar** e **comparar** os mecanismos de estrutura, processos e relacionamentos na governança de TI por meio de estudo de casos, um no Brasil (Universidade Federal de Santa Catarina) e outro em Portugal (Universidade do Minho);
4. **Propor** um *modelo* de governança de TI para universidades públicas a partir da revisão de literatura, análise de mecanismos na estrutura, processos e relacionamentos identificados no *survey,* e estudo de casos da sua adoção e implementação; e
5. **Avaliar** o *modelo* proposto com especialistas e profissionais da área TI através de *survey,* entrevistas e *workshops*.

4. Método de Investigação

Design Science Research (DSR) é um dos métodos de investigação que se tem afirmado e tem vindo ser adotado na área de sistemas de informação para a resolução de problemas organizacionais, contribuindo para a resolução de problemas complexos e em domínios específicos (Hevner, March, Park, & Ram, 2004; Kuechler & Vaishnavi, 2012). Na proposta do presente trabalho, que passe pelo desenvolvimento de um modelo para resolver um problema específico, nesse caso da governança de TI das universidades, o método *DSR* revela-se adequado para *o* artefacto que se pretende propor. Aplicar-se-á o modelo que se apresenta na Figura 1:

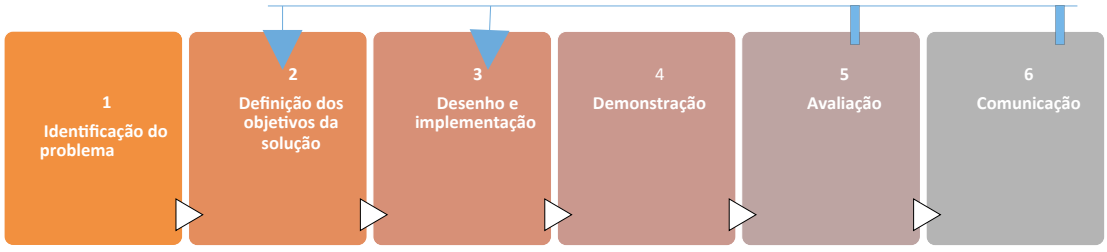


Figura 1. Etapas em Design Science Research (adaptado de Peffers et al.(2006))

A tabela 1 instancia a aplicação do método ao trabalho de investigação em causa:

***Tabela 1**. Etapas do método DSR*

Etapa		Descrição
1	Identificação do problema e motivação	O nível de maturidade da governança de TI em universidades é considerado baixo, conforme verificado pela revisão de literatura e por experiência pessoal.
2	Definição dos objetivos e solução	Propor um modelo que seja eficaz e eficiente na governança de TI em universidades públicas, incluindo orientações para a sua adoção em termos de estruturas, processos e mecanismos de relacionamento.
3	Desenho e desenvolvimento	Desenho de um modelo cujo desenvolvimento passa pela identificação dos mecanismos de governança apropriados ao sector de atividade através da administração de um survey e realização de estudos de caso em universidades públicas portuguesas e brasileiras.
4	Demonstração	Demonstração da eficiência e eficácia do modelo desenvolvido recorrendo a profissionais de TI de universidades na simulação de cenários.
5	Avaliação	Avaliação do modelo por meio de entrevistas, workshops e surveys com profissionais especialistas da indústria e de universidades.
6	Comunicação	Comunicação pela disseminação do modelo por profissionais e gestores de TI nas universidades e pela publicação de resultados blogs, fóruns, conferências e revistas.

5. Contributos Esperados

Este trabalho de investigação deverá contribuir com um modelo que inclua estruturas, processos e mecanismos de relacionamento particularmente adequados ao ambiente universitário, permitindo às universidades públicas governar o seu ambiente de TI eficaz e eficientemente. Pretende-se contribuir para o corpo de conhecimento associado à utilização de referenciais como ITIL e COBIT em função de fatores contingenciais, nomeadamente, daqueles mais relevantes no contexto das universidades públicas portuguesas e brasileiras. Espera-se, ainda, que o guia a disponibilizar para a adoção do modelo auxilie consideravelmente os gestores na implementação dos mecanismos adequados, com o menor custo e o maior impacto no cumprimento da missão da universidade pública.

Agradecimentos

Este trabalho foi suportado pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), Processo n.º 10415/13-0 e pela FCT - Fundação para a Ciência e Tecnologia no âmbito do projeto UID/CEC/00319/2013.

6. Referências

Bajgoric, N. (2014). Business continuity management: a systemic framework for implementation. *Kybernetes, 43*(2), 156-177. doi:10.1108/K-11-2013-0252

Bhattacharjya, J., & Chang, V. (2006). *Adoption and implementation of IT governance: cases from Australian Higher Education.* Paper presented at the 17th Australasian Conference on Information Systems.

Bin-Abbas, H., & Bakry, S. H. (2014). Assessment of IT governance in organizations: A simple integrated approach. *Computers in Human Behavior, 32*, 261-267. doi:10.1016/j.chb.2013.12.019

Broussard, F. W., & Tero, V. (2007). Configuration and Change Management for IT Compliance and Risk Management: The Tripwire Approach. *White Paper. IDC*.

Brown, A. E., & Grant, G. G. (2005). Framing the frameworks: A review of IT governance research. *Communications of the Association for Information Systems, 15*(1), 38.

Coen, M., & Kelly, U. (2007). Information management and governance in UK higher education institutions: bringing IT in from the cold. *Perspectives: Policy and Practice in Higher Education, 11*(1), 7-11. doi:10.1080/13603100601127915

Conger, S., Winniford, M., & Erickson-Harris, L. (2008). Service management in operations. *AMCIS 2008 Proceedings*.

De Haes, S., & Van Grembergen, W. (2004). IT governance and its mechanisms. *Information Systems Control Journal, 1*, 27-33.

De Haes, S., & Van Grembergen, W. (2005). *IT governance structures, processes and relational mechanisms: Achieving IT/business alignment in a major Belgian financial group*.

De Haes, S., & Van Grembergen, W. (2008). An exploratory study into the design of an IT Governance minimum baseline through Delphi research. *Communications of the Association for Information Systems, 22*(1), 24.

De Haes, S., & Van Grembergen, W. (2009). An exploratory study into IT governance implementations and its impact on business/IT alignment. *Information Systems Management, 26*(2), 123-137.

De Haes, S., Van Grembergen, W., & Debreceny, R. S. (2013). COBIT 5 and enterprise governance of information technology: Building blocks and research opportunities. *Journal of Information Systems, 27*(1), 307-324. doi:10.2308/isys-50422

Fernández, A., & Llorens, F. n. (2009). *An IT Governance Framework for Universities in Spain* Paper presented at the European University Information Systems Organization (EUNIS) Santiago de Compostela.

Flaviano, V., & Lucca, G. (2013). *Planejamento Estratégico de Tecnologia de Informação nas Universidades Federais Brasileiras*. Paper presented at the XXXVII Encontro da ANPAD - EnANPAD Rio de Janeiro.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75-105.

Hicks, M., Pervan, G., & Perrin, B. (2012). A study of the review and improvement of IT governance in Australian universities.

Iden, J., & Eikebrokk, T. R. (2014). Using the ITIL Process Reference Model for Realizing IT Governance: An Empirical Investigation. *Information Systems Management, 31*(1), 37-58. doi:10.1080/10580530.2014.854089

Kallaya, J., Prasong, P., & Pilastpongs, S. (2015). Information technology governance practices based on sufficiency economy philosophy in the Thai university sector. *Information Technology & People, 28*(1), 195-223. doi:10.1108/ITP-10-2013-0188

Ko, D., & Fink, D. (2010). Information technology governance: an evaluation of the theory-practice gap. *Corporate Governance: The international journal of business in society, 10*(5), 662-674.

Kuechler, W., & Vaishnavi, V. (2012). A framework for theory development in design science research: multiple perspectives. *Journal of the Association for Information systems, 13*(6), 395-423.

Lunardi, G. L., Becker, J. L., Maçada, A. C. G., & Dolci, P. C. (2014). The impact of adopting IT governance on financial performance: An empirical analysis among Brazilian firms. *International Journal of Accounting Information Systems, 15*(1), 66-81. doi:http://dx.doi.org/10.1016/j.accinf.2013.02.001

Marrone, M., Gacenga, F., Cater-Steel, A., & Kolbe, L. (2014). IT service management: A cross-national study of ITIL adoption. *Communications of the Association for Information Systems, 34*(1), 865-892.

Peffers, K., Tuunanen, T., Gengler, C. E., Rossi, M., Hui, W., Virtanen, V., & Bragge, J. (2006). *The design science research process: a model for producing and presenting information systems research*.

Pereira, R., Almeida, R., & Silva, M. M. d. (2014). *IT Governance Patterns in the Portuguese Financial Industry.* Paper presented at the 47th Hawaii International Conference on Systems Sciences, HICSS, Hawaii, USA.

Pereira, R., & da Silva, M. M. (2011). *A maturity model for implementing ITIL V3 in practice.* Paper presented at the Enterprise Distributed Object Computing Conference Workshops (EDOCW), 15th IEEE International.

Pereira, R., & da Silva, M. M. (2012a). Designing a new Integrated IT Governance and IT Management Framework Based on Both Scientific and Practitioner Viewpoint. *International Journal of Enterprise Information Systems (IJEIS), 8*(4), 1-43.

Pereira, R., & da Silva, M. M. (2012b). IT Governance Implementation: The Determinant Factors. *Communications of the IBIMA, 2012*.

Pereira, R. F. d. S., & Silva, M. M. d. (2010). *A maturity model for implementing ITIL v3.* Paper presented at the 6th World Congress on Services.

Ribeiro, J., & Gomes, R. (2009). *IT governance using COBIT implemented in a high public educational institution: a case study*. Paper presented at the Proceedings of the 3rd international conference on European computing conference, Tbilisi, Georgia.

Saleh, J. M., & Almsafir, M. K. (2013). *The Drivers of ITIL Adoption in UNITEN.* Paper presented at the Advanced Computer Science Applications and Technologies (ACSAT).

Sarkar, S. (2012). The role of information and communication technology (ICT) in higher education for the 21st century. *Science, 1*(1), 30-41.

Selig, G. J. (2008). *Implementing IT Governance-A Practical Guide to Global Best Practices in IT Management*: Van Haren.

Svensson, C., & Hvolby, H.-H. (2012). Establishing a Business Process Reference Model for Universities. *Procedia Technology, 5*(0), 635-642. doi:http://dx.doi.org/10.1016/j.protcy.2012.09.070

van Grembergen, W. (2007). *Implementing Information Technology Governance: Models, Practices and Cases: Models, Practices and Cases*: IGI Global.

Van Grembergen, W., De Haes, S., & Guldentops, E. (2004). Structures, processes and relational mechanisms for IT governance. *Strategies for information technology governance, 2*(004), 1-36.

Wan, S. H. C., & Chan, Y.-H. (2008). Improving service management in campus IT operations. *Campus-Wide Information Systems, 25*(1), 30-49.

Weill, P., & Ross, J. W. (2004). *IT governance: How top performers manage IT decision rights for superior results*: Harvard Business Press.

Zhen, W., & Xin-yu, Z. (2007). *An ITIL-based IT service management model for chinese universities.* Paper presented at the *International Conference on Software Engineering Research.*

The emotion-aware city: Development of a framework based on social media towards emotion & stress mapping

Tiago H. Moreira de Oliveira [1], Marco Painho [1]

1) NOVA IMS, Universidade Nova de Lisboa, Portugal

tiago.oliveira@novaims.unl.pt

2) NOVA IMS, Universidade Nova de Lisboa, Portugal
painho@novaims.unl.pt

Abstract

Since one of the main ambitions of a smart city is to improve urban functions and provided services, it is often perceived as a living urban fabric, in which connected urban citizens, acting as active sensors, have the capacity to contribute even more efficiently to the spatial intelligence of cities. This "immaterial" dimension is related with the need that smart cities have to assess their citizen's feelings, perception and well-being, giving rise to an emotion-aware city.

Mapping emotion builds on a tradition of studies in cognitive mapping, evaluative mapping, environmental preference and environmental affect, adding an approach in which people experience, evaluate and describe their environment "in situ" through social media.

This research aims to use an Ambient Geographic Information (AGI) approach by assembling geo-tagged data from Twitter, Flickr, Instagram and Facebook related with people's perception and feelings regarding Lisbon (Portugal), and therefore characterize its emotional dimension, by comparing these subjective observations with objective measurements (such as socio-demographic statistics, questionnaires and data retrieved from biometric sensors).

With this vision of a smart city, that is capable to interpret and harnessing the emotional states of its citizens, it is essential to find new methods and techniques to sensing affect in an urban context.

Keywords: Ambient geographic information (AGI); Social media; emotion-aware city; smart city; Ubiquitous and mobile computing.

1. Introduction

We can sum the relevance of cities in modern societies with just four numbers: cities cover only 2% of the Earth's inhabited land area; 50% of the population on the globe live in cities (80% by 2050 according to the United Nations); cities account for 75% of total consumed energy and 80% of CO2 emissions [S. Roche 2014]. The infographic bellow depicts these challenges.

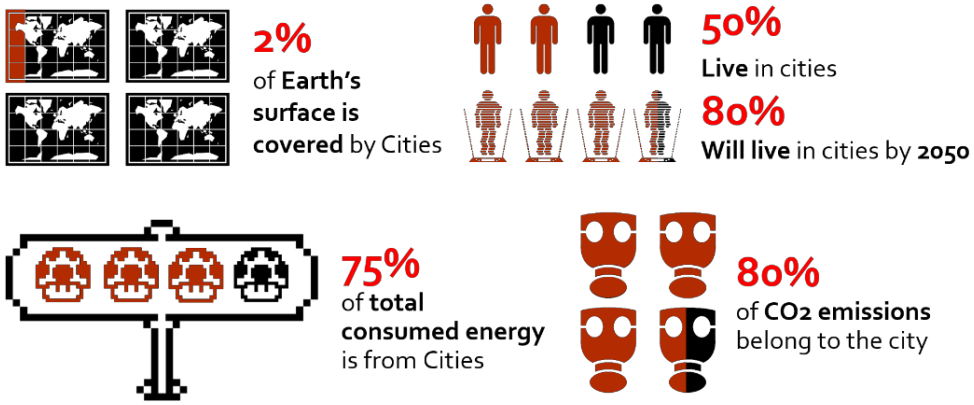


Figure 1 – Cities challenges infographic (Author)

The idea and concept of smart city itself tries to address these challenges. A smart city's main goal is to increase the quality of life for its citizens and to make the city more attractive, lively and greener [Kehoe & Nesbitt 2010]. To achieve this goal, physical sensors are deployed throughout the city to monitor various aspects such as environmental parameters (weather, pollution, etc.), traffic and the consumption of resources [Caragliu, Del Bo, & Nijkamp 2011]. However, this does not directly reflect how humans actually perceive their environment and the city's services [Nam & Pardo 2011].

These physical sensors are interconnected into a computing framework to create a big picture of the live state of the city [Guthier, Alharthi, Abaalkhail, & El Saddik 2014]. This live state, however, only includes measurable quantities and disregards how the citizens actually feel. It is likely that there exist correlations between the emotional states of the citizens and relevant statistics like well-being of the city's inhabitants or quality of living [Lathia, Quercia, & Crowcroft 2012]. When urban planners use the collected data to optimize parts of the city, the emotional state of the inhabitants can thus serve as valuable implicit feedback. This gives rise to the vision of an emotion-aware city with the ability to understand and utilize the emotional states of its citizens to enable improved planning and decision making [Guthier et al. 2014], in which urban citizens can be called upon to act as active sensors, staging their personal spaces and sharing their spatiality [M. F. Goodchild 2007].

This research paper aims to present a conceptual framework in order to assess the human/emotional component within a smart city. This work is being executed for the PhD Program on Information Management at NOVA IMS (Universidade Nova de Lisboa), which started by the end of 2014.

Since for now the main progress of this PhD research is focused on the literature review, the figure below synthesizes the main keywords retrieved from 131 papers that were analyzed for now, which establish this multidisciplinary work among GISciences, Social Sciences, Smart Cities perception, Computational Sciences tools & methods and Social Media concepts.

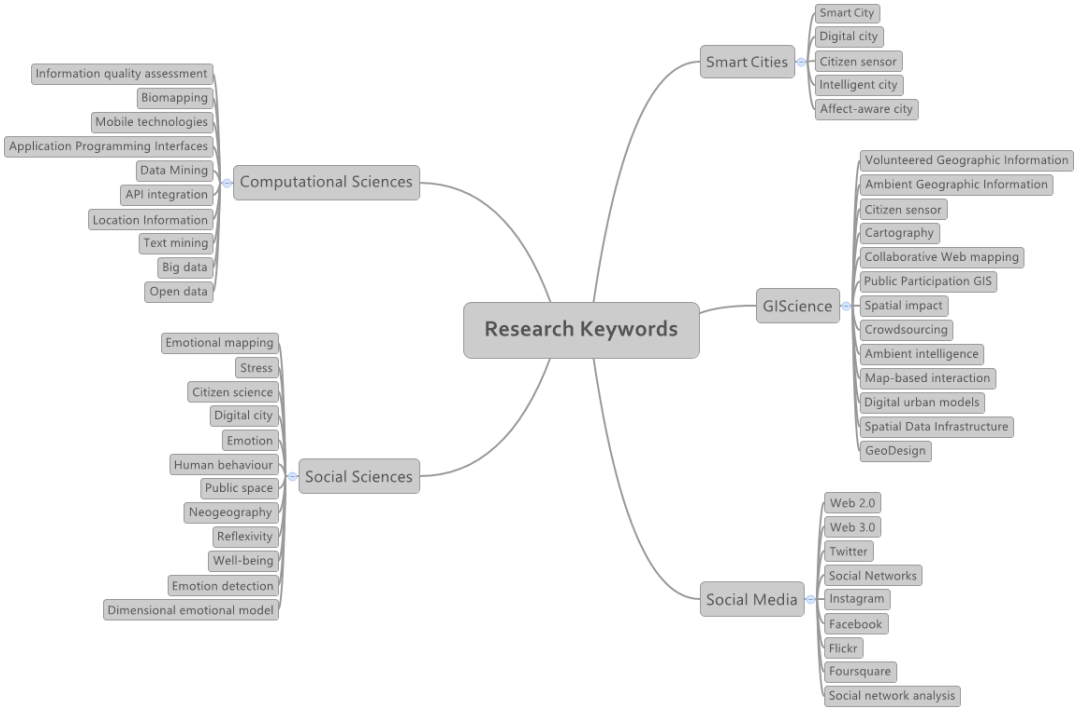


Figure 2 – Main research keywords retrieved from about 131 papers (Author)

This work will be executed in 5 Phases and 3 main deliverables (figure 3):

- Phase 1 – Literature review: in this task a literature review will be executed in order to assess the human/emotional component within a smart city, through some suggested sections, such as: definition of a smart city, more specifically its human component; the concept of volunteered geographic information (VGI) will be revisited with Ambient Geographic Information (AGI); motivations, issues and examples of emotion & stress mapping and its relevance through a smart city.
- Phase 2 – Public stream API programming: in this task, by using de social media public stream API's, a custom made code will be prepared in order to extract the spatial "emotional data" from social media.
- Phase 3 – Questionnaires: in this task several online questionnaires will be addressed, trying to capture the exact human dimension within smart cities.
- Phase 4 – Processing data and analyzing results: the results obtained will be analyzed, building the "emotional hotspots".
- Phase 5 – Objective measurements vs subjective observations: it consists mainly in comparing the results obtained through phase 2 and phase 3, analyzing the relevance of objective measurements vs subjective observations.

As for the main deliverables, they'll be represented at least by 3 Journal Papers, at least 4 Conference Papers, according to the phases on the gantt chart bellow.

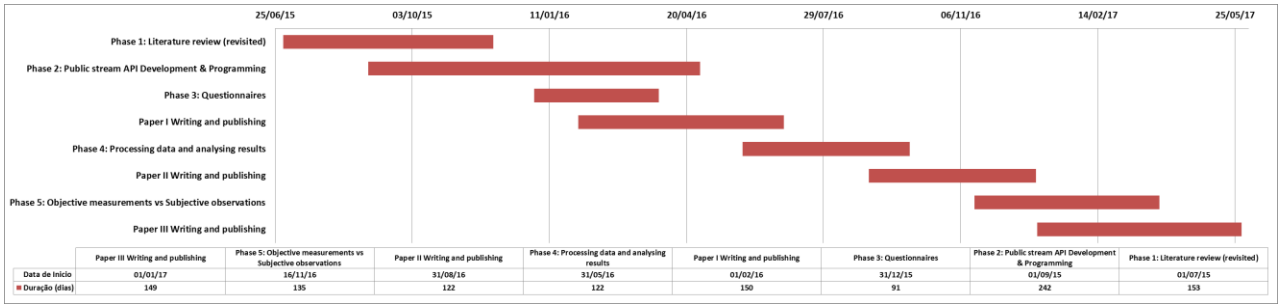


Figure 3 – Research Project Timeline (Author)

The purpose of this submission on the CAPSI Doctoral Consortium is to provide a setting where this work can collect valuable expert opinions and share new ideas, methods and applications.

2. Problem Definition

A smart city's main goal is to increase the quality of life for its citizens and to make the city more attractive, lively and greener [Kehoe & Nesbitt 2010]. To achieve this goal, physical sensors are deployed throughout the city to monitor various aspects such as environmental parameters (weather, pollution, etc.), traffic and the consumption of resources (Caragliu, Del Bo & Nijkamp, 2011). However, this does not directly reflect how humans actually perceive their environment and the city's services [Nam & Pardo 2011].

These physical sensors are interconnected into a computing framework to create a big picture of the live state of the city [Guthier et al. 2014]. This live state, however, only includes measurable quantities and disregards how the citizens actually feel. It is likely that there exist correlations between the emotional states of the citizens and relevant statistics like well-being of the city's inhabitants or quality of living [Lathia et al. 2012]. When urban planners use the collected data to optimize parts of the city, the emotional state of the inhabitants can thus serve as valuable implicit feedback. This gives rise to the vision of an emotion-aware city with the ability to understand and utilize the emotional states of its citizens to enable improved planning and decision making [Guthier et al. 2014], in which urban citizens can be called upon to act as active sensors, staging their personal spaces and sharing their spatiality [M. F. Goodchild 2007].

A smart city operates primarily in four dimensions: the intelligent city (its social infrastructure), the digital city (informational infrastructure), the open city (open governance) and the live city (a continuously adaptive urban living fabric) [S. Roche 2014]. In this context, today's actors are no longer mere consumers of urbanity, but true producers (or *produsers* [Coleman, Georgiadou, & Labonte 2009]) of these intelligent cities, since their level of spatial commitment is essential to the effective functioning of the smart city's four dimensions.

To complement this information, figure 4 identifies and clarifies the main conceptual variants and core factors of a smart city according to Nam and Pardo [Nam & Pardo 2011]: Technology (infrastructures of hardware and software); People (creativity, diversity, and education); and Institution (governance and policy).

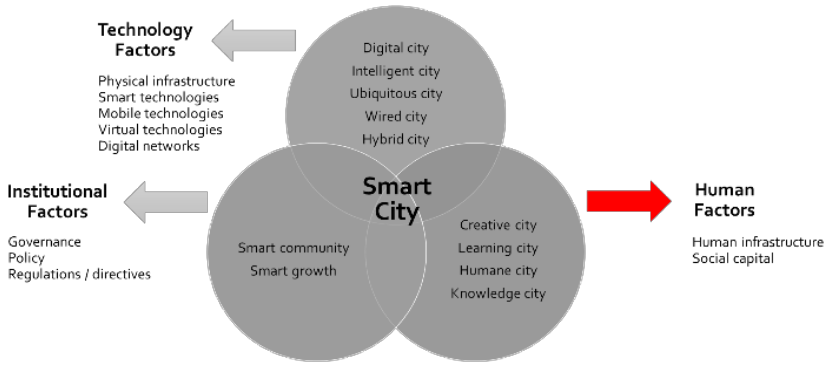


Figure 4 – Fundamental components of a Smart City (Adapted from Nam & Pardo, 2011)

The above mentioned "immaterial" component of the smart city is connected with the idea that a smart city is also a living urban fabric that is continuously being reshaped as is adaptive to change [S. Roche 2014].

In this context, there's a need to consider affect and emotion within a smart city, which are key elements towards rational decision making [Naqvi, Shiv, & Bechara 2006]. Since emotion is a central component of human behavior and, in order for a city to be truly "smart", it is important not only to assess *what* people are doing, but also, *why* they are behaving in a certain way [Dolan 2002]. Considering emotional states is essential for achieving real-time judgment and perceived life satisfaction [Guthier et al. 2014]. With this information, city planners can make use of the gathered affective data to detect positive or negative trends developing in the city, managing to take early countermeasures. Answering subjective questions such as "which part of the city is the best?", requires affect and emotion [Guthier et al. 2014].

But more important for a smart city is its capability to capture the sense of places. A city is not a machine, but rather made by people local actions and feelings. This could not be captured and represented without active citizens sensors (Volunteered Geographic Information [M. F. Goodchild 2007], Ambient Geographic Information [Stefanidis, Crooks, & Radzikowski 2013], crowdsourcing [M. F. Goodchild & Glennon 2010]) connected to location based-social networks [P. S. Roche 2012].

An active and engaged citizen is indeed the main driving force of a "smart city" [P. S. Roche 2012]. Nowadays we give a growing amount of location-based contents generated by connected – anytime and anywhere – *produsers*, mainly equipped with smartphones. The exponential growth of ambient geographic information through social networks (most of them are location-based social networks) became the basic feature of a spatially enabled society [P. S. Roche 2012]. Social networks are vessels which millions of people use to share their current thoughts, observations and opinions, and have been shown to provide more reliable and trustworthy information than traditional methods like questionnaires and other sources [Marwick & Boyd 2010].

These contributions raised through the ubiquity of mobile technologies, but accessible from different platforms, act as socio-spatial mediators [Lindqvist, Cranshaw, Wiese, Hong, & Zimmerman 2011]. Mobile technologies are definitely a valuable tool for collecting affective data in the context of an emotion-aware city [Liu et al. 2011], since they can simultaneously collect both spatial location and the user posts, which should contain emotion or mood content.

Social media generated from many individuals is playing a greater role in our daily lives and provides a unique opportunity to gain valuable insight on information flow and social networking within a society [Kar & Ghose 2010]. Through data collection and analysis of its content, it supports a greater mapping and understanding of the evolving human landscape [Stefanidis et al. 2013].

Such data conveys Ambient Geographic Information (AGI), capturing for example, people's references to locations that represent momentary social hotspots. Harvesting this ambient geospatial information provides a unique opportunity to gain valuable insight on information flow and social networking within a society, support a greater mapping, understand the human landscape and its evolution over time [Stefanidis et al. 2013].

This emergence of AGI represents a second step in the evolution of geospatial data availability, following on the heels of Volunteered Geographic Information (VGI) [Crooks, Croitoru, Stefanidis, & Radzikowski 2013]. According to Stefanidis [Stefanidis et al. 2013], harvesting and analyzing such ambient information represents a substantial challenge needing new skillsets as it resides at the intersection of disciplines like geography, computational social sciences, linguistics and computer science.

There are two major research challenges in the new field of emotion-aware cities, which are the *detection* and *aggregation* of affective data in an urban scenario [Guthier et al. 2014]. These issues are related with the specific and personal effect of appraisal and enthusiasm, which differs from person to person [Zhang & Lin 2011].

This research aims to develop a methodology towards emotion and stress mapping, in order to assess people's emotional responses to their environment as they walk through the streets of Lisbon (Portugal). Results could suggest that people's shared feelings about specific places are influenced by the particular physical properties and characteristics of a given place.

In this context, a methodology based on AGI and VGI, could produce alternative representations of space, based on individuals' georeferenced experiences, thoughts and emotions by mapping stress and emotion. The use of geospatial user-generated content, including social media information (*twitter, flickr, facebook, instagram*), could lead to a better urban planning, and how living in a urban area could relate to well-being.

3. The Hypothesis

The hypothesis for this research are as following: the first hypothesis rules that geospatial user-generated content (*subjective observations*) can truly characterize and share the *emotional dimension* of the city, in much as *objective measurements* – such as biometric sensors – or socio-demographic statistics about a specific city.

The second hypothesis states that Lisbon is a "rich" and worthy candidate to analyze as an emotion-aware city, simultaneously assessing its potential towards an (emerging) smart city.

Finally, the third hypothesis, discloses that there's a strong correlation between the main touristic sites in Lisbon and the "emotional landmarks" within the city, which can be defined as "emotional hot spots".

4. Methodological Approach

The rise of social media and the ability for analysis raises several concerns with respect to the suitability of traditional mapping and GIS solutions to handle this type of information [M. Goodchild 2010; Sui & Goodchild 2011]. We no longer map just buildings and infrastructure, but we can now map abstract concepts like the flow of information in a society, contextual information to place and linking both quantitative and qualitative analysis in human geography [Stefanidis et al. 2013].

In a sense one could consider AGI to be addressing the fact that the human social system is a constantly evolving complex organism where people's roles and activities are adapting to changing conditions, and affect events in space and time. By moving beyond simple mashups of

social media feeds to actual analysis of their content we gain valuable insight into this complex system [Stefanidis et al. 2013].

To implement an AGI-based methodology this research will focus on creating tools and methods that can collect, analyze and share information, based on geospatial user-generated content linked with social networks and media. Any spatial information related with emotion, enthusiasm, memories and stress related with the people's perception about the urban space, will be georeferenced. As a case-study, the city of Lisbon will be analyzed by using a text-mining approach in order to place in a map, any spatial *tag/post* related with emotion, either good or bad.

Figure 5 represents the methodological approach for this research. The emotional states of Lisbon citizens will be sensed through a variety of social media sources, by extracting features and applying machine learning techniques. Finally, this AGI-based data (*subjective observations*) can be compared with results from *objective measurements*, such as socio-demographic statistics, questionnaires and data retrieved from biometric sensors.

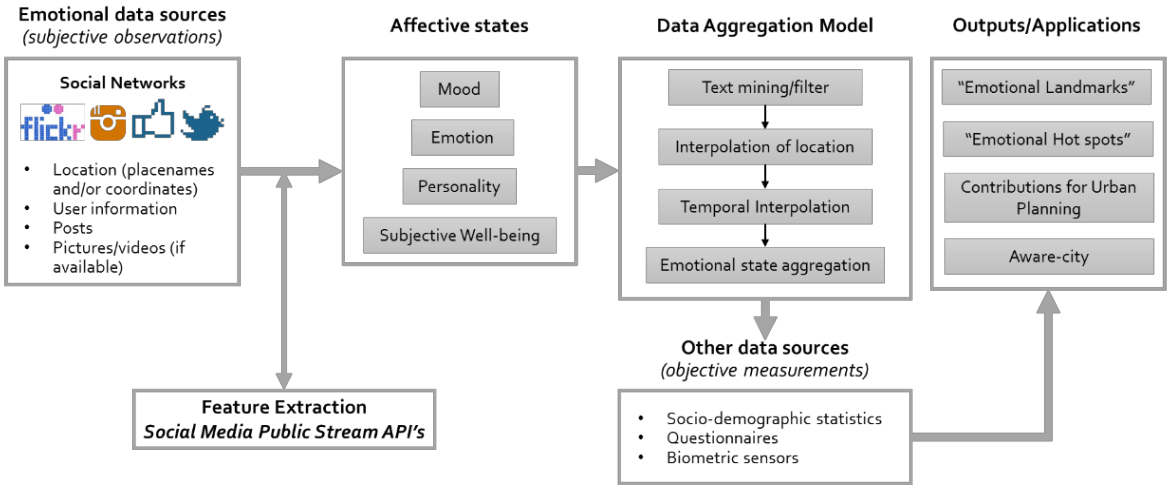


Figure 5 – Research conceptual framework (Author)

5. Expected Results

Mapping emotion builds on a tradition of studies in cognitive mapping, evaluative mapping, environmental preference and environmental affect [Weinreb 2013], adding an approach in which people experience, evaluate, and describe their environment *in situ* through social media.

With the emotional mapping of Lisbon it'll be possible to visually discern areas of strong feelings, either good or bad. These areas can be called as emotional clusters, which can be defined as the tendency shared by two or more participants in a particular place. The greater the number of people reporting a strong positive or negative feeling in the same location, the more pronounced the emotional cluster. These "hot spots", exhibit aggregations of positive ratings, negative ratings, or in some cases, a mixture of strong positive and negative ratings in the same place.

Other main goal will be to identify "emotional patterns" – those spaces where, at a specific or recurring time, a certain emotion is expressed powerfully and abundantly. If they exist: do emotional landmarks change over time? Do they change according to the observer? To language? To the time of day, week, month or year? Additionally, several groups of *produsers* will be established, based upon social-demographic characteristics, such as: gender; age; education level; motive of trip (leisure, business, e.g.); "level of acquaintance" of the place (*old-timer*, *newcomer*, tourist); origin (country and/or city).

These geospatial practices could highlight how emotions, subjectivities and spaces are mutually constitutive in particular places and at particular times, disclosing the socially encoded meanings of different kinds of bodies in specific spatial, temporal and cultural contexts. Results could suggest that people's shared feelings about specific places are influenced by the particular physical properties and characteristics of a given place, since this technique could be used to create and maintain spaces that are attractive, inviting and emotionally pleasing to a variety of users.

By engaging an emotion-aware city, new forms of communication can be engaged. Traditionally, the choice of partners for online group communication is either based on pre-existing relationships or based on similar interests or location [Ellison, Steinfield, & Lampe 2011]. In an emotion-aware city, communication groups can be formed spontaneously, based not only on a topic, but also on location and matching emotional state. These type of interactions can start interesting discussions about controversial projects or places within the city, since personal and sensitive issues are best shared with those who fell the same about a specific area of the city, creating participatory movements.

An emotion and stress mapping methodology could lead to understand, assess and evaluate the "immaterial" and emotional dimensions of the city and its spatial expression. GISciences can truly support the development of the *intelligent city* [Daniel & Doran 2013], due to crowdsourcing, VGI, AGI, including location-based social networks which stand out as key geospatial data sources indicative of the *pulse of the city*.

6. References

Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart Cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82. http://doi.org/10.1080/10630732.2011.601117

Coleman, D., Georgiadou, Y., & Labonte, J. (2009). Volunteered Geographic Information: the nature and motivation of produsers. *International Journal of Spatial Data Infrastructures Research*, *4*, 332–358. http://doi.org/10.2902/1725-0463.2009.04.art16

Crooks, A., Croitoru, A., Stefanidis, A., & Radzikowski, J. (2013). #Earthquake: Twitter as a Distributed Sensor System. *Transactions in GIS*, *17*(1), 124–147. http://doi.org/10.1111/j.1467-9671.2012.01359.x

Daniel, S., & Doran, M.-A. (2013). geoSmartCity: Geomatics Contribution to the Smart City. *Proceedings of the 14th Annual International Conference on Digital Government Research*, 65–71. http://doi.org/10.1145/2479724.2479738

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science (New York, N.Y.)*, *298*(5596), 1191–4. http://doi.org/10.1126/science.1076358

Ellison, N. B., Steinfield, C., & Lampe, C. (2011). Connection strategies: Social capital implications of Facebook-enabled communication practices. *New Media & Society*, *13*(6), 873–892. http://doi.org/10.1177/1461444810385389

Goodchild, M. (2010). Towards geodesign: Repurposing cartography and GIS? *Cartographic Perspectives*, (66), 55–69. Retrieved from http://www.cartoperspectives.org/carto/index.php/journal/article/view/93

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, *69*(4), 211–221.

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, *3*(3), 231–241. http://doi.org/10.1080/17538941003759255

Guthier, B., Alharthi, R., Abaalkhail, R., & El Saddik, A. (2014). Detection and Visualization of Emotions in an Affect-Aware City. *Proceedings of the 1st International Workshop on Emerging Multimedia Applications and Services for Smart Cities - EMASC '14*, 23–28. http://doi.org/10.1145/2661704.2661708

Kar, B., & Ghose, R. (2010). Is My Information Private? Geo-Privacy in the World of Social Media. *Ceur-Ws.org*, (Webster), 2009–2012. Retrieved from http://ceur-ws.org/Vol-1273/paper4.pdf

Kehoe, M., & Nesbitt, P. (2010). Smarter Cities Series : A Foundation for Understanding IBM Smarter Cities. *IBM Journal of Research and Development*.

Lathia, N., Quercia, D., & Crowcroft, J. (2012). The hidden image of the city: sensing community well-being from urban mobility. *Pervasive Computing*, 1–8. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-31205-2_6

Lindqvist, J., Cranshaw, J., Wiese, J., Hong, J., & Zimmerman, J. (2011). I'm the Mayor of My House:Examining Why People Use a Social-Driven Location Sharing Application. *ACM Conference on Human Factors in Computing*.

Liu, Y., Piyawongwisal, P., Handa, S., Yu, L., Xu, Y., & Samuel, A. (2011). Going Beyond Citizen Data Collection with Mapster: A Mobile+Cloud Real-Time Citizen Science Experiment. *2011 IEEE Seventh International Conference on E-Science Workshops*, 1–6. http://doi.org/10.1109/eScienceW.2011.23

Marwick, a. E., & Boyd, D. (2010). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133. http://doi.org/10.1177/1461444810365313

Nam, T., & Pardo, T. a. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. *Proceedings of the 12th Annual International Digital Government Research Conference on Digital Government Innovation in Challenging Times - Dg.o '11*, 282. http://doi.org/10.1145/2037556.2037602

Naqvi, N., Shiv, B., & Bechara, A. (2006). The role of emotion in decision making a cognitive neuroscience perspective. *Current Directions in Psychological Science*, *15*(5). Retrieved from http://cdp.sagepub.com/content/15/5/260.short

Roche, P. S. (2012). Sensing Places ' Life to make City Smarter. *ACM SIGKDD International Workshop on Urban Computing (UrbComp 2012) August 12, 2012 - Beijing, China*.

Roche, S. (2014). Geographic Information Science I Why does a smart city need to be spatially enabled? *Progress in Human Geography*, *38*(5), 0309132513517365–. http://doi.org/10.1177/0309132513517365

Stefanidis, A., Crooks, A., & Radzikowski, J. (2013). Harvesting ambient geospatial information from social media feeds. *GeoJournal*, *78*(2), 319–338. http://doi.org/10.1007/s10708-011-9438-2

Sui, D., & Goodchild, M. (2011). The convergence of GIS and social media: challenges for GIScience. *International Journal of Geographical Information Science*, *25*(11), 1737–1748. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/13658816.2011.604636

Weinreb, A. R. (2013). Mapping feeling: An approach to the study of emotional response to the built envirnoment and landscape. *Journal of Architectural and Planning Research*, 1–19.

Zhang, H., & Lin, S.-H. (2011). Affective appraisal of residents and visual elements in the neighborhood: A case study in an established suburban community. *Landscape and Urban Planning*, *101*(1), 11–21. http://doi.org/10.1016/j.landurbplan.2010.12.010

Um modelo de confiança para a gestão dos Sistemas de Informação no ambiente de Computação em Nuvem

Teófilo Branco Júnior[1], Henrique Manuel Dinis dos Santos[2]

1) Universidade do Minho – PDTSI – Programa Doutoral em Tecnologia e Sistemas de Informação Guimarães - Portugal

teofilotb@hotmail.com

2) Universidade do Minho – DSI – Departamento de Sistemas de Informação – Guimarães - Portugal
hsantos@dsi.uminho.pt

Resumo

Apesar das vantagens esperadas, a Computação em Nuvem (CN) ainda apresenta desafios em relação à desconfiança dos usuários em colocarem seus dados em computadores nos quais não possuem muitas formas de controle. Preocupações com a segurança no trato com os usuários, a falta de conformidade com normas e regulamentações, a ausência de mecanismos de controle acessíveis e a incerteza em relação ao sucesso em uma migração para um ambiente de nuvem são ainda alguns entraves para a adoção em larga escala desta tecnologia. Esta proposta de tese pretende investigar as normas e métricas que estão sendo propostas para regulamentar o uso dos SI no ambiente de CN, a fim de se estabelecer um modelo constituído por um conjunto de métricas integradas que representem as condições mínimas necessárias para tornar o ambiente de CN confiável e que possibilite uma avaliação deste ambiente pelas partes envolvidas.

Palavras chave: Computação em Nuvem, Sistemas de Informação, Segurança da Informação, Modelo de Confiança, Segurança na Nuvem.

1. Situação da Investigação

A presente investigação está inscrita no PDTSI – Programa Doutoral em Tecnologia e Sistemas de Informação da Universidade do Minho em Portugal. Os estudos afetos a esta investigação tem a previsao de duração de quatro anos e as atividades a serem desenvolvidas focam fenômenos organizacionais e sociais relacionados com a adoção e uso das Tecnologias de Informação (TI) emergentes como a Computação em Nuvem (CN).

Os trabalhos vinculados ao programa de doutoramento iniciaram-se em Outubro de 2014 e tem a previsão para sua conclusão em Outubro de 2018. O doutorando concluiu o primeiro ano que corresponde a componente curricular denominada *Curso de Doutoramento* que nclui a formação em metodologias de investigação em áreas específicas ao programa doutoral e outras áreas complementares, bem como seu plano de trabalho de tese de doutoramento a ser desenvolvido nos períodos seguintes.

Almejando uma participação no Doutoral Consortium (DC) da CAPSI 2015, o doutorando espera obter relevantes contribuições dos Srs. doutores com opiniões e críticas acerca de seu plano de trabalho e projeto de investigação, ofertando em contrapartida as reflexões acerca da temática de sua proposta.

2. Contexto e Motivação

A CN representa um modelo de serviço onde a premissa é a do fornecimento de todo o tipo de aplicações tecnológicas, através da disponibilização da infraestrutura e armazenamento de dados através da internet, pautado sobretudo na necessidade de consumo do usuário (Buyya, Broberg, and Goscinski 2011).

Na busca de maior eficiência no controle dos recursos computacionais, bem como agilidade e redução de custos, o ambiente da Computação em Nuvem pode proporcionar um melhor gerenciamento dos recursos de TI e da operacionalização dos Sistemas de Informação de maneira mais eficiente que a empregada tradicionalmente.

Entretanto, a CN apresenta desafios a serem superados em relação à desconfiança dos usuários em colocarem seus dados em computadores dos quais não possuem controle. A Internet se apresenta como um ambiente hostil e isso torna-se crítico quando se tem dados confidenciais trafegando entre os terminais e os servidores em nuvem (Mirashe and Kalyankar 2010).

A tecnologia de CN ainda causa dúvidas e desconfianças sobre a conveniência da adoção desta tecnologia, principalmente no tocante a segurança da informação. Embora promissora, esta tecnologia necessita ainda resolver diversos pontos de vulnerabilidade (Grobauer, Walloschek, and Stocker 2011)**.**

Identifica-se, portanto, a necessidade do aprofundamento dos estudos em novas investigações como o estabelecimento de padrões de conduta e de ferramental (métricas e formas de aferição) adequados para a operacionalização dos SI no ambiente de CN.

3. Fundamentos e trabalhos relacionados

Foi utilizado o quadro de referência bibliográfica para o desenvolvimento do processo preliminar de investigação, relacionado na Tabela 1, a partir de critérios padrões utilizados em um processo de revisão da literatura.

Referências	**Conceitos Chave envolvidos na Operacionalização de SI em CN**		
	Gestão do ambiente	**Segurança da Informação**	**Regulamentação e avaliação**
B. Grobauer et al, 2011			●
P. Mell and T. Grance, 2011	●		
R. Buyya et al, 2011	●		
C. Weinhardt et al, 2009	●		●
A. Lenk et al, 2009	●		
R. Buyya et al, 2009	●		
S. P. Mirashe and N. V Kalyankar,2010	●		
J. Che et al, 2011	●	●	
Y. Chen et al, 2010		●	●
M. Armbrust et al, 2009	●		●
M Kanchana et al, 2013		●	
M. Jensen et al, 2009		●	●

Tabela 1 – Matriz Referências x Conceitos-chave da revisão bibliográfica preliminar

Foram também pesquizadas normas e certificações atualizadas na área de Segurança da Informação. Manuais e recomendações relativas a Segurança da Informação no ambiente de Computação em Nuvem, foram consultadas através das publicações de órgãos normatizadores e certificadores.

Na revisão da literatura verificou-se que diversos pesquisadores, entidades e empresas tratam de assuntos relacionados a problemática da segurança e fiabilidade no ambiente de Computação em Nuvem, evidenciando que estudos ainda são necessários para tornar o ambiente tecnológico das CN mais robusto. A tecnologia de CN ainda causa dúvidas e desconfianças sobre a conveniência da adoção desta tecnologia, principalmente no tocante a segurança da informação. Embora promissora, esta tecnologia necessita ainda resolver diversos pontos de vulnerabilidade (Grobauer, Walloschek, and Stocker 2011).

4. Definição da hipótese ou do problema de investigação

Tanto para os executivos quanto para os técnicos da área de Tecnologia da Informação, torna-se um desafio descobrir como a Computação em Nuvem pode ser confiável e prover o nível de segurança necessária para a plena operação dos Sistemas de Informação. Neste sentido, há uma preocupação cada vez maior dos pesquisadores com a segurança e os aspectos relacionados com a confiabilidade do ambiente. O nível de preocupação varia desde conformidades com regulamentações até a questão da segurança no trato com os usuários finais. (Che, Duan, Zhang, & Fan, 2011)

A proposta desta investigação consiste no engajamento do pesquisador no sentido de verificar, validar e propor métricas e recomendações que contribuem efetivamente para tornar o ambiente de CN mais confiável através do estabelecimento de pârametros de verificação a serem aplicados em contratos de prestação de serviços neste ambiente.

Neste sentido, são levantadas as seguintes questões de pesquisa diretamente relacionadas com a investigação proposta:

i. É possível estabelecer parâmetros para tornar o ambiente de CN seguro o suficiente para a operacionalização dos SI?
ii. É possível formular um modelo composto por recomendações, métricas e cláusulas contratuais para que se possa verificar a consistência de um ambiente de CN para SI?
iii. Este modelo proposto, denominado pelo investigador como "modelo de confiança" pode ser automatizado?

O investigador acredita que tal modelo constituído por um conjunto de métricas integradas que representem os requisitos mínimos necessários para a operacionalização de sistemas de informação com segurança poderá ser útil para orientar os provedores, desenvolvedores de aplicações e clientes. Pretende-se ainda verificar se o mesmo poderá estar acessível as partes envolvidas de forma automatizada de forma que possibilite uma avaliação do ambiente de CN mútua e bilateral.

5. Objetivos, Metodologia e Métodos de Investigação

O objetivo geral da investigação é o da formulação de um modelo de gestão operacional constituído por um conjunto de métricas integradas que permitam monitorar o cumprimento dos requisitos minimos necessários para a operacionalização dos SI de forma segura e confiável no ambiente de CN, permitindo às partes envolvidas, a verificação deste ambiente.

Os objetivos específicos envolvem a demonstração do conhecimento dos estudos desenvolvidos neste âmbito, a implementação e aferição das métricas a serem utilizadas, a comprovação da sua utilidade destas e a aceitação das proposições produzidas pela comunidade científica.

Em relação a abordagem, trata-se de uma pesquisa com abordagem quantitativa e qualitativa. Quantitativa porque considera que pode-se tratar grande parte do problema de investigação a partir da coleta de dados em forma de números representados por medições de métricas as quais serão classificadas e analisadas. Qualitativa porque considera que existe uma uma relação entre os resultados numéricos obtidos e uma avaliação conceitual que não pode ser traduzida em números e deve ser analizada a partir de comparações de padrões qualitativos estabelecidos.

Trata-se de uma pesquisa aplicada, objetivando gerar conhecimentos para aplicação prática a ser implementada em um problema do mundo real. Do ponto de vista dos procedimentos técnicos, a pesquisa é experimental, pois a fase de produção de artefatos em que se determinam os objeto de estudo é definida a partir da análise de problemas reais, relatados e registrados através da divulgação de estudos de caso disponíveis na revisão da literatura e serão aplicados em um ambiente real para a verificação de sua ação.

5.1 A etapas da pesquisa

Para o desenvolvimento dos trabalhos, a investigação é constituída por 6 (seis) etapas que estão diretamente relacionadas com os objetivos desta investigação.

Etapa 1 - Caracterização do área de estudo - A fase inicial da investigação será apoiada na revisão da literatura, pois é necessário obter um conhecimento afeto a esta área para a identificação dos problemas e das soluções propostas por investigadores e entidades de estudo na área da CN. Este estudo tem como objetivo a obtenção de uma análise crítica a partir do ponto de vista dos investigadores que opinam sobre o assunto abordado. Entretanto, é pretendida ao longo de todo o processo de investigação, uma revisão da literatura constante sobre o estado da arte relativa ao problema de investigação.

Etapa 2 - Seleção e análise de proposições - Nesta etapa, serão desenvolvidas as atividades de análise de modelos, *frameworks*, ferramentas e técnicas propostas para a monitoração, segurança e auditoria de SI no ambiente de CN. Será conduzida também uma revisão sistemática rigorosa da literatura na busca da verificação das normas e recomendações propostas pelos investigadores. Serão também estudadas as publicações relativas a normatizações, recomendações e orientações das entidades normativas e grupos de estudo.

Etapa 3 - Elaboração de artefatos referentes a um protótipo de modelo inicial - Em um primeiro momento, a investigação tem como objetivo procurar selecionar as variáveis que influenciam a confiabilidade no ambiente de CN, definindo assim as formas de controle necessárias para que a segurança e estabilidade do ambiente de CN sejam viabilizadas. Métricas serão selecionadas, testadas e implementadas a partir da identificação de sua valia e eficácia. Será elaborado um protótipo do modelo de gestão desenvolvido a partir das recomendações e estudos de investigadores sobre o que interessa ser medido e controlado em um ambiente de CN. As métricas serão implementadas e colocadas no ambiente de CN a partir de um ponto de acesso para avaliação de seu comportamento.

Etapa 4 – Elaboração de um modelo de gestão operacional - Nesta etapa, será adotado um procedimento de estudo de caso com a variante da participação na experimentação dos artefatos em um em um campo de estudo prático onde já esteja implementada uma nuvem (um ambiente de CN) utilizando-se o método Action Design Research (ADR), no qual se previlegia a participação do grupo de trabalho nas deliberações e fases do estudo. Em uma fase inicial, o grupo de trabalho irá definir um modelo de gestão do ambiente de CN, a partir do protótipo inicial proposto e de sugestões de especialistas do próprio ambiente de trabalho. A seguir, serão selecionadas pelo grupo de trabalho, as recomendações e métricas passíveis de serem aplicadas para auxiliar a gestão do modelo de forma a poder avalia-lo. A avaliação será realizada pelo investigador conjuntamente com o grupo de trabalho e serão utilizadas como fonte referencial as

atas de reuniões do grupo, formulários de avaliação tipo *ServQual* adaptados, depoimentos pontuais e relatórios de experimentações.

Etapa 5 – Divulgação dos resultados alcançados – Esta etapa ocorre paralelamente as demais e tem por objetivo dar ciência e submeter a comunidade científica os resultados alcançados relativos a investigação em curso. Neste sentido, os registros das atividades da investigação em curso serão trabalhados e divulgados em forma de artigos para conferências, revistas e jornais relacionados com a temática da CN, para que possam ser partilhados, criticados e acolhidas sugestões relevantes e relacionadas com a investigação em curso.

Etapa 6 – Contribuição para o conhecimento - Nesta última etapa, estão previstas as atividades de consolidação do modelo proposto e da publicação da tese de doutoramento.

5.2 O Método Action Design Research (ADR)

A escolha do método ADR justifica-se pelo fato deste ser um método de pesquisa para a geração de conhecimento de design prescritiva através da construção e avaliação do conjunto de artefatos de TI em um ambiente organizacional aplicado, onde a questão de investigação e os resultados são obtidos através da participação e interação com os participantes.

As respostas exigidas resultam em um método que incide sobre a construção, intervenção e avaliação de um artefato que reflete não apenas os precursores teóricos e intenção dos pesquisadores, mas também a influência de usuários e uso contínuo no contexto (Sein et al. 2011), conforme demonstra uma ilustração do framework ADR, apresentada na Figura 1.

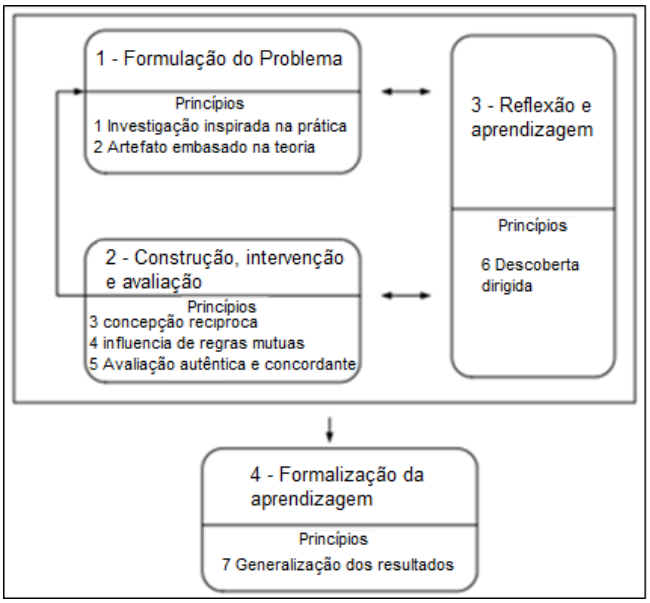


Figura 1 - Estágios e princípios do método ADR. Fonte: Adaptado de (Sein et al. 2011, pg 41)

A utilização de ADR tem o objetivo de aplicar, avaliar a construção do conhecimento na fase prática da investigação, em um estudo de caso, de forma a comprovar a aplicabilidade, eficácia e eficiência do modelo de gestão do ambiente de CN. Assim, pretende-se desenvolver o referido modelo operacional utilizando um grupo de trabalho inserido em um ambiente provedor de CN já estabelecido, através dos seguintes estágios:

Estágio 1: Formulação do Problema - No caso da investigação proposta, o fator motivador da primeira fase do *framework* ADR é um problema relacionado com os desafios em relação a

confiabilidade do ambiente de CN já antecipado pelo investigador. Desta forma, será fornecido ao grupo o problema a ser tratado e são estabelecidas as questões de motivação para a formulação do esforço de pesquisa. A entrada desta formulação vem das etapas da investigação, onde deverá ser analisado, criticado, avalizado ou complementado pelo grupo. Pretende-se neste estágio consolidar o problema da investigação. Uma vez identificado, articulado e definido o escopo, o problema serve como uma oportunidade para a criação do conhecimento.

Estágio 2: Construção, Intervenção e Avaliação - Em uma fase inicial, o grupo de trabalho irá definir um modelo de gestão do ambiente de CN, a partir do protótipo inicial proposto e de sugestões de especialistas do próprio ambiente de trabalho. O segundo estágio do ADR usa o problema e os pressupostos teóricos anteriormente levantados. Essas premissas fornecem uma plataforma para gerar o projeto inicial do artefato que será moldado no uso organizacional através dos ciclos de projeto subsequentes. Esta etapa é realizada como um processo interativo que entrelaça a construção do artefato de TI com a intervenção na organização. O resultado é a concepção realizada do artefato. O problema e o artefato são avaliados continuamente e os princípios de design são articulados para a consolidação e transformação do protótipo em modelo.

Estágio 3 – Reflexão e Estudo - A etapa de reflexão e aprendizagem envolve desde a construção de uma solução para o caso foco da investigação ao aplicar esse aprendizado a uma classe mais ampla de situações, buscando adequações e generalizações capazes de abranger uma grande aplicabilidade. Ela é contínua e paralela às duas primeiras etapas. A etapa reconhece que o processo de investigação envolve mais do que simplesmente a resolução de um problema. A reflexão consciente sobre a definição do problema, a escolha das teorias e o conjunto emergente é fundamental para garantir que as contribuições para o conhecimento sejam identificadas, a partir do envolvimento do grupo neste processo e não somente o olhar do investigador. Desta forma, pretende-se consolidar o modelo de gestão do ambinte de CN, objeto da investigação.

Estágio 4 - Formalização da Aprendizagem - O objetivo da quarta etapa do ADR é o de formalizar a aprendizagem. A aprendizagem situada a partir do projeto de ADR deve ser desenvolvido em conceitos gerais de uma solução, para toda uma classe de problemas de campo e não em relação a um caso em particular. Estes resultados podem ser caracterizados através de princípios de design justificados e com uma reflexão mais aprofundada, com o aperfeiçoamento das teorias que contribuíram para o projeto inicial.

6. Estado da Tese

Para condução dos trabalhos relativos a proposta da tese de doutoramento, são planeadas as seguintes ações a serem desenvolvidas constantes na Tabela 2.

Plano de Ação			
Ação	**Macro-atividades**	**Etapa da Investigação**	**Previsão**
Elaboração do Plano de Trabalho (Work Plan)	Obtenção de créditos curriculares no curso de doutoramento do PDTSI	Etapa 1	Setembro de 2014 a Julho de 2015
	Definição da investigação		
	Elaboração e submissão do Plano de Trabalho		
Elaboração do Projeto de Investigação	Validação dos conceitos teóricos	Etapa 2	Setembro de 2015 a Julho de 2016
	Obtenção de certificações referenciais no tema de estudo		
	Elaboração e submissão do Projeto de investigação	Etapa 3	
Desenvolvimento da investigação	Execução dos testes laboratoriais		
	Estudo de caso aplicado	Etapa 4	Setembro de 2016 a Julho de 2017
	Validação do objeto de investigação		
	Consolidação do objeto de investigação		

Participação em Conferências e Publicações	Revisão da documentação;	Etapa 5	Setembro de 2014 a Julho de 2018
	Submissão a conferências e workshops		
	Submissão a revistas e jornais		
Elaboração da Tese	Revisão da documentação	Etapa 6	Setembro de 2017 a Julho de 2018
	Elaboração e submissão da tese		
	Revisão e disponibilização a comunidade científica		

Tabela 2 - Plano de ação referente a tese de doutoramento

No processo de revisão bibliográfica preliminar, foram realizados estudos envolvendo a categorização de aspectos técnicos considerados importantes para a operacionalização dos SI em ambiente de CN e que constituem algumas das principais preocupações por parte dos investigadores como: gestão de banco de dados, gestão de usuários, bilhetagem e tarifação, segurança da informação, avaliação de desempenho e auditoria.

Estes estudos resultaram em um artigo submetido e aceito no 12º Congresso Internacional de Tecnologia e Sistemas de Informação (12CONTECSI), realizado no período de 20 a 22 de Maio de 2015 em São Paulo – Brasil e cuja apresentação se deu no dia 20 de Maio de 2015.

Um estudo mais aprofundado resultou ainda na submissão e aceitação de um outro artigo na 3rd International Conference on Cloud Security and Management (ICCSM2015) nos dias 22 e 23 de Outubro de 2015, na Universidade de Washington em Tacoma, Washington nos Estados Unidos.

O autor espera que a submissão de seu projeto a este Doctoral Consorcium da CAPSI 2015 possa vir trazer contributos para a sua melhoria e aperfeiçoamento deste projeto de investigação, tornando-o mais susceptível a uma futura aprovação de financiamento por alguma entidade ou órgão que promova a atividade científica, a partir de uma maior visibilidade e referência proporcionada pelo evento CAPSI 2015.

7. Conclusão e Contributos Esperados

Espera-se que o modelo proposto tenha êxito na presente investigação e contribua para a comunidade científica e indústria propiciando uma gestão melhor dos SI no ambiente de CN trazendo os seguintes contributos e benefícios:

i. Possibilitar a verificação dos recursos necessários a serem disponibilizados pelo provedor de serviços em CN para suportar Sistemas de Informações (SI) com qualidade;
ii. Viabilizar o direito de auditar os sistemas de SI, particularmente quando envolve um provedor para um serviço no qual o cliente necessita regulamentar o cumprimento das responsabilidades;
iii. Permitir a análise do escopo de conformidade, verificando se as regras de conformidade ao qual a organização está sujeita estão sendo impactados pelas práticas adotadas no ambiente de CN, tanto pelo prestador de Serviço quanto pelo consumidor;
iv. Permitir a análise da segurança das informações e dos dados. Potenciais usuários finais dos serviços de CN poderão ponderar quais aplicações e dados estão sendo considerados para serem movidos para serviços de CN e a que medidas estarão sujeitos para cumprimento dos regulamentos de conformidade;
v. Propiciar meios de orientação aos provedores e consumidores de soluções de SI em ambiente de CN no sentido de disporem de um ambiente mais estável e confiável;
vi. Contribuir com a comunidade científica e acadêmica através da divulgação dos resultados obtidos na investigação referentes a proposição de um modelo de gestão operacional, interagindo desta forma com a troca de experiências e de conhecimento científico;
vii. Auxiliar na formulação de novas hipóteses e teorias investigativas (estudos futuros) a partir do compartilhamento dos resultados obtidos no processo de investigação.

8. Referências

Che, J., Duan, Y., Zhang, T., & Fan, J. (2011). Study on the Security Models and Strategies of Cloud Computing. *Procedia Engineering*, *23*, 586–593. doi:10.1016/j.proeng.2011.11.2551

Chen, Y., Paxson, V., & Katz, R. (2010). *What's new about cloud computing security*. *... of California, Berkeley Report No. UCB/ ...*. Retrieved from http://www.utdallas.edu/~muratk/courses/cloud13s_files/what-is-new-in-cloud-security.pdf

Coutinho, C. P. (2008). A qualidade da investigação educativa de natureza qualitativa : questões relativas à fidelidade e validade Issues of validity and reliability in qualitative research, *12*(1), 5–15.

DeSanctis, G., & Poole, M. (1994). Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science*. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/orsc.5.2.121

Dhillon, G., & Backhouse, J. (2000). Information System Security Management in the New Millennium. *Communications of the ACM*, *43*(7), 125–128. doi:10.1145/341852.341877

Grobauer, B., Walloschek, T., & Stocker, E. (2011). Understanding Cloud Computing Vulnerabilities. *IEEE Security & Privacy Magazine*, *9*(2), 50–57. doi:10.1109/MSP.2010.115

Hevner, A., & Chatterjee, S. (2010). Design Research in Information Systems. In *Design Science in Information Systems* (Vol. 22, pp. 43–62). Springer US. doi:10.1007/978-1-4419-5653-8

Kaufman, L. M. (2009). Data Security in the World of Cloud Computing. *Ieee Security & Privacy*, *7*, 61–64 ST – Data Security in the World of Cloud Co. Retrieved from <Go to ISI>://WOS:000268639100011

Mell, P., & Grance, T. (2011). *The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology*. *National Institute of Standards and Technology*, *Information Technology Laboratory* (Vol. 145). doi:10.1136/emj.2010.096966

Mirashe, S. P., & Kalyankar, N. V. (2010). Cloud Computing. *Communications of the ACM*, *51*(7), 9. doi:10.1145/358438.349303

Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly*, *Vol 35 N 1*, pp 37–56. Retrieved from http://bada.hb.se/handle/2320/9888

ARTIGOS PROFISSIONAIS

9 TIPS ON HOW TO PITCH A PROJECT
The Collectors Bridge case study
CAPSI/2015

João Diogo Ramos[1]

1) Collectors Bridge, Portugal

jramos@collectorsbridge.com

Abstract

If you are an entrepreneur or an innovator and you have your work appraised by others, even more if you are in the field of Information Technologies, this article should be of interest and bring you a boost of motivation to your work and/or passion.

This article describes in detail the experience of the author on the preparation and successful pitching of a project called Collectors Bridge (www.collectorsbridge.com, Coimbra, PORTUGAL) that won the Best Business Plan Award on a National entrepreneurship competition named Arrisca-C (www.arrisca-c.pt, 2013 edition).

It starts by introducing the project and then describes the pitch performed, analysing the main decisions and their motivation. From there the author tries to summarize the main success factors an entrepreneur should be concerned about when on similar situations. In the end, based on the experience described, some of the intrinsic characteristics that project promoters should cultivate to enhance their success chances are highlighted.

In haste, if you have a great idea and you need to pitch it to take it forward, reading this paper should not do you any harm.

Keywords: entrepreneurship, pitching, competition, collect and communication

1. Personal and Organizational Context

I have been full-time involved in start-ups since at least 2005. I have started my career in a start-up. I have been an entrepreneur (whatever that really is… I think of it as someone who is resilient and drives projects forward in creative ways!) since I can remember. My first job after graduating in Informatics Engineering from Coimbra's University in 2001 was in a very innovative project at the time called GlobalShop [GS 2001]. GlobalShop was an MBA project that evolved into a Group of companies under the holding name of Global S, and it was funded by the industrial Portuguese Group Salvador Caetano [SC] together with Sonaecom and GE Capital. At GlobalShop I was involved in the development of some virtual stores (where people could go and buy the brand's products) and in my thesis that consisted on the development of a cross selling module. Such module was responsible for suggesting other products to buy based on other shopping basket's contents. Before Globalshop I had already been involved in several different "ventures". At the age of 20 I had already launched a computer store in our village

with two friends – one of them still continues that initiative. On top of that, being born in a small place, on several occasions during my adolescence I had to do additional work to do simple activities such as practicing Karate or Table Tennis – both of which I created from scratch: a karate dojo to have classes and a "semi-professional" table tennis team to compete.

After GlobalShop I joined Critical Software [CSW]. Critical is a company famous for having as its first customer the US Space Agency: NASA. Critical did a huge investment in quality as a differentiator from the very beginning (it is one of the few CMMI level 5 companies in Europe) and I consider it to be one of the best engineering schools I could have found. At Critical I performed development, consulting, project/product management and business development.

In 2005 I had a first meeting that would lead in 2008 to the creation of a spin-off called Critical Health (currently renamed to Retmarker, [CHT]) focused in Healthcare. Retmarker is mainly supported on a technology for Ophthalmology with the same name that the company develops and commercializes in the global market and that allows to reduce costs while addressing blindness prevention. At Retmarker I started as Business Unit Manager and in 2012 assumed the role of CEO of the company. Retmarker has been an amazing learning experience where I have confirmed that having ideas and transforming them into full products are two completely different things. At a personal level, it has been a huge career opportunity and allowed me to accomplish from scratch several other things: I became the principal inventor in a US patent; co-author in several publications and I was also the coordinator of the initial certification of our Quality Management System and Retmarker's CE marking as medical device.

But on top of all my activities, in 2013 I had another idea… I thought it was just another idea, but after a few days, awkward enough, I could not kill it! The more I thought of it, the more it made sense. The more I talked about it with friends, the more it got improved and the better the feedback I was getting. Eventually, I had one conversation with a senior colleague that had just concluded his MBA and the conversation ended with him asking "what do I need to do to be part of this project?". It was the moment I decided to start writing a business plan for what would become Collectors Bridge [C'B]!

2. Collectors Bridge origin (aka the Problem)

A few months before I had just gone through another project… This time it was the restoration of an 80's video-arcade machine. Sharing it with my friends through social networks, the impact and feedback was quite fascinating and I was impressed with the passion generated by a few hours of my free time. It was something that I had done several times before in my life - e.g. I also restored a 1957 Volkswagen Beetle.

On the other hand, as I learned to program a computer at young age (in the 80s) by spending hours and hours in front of a TV set using a Sinclair ZX Spectrum, I had decided to start a new collection as several years before in an interview I had claimed to be from the "Spectrum Generation". What I did not know was that this new collection of Sinclair and Timex clones (mostly made in Portugal!) would be the most important one in my life until now. Sharing this passion for the Spectrum computers online and in conversations with several others friends, I realized that I was really not alone. Having to spend hours and hours studying the tiny details to be able to track and buy the computers I wanted and spent even more hours negotiating and bidding for them, I thought: "there must be a better way!". I was eager to learn more and I saw people available to teach me! I observed people driven by emotion, originated in a glimpse of a childhood object, to come forward and engage in passionate conversations.

Discussing it with other colleagues I realized there was a business opportunity. I identified the competences I needed to have on the team, invited the best individuals I knew and welcomed them on-board. We noticed we had access to important passionate collectors in our network. We were also able to set up a compelling value proposition to all the players involved in the collecting arena with no need to ask for any upfront investment. We started looking for ways to finance the company and looked at entrepreneurial competitions that would provide awards in form of money that could be used freely in the development of the company according to our strategy. We clearly avoided awards that would provide only consultancy, market research or similar services, as that was not our target. After a lot of consideration, we identified Arrisca-C [ARRISCA-C 2013] and thought it would be our best (and only) bet: all-in! We applied for the competition and chose the category that would provide the best award and the one that could be better justified: "Business Plan competition". As part of the evaluation process, we were called to do a pitch were we had 10 minutes to convince the Jury to choose us.

3. Pitching session (aka the Solution)

The pitching session was, as expected, an exciting moment. Being emotionally driven as I am, the minutes before it were very tense, but as soon as I started talking, things were all ok. The stage and spotlight was mine and it felt pretty good. Bear in mind that this was our plan A, B or even C. Our capacity to invest on the project depended on being successful on these 10 minutes! I had prepared myself the best I could. As I had learned at Coimbra's University several years before, I spent the night before rehearsing the presentation. It had been years since I had done it with this intensity and relevance but even the improvise needed to be trained.

In advance to the presentation we prepared several initiatives to surprise and show our commitment to the Jury: a) we started by updating our LinkedIn profiles to include Collectors

Bridge; b) we registered all internet domains relevant for the project; c) we launched a landing page announcing the project and gathering e-mails from people interested; d) we did a survey with collectors to validate our assumptions; e) we had a friend designing our logo; f) we created a proof-of-concept, intentionally copied from Facebook, as to illustrate a simple to use platform, to have mock-ups to use in the presentation and to validate a possible technical architecture.

Regarding the presentation, we decided to structure it raising the main questions we thought the Jury wanted to clarify and then answering them with data. Those questions and their rationale were: 1) *"How we identified the opportunity?"* With this question we intended to share our dream, highlight the passion for the project and to create an excellent first impression; 2) *"Why is there an opportunity?"* With this question we wanted to make sure they agreed there is an opportunity, so there is a problem, affecting lots of people and there is not a good solution to address it; 3) *"People make companies... Who is the promoting team?"* With this question we wanted to highlight whom we were and do it early in the presentation. We wanted to stress the asset of having an abnormally mature team like ours in such ambitious project; 4) *"What is the C'B and what have we been doing?"* With this question we wanted to show all the work we had done prior to the pitching session and thus explain how we were validating our assumptions (eg. by surveys, landing page, etc.), we explained what was the target market and the data supporting it, we summarized the business model and introduced several financial scenarios showing that there was break-even even in our most conservative scenario; 5) *"Where will the award money be applied?"* With this question we wanted to show the Jury how much this was important to us (even independently of the money) and what we would do with if they chose us; 6) *"Why choose C'B?"* This question was basically a final wrap-up and an opportunity to tease the Jury about choosing us (more on that later!).

The first thing I needed to accomplish in the presentation was to create a first good impression, as there is only one chance… I wanted to surprise the Jury, make them remember us! To try to accomplish this, the first thing I ensured was that we had the entire founding team in the room to show our full commitment to the project. On top of all, my colleagues were wearing a Collectors Bridge t-shirt created specifically to this purpose. I was wearing a t-shirt about the Sinclair computers that I collect. I started the presentation by telling the story of a little boy that grew up to become an engineer and how during that process was highly affected by a device called ZX Spectrum that taught him lots of things. Obviously it was a real story and my personal one as this was indeed in the origins of the project. By using a children's photo I was appealing to their human emotions as we were all a children with toys once. I explained that I had started collecting the computers and than I could not find a good way to share my passion and learn from others in a proper way. After explaining all of this, I removed my t-shirt and I

had a Collectors Bridge t-shirt like the one my colleagues were wearing and I said "*now is the moment I step out of my collector role and where I shift to my business role*". By doing something visual, something simple - but that they were not expecting - I caught their attention for the rest of the presentation. The remainder of the pitch went as expected and although I had something like 20 slides (which normally is absolutely not recommended for a 10 minutes presentation!), they were prepared in such a way that I was able to conclude it within time.

After the presentation came the question & answers on which we were able to create a good interaction with the Jury and clarify their doubts. Bear in mind that this is a project about passionate subjects and everyone collects or collected something throughout their lives which makes easier to establish emotional bonds with other humans as it eventually happened. Another innovative thing we did (based on our experience in similar activities) was to prepare a one-pager with a kind of appraisal matrix that we filled in with what we believed were the positive things about our project for each criterion. The rationale was to facilitate the Jury's work and naturally to try to persuade them by presenting in advance our arguments. On top of this all, the first question the Jury did was the one that we had anticipated and that was related to our commitment to the project. When the question was raised I used the slide I had prepared!

4. Results

The result was very positive: we created a good impression, we clarified the doubts and a few months later in a public ceremony we received the award. Immediately after, we formalized the inception of the company. The company is alive and growing at a steady pace and our current focus is on boosting some communities and evolving the platform as needed.

5. Final Thoughts

Trying to summarize the rationale for the decisions we took in order to help others, I will now present the nine main tips that I believe entrepreneurs should focus on when on similar situations. All of these are things that can be trained and you should focus on them to reach similar objectives. 1) *"Be passionate"*: speak with passion, as if you do not believe in it, no one will!; 2)*"Be prepared"*: rehearse the presentation in front of a mirror and/or friends until exhaustion… and then start over again! Improvise needs to be trained; 3)*"Be creative and innovate"*: surprise your audience, e.g. something visual as changing a t-shirt in the middle of the presentation. Be prepared to make an amazing first impression; 4)*"Bring data"*: identify your strengths and weakness and elaborate on them. Support your conclusions and hypothesis and try to show the advantages, e.g. showing that you are still profitable in your worst-case scenario; 5)*"Demonstrate knowledge"*: do it in subtle ways, e.g. I studied Museology to

understand the roots of collecting and to be able to discuss it with experts; 6)*"Anticipate Doubts"*: if you are able to anticipate the Jury's concerns you will surprise them – e.g. we had the luck of achieving that with the question about our commitment to the project. Try to be one step ahead, always!; 7)*"Facilitate other people's work"*: e.g. by looking out the evaluation criteria and filling it in with what you want to stress as in any self-appraisal; 8)*"Reduce liability"*: by saying what you will do with the award you reduce the concern on choosing you. In our case we had decided to hire people (for the daily operation and to develop the platform) which seemed a good investment and one easily justifiable; 9)*"Tease and provoke others"*: you need to be confident and believe in yourself and in your team. Afterwards you can provoke others to make them think they should bet on you as we did in our final message showing we were willing to risk and asking the Jury if they had the courage to do it also.

Being an entrepreneur is not for everyone. You have to live with uncertainty, with risk, with the fear of failing… constantly! It is indeed a tough way of living. But when you accomplish something, even small things, the satisfaction from it compensates it all in my personal opinion. I believe that anyone can be an entrepreneur and I do not believe that there is a set of specific things you need to have or demonstrate to become one. What I think is that if you have a certain set of skills it may become easier to reach certain objectives. Nevertheless, all can be taught and trained. I believe that to increase your chances of success you should a) "*try to excel in terms of communication and networking*": go out and talk to people that may help you but never make them lose their time. Find ways to get their attention and seek for their advice. Do not be shy as the "no" is granted and networking is what drives the World; b) "*stay self-confident and resilient*": it is a tough but satisfying way of living; c) "*Be flexible*" as you will have to multi-task and do everything in a start-up; and, most importantly, d) "*open your eyes and be eager to learn*" as life is made out of learning and that helps to identify new opportunities.

6. References

GS, GlobalShop, web.archive.org/web/20010609091549/http://www.globalshop.pt/, 2001

SC, Salvador Caetano Group, www.gruposalvadorcaetano.pt, Porto, PORTUGAL

CSW, Critical Software, www.criticalsoftware.com, Coimbra, PORTUGAL

CHT, Retmarker / Critical Health, www.retmarker.com, Coimbra, PORTUGAL

C'B, Collectors Bridge, www.collectorsbridge.com, Coimbra, PORTUGAL

ARRISCA-C, www.arrisca-c.pt/ ; web.archive.org/web/20141025033143/http://www.uc.pt/gats/eventos_e_iniciativas/a_decorrer/Arrisca_C_2013, 2013

The author would like to thank all the friends, individuals and organizations that have been supportive of the project. Also send a very warm thank you for my fellow Collectors Bridge's team members. And, last but not least, our USERS, the reason we created the platform: thank you!

Angariação e retenção de clientes numa maternidade privada

Maria Rodrigues [1], Guilherme Victorino [2], Roberto Henriques [3].

.

1) Nova IMS, Portugal
mrodrigues@novaims.unl.pt

2) Nova IMS, Portugal
gmvictorino@novaims.unl.pt

3) Nova IMS, Portugal
roberto@novaims.unl.pt

Resumo

Este trabalho apresenta o caso de um prestador de saúde privado, com maternidade, da zona da grande Lisboa, cujo número de partos tem vindo a decrescer. Trabalhou-se um conjunto de dados de utentes da especialidade de Ginecologia/Obstetrícia (GIN/OBS), a partir do qual se construiu uma metodologia de análise inovadora na aplicação de Customer Relationship Management (CRM) a esta especialidade médica. A criação de perfis de utente, através da construção de métricas agregadas, permitiu aferir condicionantes do negócio, como a utilização de Entidades Financiadoras de Referência (EFR's) e o desempenho de médicos em número de partos, a georreferenciação de utentes, e a segmentação de clientes por valor. Este conhecimento, em conjunto com dados da literatura e da análise do mercado das maternidades privadas, permitiu definir diretrizes de atuação de marketing que podem ser aplicáveis a vários níveis da organização, visando o aumento da quota de mercado de partos do prestador.

Palavras chave: CRM na Saúde, Data Mining, Ginecologia/Obstetrícia, Segmentação, RFM.

1. Introdução

Este trabalho apresenta o caso de um prestador de saúde privado com maternidade da zona da grande Lisboa que pretende caracterizar as suas clientes de Ginecologia/Obstetrícia para melhor poder orientar a sua estratégia de marketing de angariação e retenção de clientes na sua maternidade. Numa altura em que o contexto da natalidade em Portugal é um dos assuntos mais preocupantes a nível nacional, devido ao decréscimo contínuo deste índice desde a década de 60 e à queda cada vez mais acentuada na última década, urge ao sector hospitalar rentabilizar a capacidade instalada das suas maternidades para lutar num mercado de potencial cada vez menor e cada vez mais concorrencial. Este prestador é composto por várias clínicas e um hospital com maternidade, e por motivo de confidencialidade será doravante designado por Prestador A.

2. Metodologia

O primeiro passo desta metodologia compreendeu a caracterização dos dados disponibilizados, o seu tratamento com vista à obtenção de um conjunto de dados consistente. Num segundo passo foi realizada a construção da base de dados dimensional, introduzindo variáveis de identificação das utentes grávidas e de caracterização do seu comportamento de compra. A identificação das utentes grávidas é feita pela identificação de atos realizados que são exclusivos da Obstetrícia, como apresentado na Tabela 1, e de acordo com a metodologia apresentada por Isken & Rajagopalan (2002).

"Acomp. gravidez"	"Parto normal"	"Cesariana"
Amniocentese (2º trimestre)	Parto distocico	Cesariana c/1 RN
Amniocentese (3º trimestre)	Parto distocico s/anest. c/1 RN	Cesariana c/2 RN
Consulta de Obstetricia	Parto normal	Cesariana
Consulta Diagnostico Pre-Natal	Parto gemelar normal por c/gemeo	
Consulta materno fetal	Parto eutocico c/anestesia c/1RN	
Eco obstetrica		
Eco obstetrica - 1º Trimestre		
Eco obstetrica - 2º Trimestre		
Eco obstetrica - 3º Trimestre		
Monitorizacao fetal		
Monitorizacao fetal externa		
N.s.t/ecg fetal		

Tabela 1: Tipos de atos inequívocos de gravidez

O comportamento de compra relativo à gravidez é caracterizado pela criação de quatro tipos de utentes: Não grávidas, "SóParto" (utente grávida que só realizou parto), "SóAcompanhamento" (utente grávida que só realizou acompanhamento de gravidez), "Ambos" (utente grávida que realizou acompanhamento e parto). Esta base de dados serve o propósito de criação de uma ABT (*Analytical Base Table*) para realização da análise de segmentação por *data mining*, permite a realização de consultas ad-hoc à base de dados e a extracção de dados para a análise de georreferenciação das utentes grávidas.

Num terceiro passo, realizaram-se consultas ad-hoc à base de dados construída, com base em questões de negócio específicas da especialidade de GIN/OBS, e que compõem a análise exploratória dos dados. Num quarto passo foi realizada a análise de tipos de utentes grávidas por área geográfica de residência, usando ferramentas de georreferenciação. Com esta análise pretendeu-se georreferenciar as moradas das clientes para visualmente identificar as zonas de maior densidade em mapas de choropleto. Utilizou-se o *software* ArcGIS, versão 10.2.1 da ESRI, e a plataforma Google Earth. As moradas são caracterizadas pelos códigos postais fornecidos nos dados originais. Para efetuar uma análise mais pormenorizada representou-se a localização das clínicas e do hospital em estudo no Google Earth. Uma vez que o Google Earth

contém informação relativa aos nomes das freguesias é possível analisar esta informação com uma perspetiva de pormenor personalizada.

Por último, efectuou-se a segmentação de clientes por valor através de técnicas de *data mining* [Berry e Linoff 2004]. Esta metodologia foi baseada nos trabalhos de [Isken e Rajagopalan 2002], [Prather et al. 1997] e [Goodwin e Maher 2000], que aplicaram técnicas de *data mining* a dados clínicos de Obstetrícia. A metodologia de segmentação por valor realizada baseia-se em aplicações de clustering a clientes na área da saúde com aplicação do algoritmo k-means a variáveis RFM (*Recency*, *Frequency*, *Monetary)*, em [Wu, Lin, & Liu 2014], [Chen, Cheng, Lai, Hsu, & Syu 2012]. Também noutras indústrias esta metodologia foi aplicada [Cheng & Chen 2009], [Hosseini, Maleki, & Gholamian 2010], tendo-se revelado eficaz na determinação da lealdade dos clientes. Neste caso a integração entre RFM e clustering pretende fornecer inteligência de marketing para ajudar no desenvolvimento de estratégias mais eficazes [Birant 2011]. As variáveis RFM para cada cliente foram calculadas com base na identificação de episódios, em que um episódio é o conjunto de atos realizados numa data concreta Juntamente com R,F e M utilizaram-se as variáveis TipodeUtente e FaixaEtária, criadas para este propósito a partir de variáveis originais. Com estas 5 variáveis criou-se uma ABT (*analytical base table*) que foi carregada em Enterprise Miner 12.3. da SAS, onde se realizou a segmentação.

3. Resultados

3.1. Análise de Entidades Financiadoras de Referência

No que diz respeito às EFR's, a maioria das clientes usufrui de apenas uma EFR na especialidade GIN/OBS, e esta encontra-se relacionada com o tipo de utente. O tipo SóParto utiliza maioritariamente Seguros, por oposição ao tipo Ambos que utiliza maioritariamente o tipo A. O tipo SóAcompanhamento é maioritariamente do tipo A, tendo-se identificado que as utentes abandonantes possuem EFR maioritariamente OutrosUtentes e Particulares.

3.2. Desempenho de médicos

No que diz respeito à análise do desempenho de médicos, o aumento do número de médicos a realizar partos em 2013, mais 24% que em 2012, não teve um aporte positivo no número de partos anual, que desceu cerca de 12%, podendo-se concluir que ter médicos mais dedicados ou mesmo em exclusividade e a realizar maior número de partos poderá ser importante na fidelização de clientes.

Uma vez que uma das principais motivações para escolha de uma maternidade é, para o setor privado, a garantia de ser o obstetra que acompanhou a gravidez a realizar o parto [DECO, 2014], recomenda-se envolver o médico, respeitando as determinações deontológicas médicas na cultura da organização e no esforço e nas ações de marketing com vista ao fortalecimento da relação emocional do cliente ao médico e à organização [Kotler, Shalowitz, & Stevens 2008]; usar como indicador de desempenho os perfis de médico por tipo de utente, incentivando o número de utentes do tipo Ambos; analisar o tipo de vínculo do médico com o prestador, se trabalha a tempo total ou parcial, e outros locais onde trabalha, relativamente à taxa de abandono de utentes que são seguidas no prestador mas que optam por outro local para ter o parto.

3.4. Georreferrenciação

Com esta análise determinou-se que a morada da utente ser perto de uma clínica do prestador pode ter relação com a realização de parto, pelo que as clínicas podem ter um papel fundamental na comunicação com as suas utentes com vista à sua retenção, e também na angariação de novas clientes. Por esse motivo recomenda-se apostar na comunicação dos serviços de Obstetrícia nas clínicas, e em especial às utentes de Ginecologia, com foco na recente renovação da maternidade, uma vez que os factores de escolha de um prestador são a reputação e imagem do hospital, a publicidade e informação disponível sobre o hospital Yavas e Sahmwell [citados em Fernandes 2010]; e apostar em divulgação local e participação em atividades/eventos promovidos pelos concelhos/freguesias próximas das suas clínicas de forma a reforçar a ligação com as mulheres das comunidades onde está inserido o prestador [Dearing 1987], por forma a captar novas clientes.

3.5. Segmentação de clientes

A segmentação de clientes por valor permitiu identificar 3 clusters com características distintas. A contribuição desta análise é perceber que não há clusters de valor elevado a não ser o Cluster 2, que engloba as utentes que realizaram parto na instituição, e que as senhoras que realizam SóAcompanhamento, no Cluster 3, têm na sua maioria recências baixas e frequência altas, o que indica que estão a fazer acompanhamento continuado

Para o cluster 1, senhoras com mais de 50 anos e com baixo valor, recomenda-se envolver a família para se tornarem clientes do prestador, e desenvolver a relação noutras especialidades, uma vez que serão à partida clientes com maiores necessidades médicas de curto e médio prazo, à medida que a idade avança [Dearing 1987]. Para o cluster 2, as clientes que tiveram parto na

instituição, do tipo Ambos e SóParto, com valor alto, recomenda-se como primeira abordagem, canalizar para a especialidade de Pediatria, e no futuro, para a realização do segundo parto no prestador. Estas são as clientes com grande envolvimento emocional com o prestador, e que podem ser os drivers de trazer toda a família e referenciar o prestador a amigos. São estas as clientes mais importantes. Para o cluster 3, composto pelas clientes SóAcompanhamento, com recências baixas e frequência altas, a recomendação é a mesma que para o cluster 2. e que provavelmente irão realizar o parto no prestador

4. Conclusão

Este estudo pretendeu estabelecer, por um lado, uma metodologia de análise de clientes de GIN/OBS que fornecesse, por um lado, conhecimento acionável em termos de marketing para estabelecimento de recomendações ao nível do CRM e de alterações organizacionais, que, com suporte em literatura, pudessem ser usadas para reter clientes atuais e para encontrar formas de angariar novos clientes. Por outro lado, estabelecer uma análise abrangente de dados da especialidade de GIN/OBS centrados no cliente, evidenciando a importância das análises de dados organizados numa perspetiva de marketing, com vista à compreensão do comportamento do cliente. Esta análise foi efetuada com recurso à criação de tipos de utente que caracterizam a realização de acompanhamento de gravidez e/ou de parto no prestador, e que permitem determinar se as variáveis em estudo estão relacionadas com o comportamento das utentes.

Conclui-se que os dois objetivos foram conseguidos; a metodologia utilizada permitiu fundamentar as recomendações baseadas em literatura e as análises realizadas aportaram conhecimento sobre o comportamento de compra dos clientes. Quanto à definição de tipos de utentes, verifica-se que apoia a compreensão mais profunda do comportamento relativamente ao acompanhamento de gravidez e realização de parto no prestador, permitindo criar perfis de análise que indicam se as variáveis em estudo são ou não relevantes para a utente realizar só acompanhamento de gravidez, só parto, ou ambos, no prestador.

As principais dificuldades de concretização do projeto relacionam-se com a transformação dos dados organizados numa ótica de gestão para uma ótica de marketing, centrada no cliente. O conhecimento da especialidade clínica e das particularidades dos dados fornecidos foi fundamental na extração de conhecimento útil e que está de acordo com a realidade do negócio. Recomenda-se em projectos onde se verifique esta condição consultar especialistas clínicos e realizar entrevistas no prestador por forma a interpretar corretamente os dados fornecidos uma vez que a construção da base de dados dimensional é o passo chave para a análise dos dados.

5. Referências

Birant, D. (2011). Data Mining using RFM Analysis. Knowledge-Oriented Applications in Data Mining.

Chen, Y. S., Cheng, C. H., Lai, C. J., Hsu, C. Y., & Syu, H. J. (2012). Identifying patients in target customer segments using a two-stage clustering-classification approach: A hospital-based assessment. [Article]. Computers in Biology and Medicine, 42(2), 213-221.

Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. [Article]. Expert Systems with Applications, 36(3), 4176-4184.

Dearing, R. H. (1987). Marketing to attract women. Health progress (Saint Louis, Mo.), 68(1), 26-28.

DECO (2014). Preços de partos: nascer em Lisboa custa mais. http://www.deco.proteste.pt/saude/hospitais-servicos-saude/noticia/precos-de-partos-nascer-em-lisboa-custa-mais

Fernandes, T. S. (2010). As perceções dos utentes dos hospitais públicos e privados portugueses: uma perspetiva comparativa (Projeto de mestrado não plublicado). ISCTE Business School, Lisboa.

Goodwin, L., & Maher, S. (2000). Data mining for preterm birth prediction Paper presented at the Proceedings of the 2000 ACM symposium on Applied computing

Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. Expert Systems with Applications, 37(7), 5259-5264.

Isken, M. W., & Rajagopalan, B. (2002). Data mining to support simulation modeling of patient flow in hospitals. [Article]. Journal of Medical Systems, 26(2), 179-197.

Kotler, P., Shalowitz, J., & Stevens, R. J. (2008). Strategic Marketing for Healthcare Organizations. In I. John Wiley & Sons (Eds.)

Prather, J. C., Lobach, D. F., Goodwin, L. K., Hales, J. W., Hage, M. L., & Hammond, W. E. (1997). Medical data mining: knowledge discovery in a clinical data warehouse. Proceedings : a conference of the American Medical Informatics Association / ... AMIA Annual Fall Symposium. AMIA Fall Symposium, 101-105.

Wu, H. H., Lin, S. Y., & Liu, C. W. (2014). Analyzing Patients' Values by Applying Cluster Analysis and LRFM Model in a Pediatric Dental Clinic in Taiwan. Scientific World Journal, 7.

Qual o Impacto das aplicações em "Software as a Service" (SaaS) no negócio das PME ?

CAPSI 2015 - Fórum de Profissionais de Sistemas de Informação

Jorge Carrola Rodrigues

NOVA IMS, Information Management School, Universidade Nova de Lisboa, Campus de Campolide, 1070-312 Lisboa, Portugal

jcarrola@novaims.unl.pt

Resumo

Neste artigo abordam-se fatores de adopção de soluções empresariais desde a época dos "Application Service Providers" (ASP) e que diferenças existem para a atualidade do "Software as a Service" (SaaS). São apresentadas as óticas de Fornecedores e de Clientes, nomeadamente com exemplos de Casos de Estudo realizados e como os Fatores que são considerados mais importantes por parte dos Profissionais de Tecnologias de Informação (TI) nem sempre têm a mesma valorização por parte dos Decisores de Negócio das empresas. Outros aspetos que são também abordados é a tendência crescente para os modelos Híbridos (mix de sistemas em "On-premise" e em SaaS) e como estão a ser impactados os modelos de negócio, quer dos fornecedores, quer das organizações clientes, e a transformação organizacional, de processos e de gestão que passa da gestão de "assets" para a gestão e contratualização de serviços ou um misto dos dois. Finalmente discute-se o que está a ser feito para medir o impacto do uso e o valor resultante, na maior competitividade e performance empresarial das pequenas e médias empresas (PMEs), com aplicações em SaaS.

Palavras chave: "Software as a Service", SaaS, PME, fatores de adopção, valor, performance empresarial

1. ASPs e SaaS - semelhanças e diferenças nos fatores de adopção

Considerou-se útil remontar até há quinze anos atrás, aos primórdios do que é hoje a disponibilização de aplicações informáticas como serviço, que na altura se designavam de "Application Service Providers" (ASP). No final dos anos 90 os ASP surgem como a grande promessa para o acesso a aplicações, por parte das Pequenas e Médias Empresas (PME), que usualmente apenas estavam acessíveis a grandes organizações.
Os valores estimados para o Mercado de ASP oscilariam entre $US 150 Milhões a $US 132.000 Milhões, entre 1999 e 2006, de acordo com vários investigadores [Lacity 2001], [Kern et al. 2001] e analistas de mercado. No entanto, apesar o conceito de ASP parecer estar a ser destinado a vingar, a questão que se colocava muitas vezes era se realmente os Clientes estavam à procura de ASPs. No início de 2000 a percepção era a de que ainda havia muito poucos clientes e o Gartner Group [Kern et al. 2001] previa que 60 por cento dos fornecedores ASP abandonariam este negócio.

Em Portugal também se assistiu ao crescimento do mercado de ASP, a partir de 1999, inicialmente com os operadores Telepac (TASP), Novis (Rent@soft), Hiperbit.com (YASP) e a e-Chiron, a que se seguiram alguns outros [Caçador 2000]. Tendo o Autor colaborado com um destes operadores desde a sua génese ("start-up"), cedo se constatou que o que na altura era um sonho rapidamente se revelou um modelo de negócio para o qual o mercado nacional ainda não

estava preparado, em especial ao nível das PME, pois entre as principais limitações da altura contava-se a diminuta ligação das empresas à Internet ou a capacidade das comunicações. Era aliás caricato, nalgumas situações, chegar-se às instalações das empresas destinatárias e verificar que não havia suficiente sinal da rede de comunicações para se conseguir ter um acesso conveniente ao "datacentre" onde as aplicações estavam instaladas. Por outro lado e como se estava nalguns casos a falar de aplicações na área financeira ou de gestão de sistemas da Qualidade, o receio por parte dos gestores das potenciais empresas Clientes sobre a localização dos dados e sua privacidade era um obstáculo muito grande à adopção.

A principal questão na altura era também qual ou quais as razões que levavam à adopção de ASP, principalmente no caso das PME e no início dos anos 2000, sendo que os fornecedores já consideravam que o controle de Custos permitido (previsibilidade sobre o budget de TI) [Udo 2000] era um dos principais; mas outros eram também apontados, tais como a falta de recursos qualificados para gerir as TI ou a focalização da empresa no "core" do seu negócio ou a possibilidade de efetuar o "outsourcing" que facilitasse a reengenharia dos processos de negócio. No entanto ao mesmo tempo o receio por parte dos clientes, de perder o controle sobre as suas aplicações era dado como um fator para a não-adopção. Sensivelmente uma década depois do emergir dos ASP, começou a generalizar-se a terminologia em torno de "Cloud Computing" e em particular, de "Software as a Service" (SaaS).

Comparando o ASP com o SaaS, as principais diferenças situam-se na escalabilidade e no facto de o ASP estar principalmente baseado na instanciação "single-tenant" ao invés da "multi-tenant" mais típica do SaaS, embora muitas vezes se utilize erradamente esta designação quando na realidade o que se está é a fazer "hosting" de uma única aplicação. Em especial a partir de 2010 [Rodrigues et al. 2014] começa a notar-se o interesse em estudar os Fatores de Adopção de SaaS. Do lado dos Fornecedores começou a assistir-se a uma crescente "pressão" para direcionar os Clientes no sentido de adotarem este modelo de distribuição.

Num estudo [Johansson 2013] realizado junto de 20 especialistas de um grande fornecedor de SaaS, as três razões mais apontadas (num total de dez) para a adopção ou não de aplicações de negócio ERP em SaaS foram as relacionadas com Custos/Factores Financeiros, Segurança e Disponibilidade. Todos os participantes no estudo identificavam a redução de custos como um beneficio para adoptar um ERP em SaaS mas ao mesmo tempo apontavam os factores de segurança, confidencialidade e proteção dos dados dos Clientes como factores que frequentemente vinham à discussão no sentido de serem inibidores para a adopção e uso. O outro factor que era também frequentemente apontado era o da disponibilidade e acesso às aplicações, nomeadamente com receios de "downtime" ou de falhas nas comunicações.

Recentemente o investigador realizou outro estudo [aceite e em fase de pré-publicação], complementar ao de 2013, com a visão de Clientes do mesmo fornecedor, de forma a confirmar a importância ou não destes factores, o qual trouxe algumas surpresas. Nomeadamente, apesar de se confirmar a importância do factor Financeiro/Custo, verificou-se que a importância relativa da Segurança era muito menor, sendo substituída por um factor não considerado pelo Fornecedor, que é a importância e confiança no Integrador de Sistemas (ou parceiro). Curioso foi o facto de um dos Clientes, no caso um "Partner" de uma empresa de Auditoria e serviços de "Business Process Outsourcing" referir que a Segurança era apenas um tema de discussão dos informáticos, dado que do seu ponto de vista o que era realmente importante eram os factores financeiros pois o nível de confiança no Fornecedor e no parceiro Integrador de Sistemas era tal

que não se punha em causa qualquer questão relativa à Segurança. Em terceiro lugar também se destacou a importância da Disponibilidade, existindo um ponto de vista positivo, dado que um dos Clientes referiu que era uma garantia e de que não necessitava portanto de estar a investir em equipamento de "alta disponibilidade".

2. Os modelos de adopção "Cloud pura", "On-premise" e Híbridos

Uma das discussões a que se assiste frequentemente é a de qual o melhor modelo de adopção e uso das aplicações de negócio (ERP, HRM, SCM, CRM, etc), a qual é frequentemente centrada no dilema da escolha entre a hipótese "On-premise" versus SaaS. Na maior parte das vezes os fornecedores e integradores também não ajudam a esclarecer esta discussão e a dar toda a informação para que o Cliente possa escolher a opção mais adequada e para tornar esta discussão mais complexa começa também a ouvir-se falar, cada vez mais, em modelos Híbridos. Vejamos então quais os tópicos que poderão estar em jogo na resolução desta equação que implica sempre investimentos consideráveis e onde os da tecnologia são apenas uma parte.

Num estudo [Ruivo et al. 2015] realizado com base em entrevistas a dezenas de CXOs (CEO; CFO; CMO; CIO, etc), aquando da feira CEBIT 2014, de PMEs e grandes organizações em vários sectores, sobre os modelos de adopção de ERP, resultaram algumas conclusões interessantes, das quais se realçam:

1) O modelo "On-premise", monolítico e baseado num único fornecedor, continua a ser predominante, dado também que em muitos casos ainda se está a assistir à implementação das aplicações de negócio como forma de normalizar e "standardizar" processos;

2) Os modelos mais "Cloud pura" terão tendência a predominar em pequenas ou médias organizações com menores níveis de sofisticação dos seus processos;

3) O modelo de ERP Híbrido (parte "On-premise" – nomeadamente os módulos mais críticos para o negócio, complementado com alguns módulos em SaaS) será o dominante no prazo de 10 anos, em especial nas médias e grandes organizações.

Quando inquiridos sobre o/os principais fatores que motivam a mudança das aplicações de negócio para a "cloud", foi também indicada a pressão do negócio e dos principais "stakeholders" da organização para reduzir os custos de TI, apesar de ser crucial a forma como as necessidades do negócio serão satisfeitas, suportadas e geridas, incluindo com uma visão clara dos custos e benefícios ao longo do ciclo de vida.

3. O Impacto no Negócio e Organização

Entre os maiores constrangimentos das organizações para acelerarem a adopção e uso de modelos mais modernos de aplicações de negócio (sejam "Cloud pura" ou Híbridos) contam-se a falta de visão clara sobre os "roadmap" dos fornecedores e a falta de planos para gerir a mudança [Ruivo et al. 2015], nomeadamente ao nível da preparação dos seus RH, das implicações nos seus processos de negócio e na forma de gerir a tecnologia, entre outros.

Em particular sobre este segundo aspecto, intrínseco à organização, os fornecedores têm aqui uma oportunidade a explorar para conseguir aumentar o nível de adopção do SaaS. Alguns dos maiores fornecedores, como é o caso da EMC e da Microsoft [EMC 2013], [HBR 2015], já têm

programas específicos (“Cloud and IT-as-a-Service for Business Transformation”; “Cloud Profitability Program”) que abordam estes temas, nalguns casos mais direcionados para o interno das suas organizações ou para o seu canal de parceiros, pois também ainda é grande a necessidade de efetuar a transformação ao nível interno dos fornecedores.

Neste processo de transformação para as organizações, devem ser contemplados vários aspectos, de forma a retirar todos os benefícios e reduzir os riscos, entre os quais destacamos os seguintes:

. Impacto na Direção de Informática:
Nas maiores organizações o impacto desta transformação faz-se sentir desde logo pela necessidade e relevância de algumas funções. Por um lado surgem novas funções, como é o caso dos “Cloud architects” (Arquitectos de sistemas “Cloud”) e, por outro, ganham relevância outras como os “Service managers” (Gestores de Serviços). Também o papel dos Responsáveis de TI (Diretores/CIOs) torna-se cada vez mais um papel estratégico e não apenas um de suporte - ao evoluírem de organizações centradas na tecnologia, para organizações centradas nos serviços de TI, os departamentos de TI que usualmente são vistos como centros de custo, têm uma grande oportunidade em tornarem-se os parceiros estratégicos de negócio para a organização. O foco deixa de ser a gestão e disponibilização de “assets” tecnológicos para uma gestão de oferta e procura de serviços para o negócio;

. Impacto nos Processos, Organização e Recursos Humanos:
Com a maior flexibilidade permitida pela contratação de serviços de TI / SaaS, surgem oportunidades para fazer alteração ou reengenharia de processos. Há também impacto ao nível da organização e dos recursos humanos, sendo necessário prever que o uso dos sistemas em SaaS necessita de ser potenciado pela preparação (“readiness”) das pessoas, na sua utilização. Abrem-se também oportunidades que nalguns casos não existiam, como a possibilidade de acesso às aplicações em qualquer local ou a disponibilização em dispositivos móveis;

. Aspectos de Contratação e Financeiros, “Governance” e Conformidade:
Um dos aspectos a ter especial cuidado é o da contratação. A organização passa a consumir serviços, os quais podem variar de acordo com o número de utilizadores ou de aplicações/funcionalidades/transações e torna-se portanto necessário acautelar nos contratos esta flexibilidade e o impacto nas condições financeiras, bem como os níveis de SLA e penalizações por deficiência ou falha nos serviços. Mesmo ao nível interno da organização, o modelo financeiro também muda, pois passa a ser possível imputar valores de utilização de recursos TI, por departamento, de acordo com os consumos reais.

Os aspectos de “Governance” e Conformidade são também pontos a ter em especial conta, dada a mudança de gestão de “assets” em “casa própria” para serviços em fornecedores externos. Ao assumir uma arquitectura de sistemas baseados em SaaS e modelos Híbridos, as organizações não estão apenas a mudar a forma de disponibilização dos recursos de TI e têm uma oportunidade estratégica para, de uma forma holística, tornarem-se mais ágeis, com uma capacidade de resposta mais rápida às condições de mercado e portanto obterem uma maior competitividade.

4. O Impacto do Uso e o Valor resultante de aplicações em SaaS, na Performance Empresarial das Organizações

Ao nível do Valor que é obtido pela utilização de SaaS e do impacto que isso tem na performance das Organizações, este ainda é um campo pouco trabalhado do ponto de vista científico. Existem muitos estudos, apontando vantagens e benefícios, nomeadamente realizados pelos mais reputados Consultores e Analistas de mercado, em muitos casos também co-participados pelos fabricantes, mas que carecem de confirmação científica.

Como se viu um dos principais impactos esperados pelas Organizações é a redução de Custos. E sobre esse ponto acrescento um exemplo de Caso de Estudo realizado pelo Autor, em 2010 [Rodrigues 2010], sobre a implementação e uso de uma solução de CRM para a IDC Portugal, disponibilizada em modo de aluguer em "hosting". Entre as vantagens apontadas e percecionadas pelos responsáveis da IDC Portugal estavam a rapidez de implementação e a garantia de escalabilidade com custos mensais bem identificados e previsíveis. Estes responsáveis apontaram como resultados, uma poupança da ordem dos 7% nos custos operacionais (ao optar pelo "hosting" versus uma implementação tradicional) e um aumento estimado em 10% na produtividade.

Mas torna-se necessário criar modelos que expliquem estes impactos de forma mais generalizada e com rigor, sendo que desde logo surge a primeira questão: Como se mede Performance Empresarial? Os investigadores que já realizaram trabalhos em grandes organizações ou corporações optam em muitos casos por medir o impacto na valorização bolsista da empresa, mas esse não é o critério que deva ser aplicado para a esmagadora maioria das PME.

Alguns investigadores apontam para os factores Rentabilidade (volume de negócios, margem de lucro e variação anual da margem de lucro) e Performance Competitiva (quotas de mercado; variação das quotas de mercado), como forma de medir o impacto na performance das empresas que utilizam SaaS. Mas há outros factores que é necessário ter em consideração, para que se possam obter conclusões reais, nomeadamente, em que sectores de mercado estamos a fazer a análise, qual o grau de comparação do universo de empresas, geografias, maturidade organizacional e de processos, quais os tipos de aplicações SaaS envolvidas e qual o período de tempo de utilização (1 ano? 2 anos? 3 anos?).

Daí que toda a investigação que possa ser realizada sobre o impacto no Valor e na Performance Empresarial das organizações, resultantes da adopção e uso de aplicações de gestão em modelos de distribuição distintos do tradicional "On-premise", será um importante contributo para ajudar tanto as organizações Clientes como as de Fornecedores e Integradores a terem uma visão mais clara dos modelos a adotar de acordo com a situação específica de cada organização.

5. Conclusões

Os fatores de adopção e uso de aplicações em SaaS, passados quinze anos desde que surgiram os primeiros ASP, continuam a girar em torno dos custos e razões de ordem financeira, da segurança e privacidade dos dados e também, de forma importante, na confiança sobre a idoneidade e capacidades futuras dos Fornecedores e Integradores de sistemas.

Depois de uma fase inicial em que se confrontavam os modelos “On-premise” e “Cloud pura”, atualmente fala-se muito em modelos Híbridos de adopção da tecnologia de aplicações de negócio, que deverão ser preponderantes para as médias/grandes organizações. Estas preferem continuar a manter alguns dos sistemas considerados mais críticos em “On-premise”, complementando com outras soluções em SaaS para áreas funcionais mais operacionais. Ao invés, nas pequenas organizações, dado o menor nível de sofisticação de processos e sem especialistas para fazer a gestão dos sistemas de informação, predominará a adopção de soluções “Cloud pura”.

Outro aspeto importante é o facto de a mudança de paradigma de exploração de soluções de software em modo “On-premise” para modelos em “Cloud pura” ou Híbridos ser mais do que uma decisão meramente tecnológica para passar a ter impacto em toda a Organização, nos Recursos Humanos, nos Processos e no “Governance” da empresa, pelo que se devem considerar estas mudanças numa perspectiva holística, de forma a extrair todo o potencial de benefícios realizável e reduzir os riscos associados.

Este artigo apresenta algumas limitações, dada a natureza ainda prospectiva de alguns trabalhos que o investigador está a realizar nos domínios da demonstração do Valor do SaaS e do impacto na Performance Empresarial das Organizações, especialmente das PME. Um desses trabalhos, aceite e em fase de pré-publicação refere-se às diferentes perspectivas de adopção de aplicações em SaaS, por parte dos Clientes face aos prestadores de TI. Existem também outros trabalhos que será importante desenvolver, tais como sobre os objectivos e factores críticos de decisão na adopção de serviços de TI em modelo Cloud-puro, on-premise e hibrido ou sobre o impacto no negócio com a prestação de serviços de TI baseada em modelos SaaS, nos vários modelos (Cloud pura ou hibrida), nas áreas operacionais, departamental ou por unidade de negócio, ou a nível estratégico/empresarial. Estes trabalhos serão também importantes para credibilizar e reforçar o esforço que os Fornecedores e Integradores estão a fazer para acelerar a transição para estes novos modelos de distribuição de software, por parte das Empresas.

6. Referências

M. Lacity and L. Willcocks, “Global information technology outsourcing: in search of business advantage”. Wiley, 2001

T. Kern, M. C. Lacity, L. Willcocks, R. Zuiderwijk, and W. Teunissen, "ASP Market Space Report 2001. Mastering the Customers’ Expectations". GMG report, 2001

Fátima Caçador, “Um Mercado a começar”. Publico http://www.publico.pt/noticias/jornal/um-mercado-a-comecar-150911 , 2000

G. G. Udo, "Using analytic hierarchy process to analyze the information technology outsourcing decision". Industrial Management + Data Systems, 2000

Jorge Rodrigues, Pedro Ruivo, Tiago Oliveira, “Software as a Service Value and Firm Performance - a literature review synthesis in small and medium enterprises”. 2014

Björn Johansson, Pedro Ruivo, “Exploring Factors for Adopting ERP as SaaS”. 2013

Pedro Ruivo, Jorge Rodrigues, Tiago Oliveira, “The ERP surge of hybrid models - an exploratory research into five and ten years forecast”. 2015

EMC, “Introducing Cloud and IT as a service for business transformation”. 2013

“The Digital Transformation of Business”. HBR - Harvard Business Review Analytic Services Report, 2015

Jorge Rodrigues, “IDC Portugal aumenta produtividade e reduz custos operacionais com Microsoft Dynamics CRM em modelo “hosted””. Microsoft Customer Case Study, 2010

Technical Writing for Portuguese Product Companies

Joaquim Baptista [1]

1) Knowledge management consultant, Portugal

px@acm.org

Abstract

The economic crisis in Portugal stimulated the creation of product companies that sell worldwide. Over time, these companies accumulate a capital of knowledge (technical, domain, user) that benefits from governance and engineering by dedicated professionals, traditionally called technical writers.

Technical writers have the ability to learn, organize, and transform knowledge, using a rare combination of technical, language, abstraction, and people skills. There is a growing practice of technical writing in Portuguese companies, and professional communities of writers, but University-level training is mostly absent.

Sophisticated Portuguese product companies would benefit from further investment in knowledge management practices, using senior talent to systematically audit, retain, and improve knowledge capital.

Keywords: technical writing, knowledge management, product company, Portugal

1. Context

The economic crisis in Portugal spurted a tendency for internationalization in Portuguese companies. New companies, often immersed in the so-called startup culture, start with the intention to sell abroad, sometimes almost exclusively. The core of these companies is often a single software product, initially created by a small core team of passionate developers.

Over time, product companies hire more people with an ever-widening range of roles and skills. These new people and roles impose their own forces onto the growing product:

- Marketing wants new flashy features that lure buyers and differentiate the product from the competition.
- Services people want to foster an impression of professionalism in front of customers, and may take shortcuts to do so.
- Support people want to solve problems quickly and close tickets, especially when a customer is shouting over the phone.

- Developers want to protect the architecture and stability of the product, but at the same time they must deal with the questions and requirements asked by customers and other teams.

2. Problem

How do all the different people access the knowledge needed for their tasks? What started as a non-issue when the company was a small team grows in complexity as the company and the product grow.

One particular issue that stresses the knowledge of the company is the emergence of the sophisticated user, who tries to use the product in ways that go beyond the experience of the typical company employee. The company as a whole struggles to meet the unexpected demands of these users. As a result, the company feels a growing need to capture and structure its own knowledge, including both implicit knowledge (in people's heads) and explicit (in documents and other media).

3. Typical methods

Product companies typically employ a combination of the following methods to capture and disseminate knowledge:

- Marketing creates demos and short articles that show off new flashy features. However, these typically avoid any hint of complexity to avoid scaring potential buyers.
- Services people create and deliver training, including classroom training with exercises, in-depth video tutorials, or step-by-step tutorials. These will not cover all possible usage scenarios, and students struggle to recall the training and to apply the training to their specific needs.
- After explaining parts of the product to customers and new employees, developers write technical information for technically minded people, including themselves. Developers also write a user guide (or embed the guide into the product), but the result is often not helpful enough for the actual users.
- Developers improve usability and user experience, therefore trying to incorporate more domain knowledge in the product [Armour 2004]. The ultimate objective is to make the product so simple that neither training nor manual is needed. This strategy can only succeed if the product does not grow functionality over time.

- Management adopts social tools that go beyond email to streamline communication and capture conversations. However, useful conversations require community building and curating the resulting knowledge. People stop using the social tools if questions are left unanswered.

4. Solution

Given enough time, the growing complexity of product and company limits the effectiveness of naïve solutions. The solution is to dedicate people to the task of engineering the flow of knowledge of the company. Traditionally these professionals have been called technical writers, but also technical communicators, information developers, and knowledge managers, among other titles.

Technical writers have the ability to learn, organize, and transform knowledge. Their set of skills includes the following:

- Interview people, typically developers and users.
- Develop models of users and user tasks.
- Learn by using the product, and advocate improvements on behalf of users.
- Organize knowledge into a coherent whole.
- Filter and transform the knowledge into the needs of specific user profiles.
- Devise a maintainable documentation set for delivering the knowledge that fulfills the business objectives.
- Write clear and objective text that fulfills a specific purpose, for example guiding a user to complete a task.
- Create effective media, catering for future change.
- Update the information deliverables, balancing priorities and resources with the need to keep the documentation set useful and coherent.
- Collaborate with other writers to maintain tone, style, and structure.
- Use professional tools that enable and streamline all of the above.

The documentation set may include online help, guides, tutorials, illustrations, videos, demos, lessons, web sites, and so on, but also forums, wikis [Baptista 2010], or even training [Baptista 2004]. The correct mix depends on the business priorities but also on the resources available.

5. Issues

Reference technical writing books [Hackos 1994] describe guidelines for short-term consulting projects in great detail. Content strategy [Halvorson 2009] provides guidelines for consultants that must maintain large sites. However, there is a scarcity of reports of the long-term processes required to maintain mature evolving products, such as [Colwell 2005].

Product companies typically sell to a global market, with most of their revenue coming from abroad. Therefore, they develop most communication materials in English, and sometimes translate to other languages.

While employees can often write a document in English, there are two problems looking forward: maintaining the document over time, as changes and additions appear and as features creep; and collaborating with several writers to maintain the document, while preventing the result to resemble patchwork.

Translation often exacerbates any issues by penalizing inefficient processes that recreate knowledge as needed while benefiting from efficient processes that reuse knowledge as much as possible. These processes benefit from tools outside the mindset of the office productivity software.

Writing teams seldom have the luxury to have enough manpower to polish materials into perfection. Most of the times, teams have to settle with some soft of bare minimum, and then wisely applying extra resources to areas where they will matter the most. Even a team of one must allocate its resources wisely.

Since technical writing is a little known area in Portugal, writers can easily receive unrealistic requests. For example, create a two-minute video that magically explains away the complexity of a 10-year-old product, or write a short guide for a domain that befuddles users, developers, and management.

While most functional areas can get away with understanding their part of the problem, writers must face the full complexity facing the company, or at least its users. This complexity can be overwhelming.

There is no formal training in technical writing in Portugal, although you can find people that enjoy writing and communicating. Most practitioners transitioned from other professional areas, although you can also find people that received training abroad. Junior practitioners lack the training and experience to tackle difficult problems, while senior-level skills are rare.

6. Practice in Portugal

Conversations within the community of technical writers in Lisbon have shown a surprising diversity of knowledge governance scenarios. These mature companies based on a software product have unique challenges that require unique technical and human solutions:

- Nokia Siemens Networks (and now Coriant) employs as many as 40 writers working in several separate products. In general, writers are part of a well-defined process that streamlines production. They also tried embedding writers in agile teams.
- Altitude Software has a documentation team with 3 to 6 writers, which became responsible for developing user guides, online help, training lessons, and illustrations [Baptista 2014]. They use wikis and XML-based tools [Baptista 2008] to promote reuse and limit the cost of localization.
- AnubisNetworks (now a Bitsight company) uses the product manager to manage the documentation as a part-time effort. Developers and marketing write most documents, but the product manager adds oversight and consistency.
- SISCOG uses a full time writer to maintain online helps, and a part-time writer to improve marketing materials. SISCOG is now investing in tools and processes to promote reuse.
- VisionBox started a writing team of 3 to impose consistency and control over the documents created by developers. They are considering wikis for documentation.
- OutSystems has a growing team of technical writers that emphasize user experience and video-based training.
- Feedzai is struggling to start a writing team to build on the documentation created by developers.

7. Recommendations

Sophisticated product companies accumulate significant intellectual capital, and therefore have sophisticated knowledge needs. These companies would benefit from bringing in senior talent that can audit the current situation of knowledge management and propose improvements. These improvements may reassign responsibilities, change processes and workflows, introduce new tools, and hire or train new skills.

Larger companies may justify a permanent senior position for knowledge management, with a corresponding team. Smaller companies may use consultants to identify needs, structure

deliverables, introduce tools, and train people to create the deliverables. Companies may also use consultants to create some deliverables, and then train people to maintain them.

Companies may also use experienced writers, acting as editors or consultants, to set up and contribute to a framework that enables other professionals to share their knowledge in a way that resonates with the culture and needs of each company or community. Such a framework requires a mix of tools, practice, and training, but also governance, writing, and editing.

Consolidating and optimizing the existing intellectual capital enables companies to leap forward, by doing more business, or by tackling harder problems.

8. Acknowledgements

The author wishes to thank the "Technical Writers @ Lisbon" for motivating this article and generously sharing their knowledge, and the anonymous reviewers for their helpful suggestions.

9. References

Armour, P. G., *The Laws of Software Process: A New Model for the Production and Management of Software,* Auerbach Publications, Florida, 2004.

Baptista, J., Just-in-time training workshops. *STC Management SIG News 8, 2* (2004), 14-16.

Baptista, J., Pragmatic DITA on a budget, in *Proceedings of the 26th annual ACM international conference on Design of communication* (Lisbon, Portugal, September 22-24, 2008). ACM, New York, NY, USA, 193-198.

Baptista, J., The birth of a company-wide wiki, in *Proceedings of the Workshop on Open Source and Design of Communication* (Lisbon, Portugal, November 8, 2010). ACM, New York, NY, USA, 7-10.

Colwell, R. P., *The Pentium Chronicles: The People, Passion, and Politics Behind Intel's Landmark Chips,* Wiley-IEEE Computer Society Pr., 2005.

Hackos, J., *Managing Your Documentation Projects.* John Wiley & Sons, Inc., New York, NY, USA, 1994.

Halvorson, K., *Content Strategy for the Web,* New Riders Publishing, Thousand Oaks, CA, USA, 2009.

Associação Portuguesa de Sistemas de Informação

"...ligamos instituições do ensino superior, empresas e administração pública"

FCT

Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

20 anos

noesis

FCA