Catarina Inês Marques de Lucena

Mestre em Engenharia Electrotécnica e de Computadores

Framework for collaborative knowledge management in organizations

Dissertação para obtenção do Grau de Doutora em

Engenharia Electrotécnica e de Computadores - Especialidade em Sistemas de Informaąão Industriais

Orientador:	Ricardo Luís Rosa Jardim Gonçalves, Associate Professor, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa
Co-orientador:	João Filipe dos Santos Sarraipa, Invited Researcher, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Júri

Presidente:	Prof. Doutor Fernando José Pires Santada
Arguentes:	Prof. Doutor Nejib Moalla
	Prof. Doutor João Pedro Mendonça de Assunção da Silva
Vogais:	Prof. Doutor Driss Quazar
	Prof. Doutor Virgílio António da Cruz Machado
	Prof. Doutor João Francisco Alves Martins
	Prof. Doutor Carlos Manuel de Melo Agostinho
	Prof. Doutor Ricardo Luís Rosa Jardim Gonçalves

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Dezembro, 2016

Framework for collaborative knowledge management in organizations

Este documento foi gerado utilizando o processador (pdf)LaTeX, com base no template "unlthesis" [1] desenvolvido no Dep. Informática da FCT-NOVA [2].
[1] https://github.com/joaomlourenco/unlthesis [2] http://www.di.fct.unl.pt

To my loyal writing companions ☺:
Buli and Francisca.

Acknowledgements

My acknowledgments to Marcelo, my boyfriend and companion, by his total encouragement. He helped me to face all the difficulties met in this period.

My acknowledgments to my parents and sister for their constant encouragement and support. I know that they are very proud.

A special acknowledge as well for José and Bia, for helping me holding the situation at home and give me time to finish this work that without them would not exist.

My appreciation to my supervisor, Ricardo Jardim-Goncalves, and co-supervisor, João Sarraipa, for gave me acknowledge of this research area. They trusted my skills and allowed me to freely explore it accordingly with my motivations and instincts. They pushed me to be better and better.

I also want to acknowledge Carlos Agostinho contribution to this work. I grew a lot while working with him in my last year at GRIS. The experience that this period gave was very valuable.

Then, I want to give a special thanks to my ex-colleagues at GRIS, my ex-research group, especially those who worked closer to me, Raquel, Pedro, Fábio, Marcio, José e Paulo. With them I felt that was impossible to have a boring working day.

I want to thanks João Serras, by Manager at Synopsys, by his support. For someone who knows so few about this research work, he really pushed me to finish it.

Finally, my great thanks to both my institutions, Uninova and FCT for hosting this research.

Abstract

Nowadays organizations have been pushed to speed up the rate of industrial transformation to high value products and services. The capability to agilely respond to new market demands became a strategic pillar for innovation, and knowledge management could support organizations to achieve that goal. However, current knowledge management approaches tend to be over complex or too academic, with interfaces difficult to manage, even more if cooperative handling is required. Nevertheless, in an ideal framework, both tacit and explicit knowledge management should be addressed to achieve knowledge handling with precise and semantically meaningful definitions. Moreover, with the increase of Internet usage, the amount of available information explodes. It leads to the observed progress in the creation of mechanisms to retrieve useful knowledge from the huge existent amount of information sources. However, a same knowledge representation of a thing could mean differently to different people and applications.

Contributing towards this direction, this thesis proposes a framework capable of gathering the knowledge held by domain experts and domain sources through a knowledge management system and transform it into explicit ontologies. This enables to build tools with advanced reasoning capacities with the aim to support enterprises decision-making processes. The author also intends to address the problem of knowledge transference within an among organizations. This will be done through a module (part of the proposed framework) for domain's lexicon establishment which purpose is to represent and unify the understanding of the domain's used semantic.

Keywords: Knowledge Creation, Tacit Knowledge, Explicit Knowledge, Collaboration, Knowledge Management, Ontology Learning, Machine Learning, Data Mining.

RESUMO

Hoje em dia, as organizações são obrigadas a acelerar o seu ritmo de criação de produtos e serviços de elevada qualidade e valor económico. A capacidade de responder com agilidade aos novos requisitos do mercado tornou-se um pilar estratégico para a inovação e a gestão do conhecimento do domínio em que a organização se encontra inserida poderá suportar esse objectivo. No entanto as abordagens de gestão de conhecimento actuais tendem a ser muito complexas ou académicas, com interfaces difíceis de gerir. Essas características são ainda mais notórias quando a gestão colaborativa e cooperativa de conhecimento é necessária. Num quadro ideal, tanto o conhecimento tácito e explicito devem ser abordados para alcançar uma gestão de conhecimento em que definições precisas e significativas dos conceitos do domínio são tidas em consideração. Isto tem ainda mais impacto com o aumento explosivo da quantidade de informação disponível devida à utilização crescente da Internet. Deste modo, muitos mecanismos para a aquisição de conhecimento útil a partir dessas fontes têm sido observados. No entanto, a representação do conhecimento de uma mesma entidade pode ter significados diferentes para diferentes pessoas e aplicações.

Contribuindo para esta direcção, esta tese propões uma *framework* capaz de gerir o conhecimento detido pelos especialistas e documentos do domínio através de um sistema de gestão de conhecimento e transforma-lo em ontologias explicitas. Isto permitirá a construção de ferramentas com capacidade avançada de raciocínio que poderão apoiar as empresas em processos de tomada de decisão. O autor também tem a intenção de abordar o problema de transferência de conhecimento dentro e entre organizações. Isto será feito através de um modulo (parte da *framework* proposta) para o estabelecimento de um léxico de domínio, cujo objectivo é unificar a compreensão da semântica utilizada.

Palavras-chave: Criação de Conhecimento, Conhecimento Tácito, Conhecimento Explicito, Collaboração, Gestão de Conhecimento, Aprendizagem de Ontologias, Aprendizagem de Máquinas, Mineração de Dados.

CONTENTS

List of Figures **xix**

List of Tables **xxiii**

Acronyms **xxv**

1 Introduction **1**

1.1 Motivation 2

1.1.1 Knowledge Management 2

1.1.2 Knowledge Acquisition 3

1.1.3 Knowledge sharing 5

1.2 Vision 7

1.3 Adopted Research Method 7

1.4 Research Problem and Hypotheses 9

1.4.1 Research Questions 9

1.4.2 Hypotheses 9

1.5 Thesis Plan Outline 10

1.5.1 Section I - Introduction 10

1.5.2 Section II - Background Observations 11

1.5.3 Section III - Contributions & Validation 11

1.5.4 Section IV - Conclusions 12

2 Knowledge Creation Process **13**

2.1 Knowledge Creation Requirements 15

2.2 Role of Managers to Promote Knowledge Creation and Transfer in organizations 16

2.3 Knowledge Life Cycle 16

2.3.1 Knowledge Acquisition 17

2.3.2 Knowledge Representation and Formalization Degrees 21

2.3.3 Knowledge Use 23

2.3.4 Knowledge Maintenance 24

2.4 Discussion 24

3 Knowledge Management and Collaborative Systems 25

3.1 Knowledge Management Systems 25

3.1.1 Knowledge Management systems selection criteria 26

3.1.2 Knowledge Management systems typologies 29

3.2 Collaborative Systems 34

3.3 Interoperability as a Collaboration enabler 36

3.4 Discussion 38

4 Business Intelligence 39

4.1 Knowledge Management and Business Intelligence Integration 41

4.2 Data Mining 41

4.2.1 Association 41

4.2.2 Classification 42

4.2.3 Clustering 43

4.2.4 Forecasting 44

4.2.5 Regression 44

4.2.6 Sequence discovery 44

4.2.7 Visualization 45

4.3 Machine Learning 45

4.3.1 Artificial Neural Networks 46

4.3.2 Fuzzy Logic 47

4.3.3 Decision Tree 47

4.3.4 Bayesian Learner 48

4.3.5 K-Nearest Neighbor Learning 48

4.3.6 Markov Chains 49

4.3.7 Genetic algorithm 49

4.4 Ontology Learning 50

4.4.1 Ontology learning from structured data 51

4.4.2 Ontology learning from semi-structured data 52

4.4.3 Ontology learning from unstructured data 54

4.5 Discussion 55

5 Framework for Organizations Knowledge Management 57

5.1 Frameworks and Models for Knowledge Management 57

5.2 Proposed Framework for Organizations Knowledge Management 62

5.2.1 Knowledge Processes 64

5.2.2 Business Processes 67

5.2.3 External Processes 71

6 Conceptual Approach for Domain Experts Knowledge Management 75

6.1 Related Work 75

6.2 Framework for Knowledge Management 76

6.2.1 Wiki-based front-end contents formalization methodology . . . 77

6.2.2 Synchronization between wiki-based front-end modules and ontologies 79

6.3 Supporting The EISB During the ENSEMBLE Project 80

6.3.1 FinESPedia 80

6.3.2 Application of the methodology 81

6.3.3 Ensemble's Knowledge base synchronization 84

6.3.4 Enabling the integration with other works 84

6.4 Discussion 85

7 Definition of a Domain Glossary for Training Support 87

7.1 ALTER-NATIVA 88

7.2 Knowledge Management Approach 90

7.3 Thesaurus building by domain experts 93

7.3.1 Delete a concept 93

7.3.2 Create a concept 94

7.4 Meta-Knowledge Management 97

7.5 ALTER-NATIVA Services 97

7.5.1 Knowledge Base Services 97

7.5.2 Recommendation Services 98

7.6 Discussion 100

8 Requirements Management Using Semantic Wiki Modules 103

8.1 Going beyond traditional requirements management tools 104

8.1.1 Semantic Wikis for Requirements Management Support 106

8.2 Requirements Management Methodology 107

8.3 BDD based Requirements Management Framework 108

8.3.1 Instantiation of the Requirements Management Framework . . . 110

8.4 Implementation of the Requirements Management Framework in Industry 113

8.4.1 **Wiki front-end** 113

8.4.2 **Ontology** establishment for requirements specification 116

8.4.3 **Synchronization Module** 118

8.4.4 **Reasoning and Decision-Making** 119

8.4.5 Execution and Validation of Results 122

8.5 Discussion 123

9 Processes Modeling Approach 125

9.1 General Approach Model to Instantiate the OSMOSIS Processes 126

9.2 Liquid Sensing Enterprise Modelling Environment 129

9.2.1 OBM-OTIM Process Transformation 129

9.3 OSMOSE Process Manager Modelling Toolbox 131

9.3.1 General Overview 132

9.3.2 Modelling a Process 133

9.4 OSMOSE Use Case Example 136

9.4.1 OSMOSE Business Model 136

9.4.2 OSMOSE Technological Independent Model 137

9.4.3 OSMOSE Technological Specific Model 138

9.5 Discussion 138

10 Services Modeling and Information Exchange 141

10.1 Web Services Interoperability 142

10.2 Approach for the interoperability of the information exchange 143

10.3 Modeling Morphisms to Enable Sustainable Interoperability 144

10.3.1 Knowledge Enriched Tuple for Mapping Representations 144

10.4 Communication Mediator 145

10.5 Model Transformations 146

10.6 ISOFIN cloud Application Scenario 147

10.6.1 Domain Reference Lexicon establishment 148

10.6.2 Mapping Tool 148

10.6.3 Mappings Storage 152

10.6.4 Generate SBS WSDL file 153

10.7 Discussion 154

11 Customers' Feedback Analysis 155

11.1 Framework for customers' sentiment analysis 155

11.1.1 Methodology for Domain's Lexicon Building 158

11.2 Portuguese hospitality industry customers' comments analysis 161

11.2.1 Data collection and pre-processing 162

11.2.2 Data analysis 163

11.2.3 Knowledge Acquisition and Decision Making 166

11.3 Discussion 166

12 Conclusions 167

12.1 From a research question to validation 168
12.2 Integration with other research activities 169
12.2.1 ENSEMBLE 169
12.2.2 ALTER-NATIVA 170
12.2.3 ISOFIN 171
12.2.4 OSMOSE 171
12.2.5 SEMANTIK 173
12.3 Publications Summary 174
12.4 Future Work 175

Bibliography **177**

List of Figures

1.1 Motivation versus easiness to share. 6
1.2 Adopted Research Method. 8
1.3 Thesis Plan Outline. 10

2.1 Socialization, Externalization, Combination and Internalization Process of Knowledge Spiral. 14
2.2 Knowledge Life Cycle. 17
2.3 Tacit and Explicit Knowledge Acquisition. 18
2.4 An Overview of the Steps That Compose the KDD Process (from [62]). . . 20
2.5 Knowledge Formalization Degrees. 22

3.1 Hierarchical structure for Knowledge Management System typology selection (bases on [83] [87]) 27
3.2 Function view of document management. 30
3.3 A pyramid collaborative model (based on [133]). 34

4.1 Regression applied to Forecasting. 45

5.1 Framework for Knowledge Management: An European Perspective. 58
5.2 Maier's Framework for Knowledge Management [268]. 59
5.3 Wiig's Three Pillars of Knowledge Management [269]. 60
5.4 Stankosky's Four Pillars of Knowledge Management. 61
5.5 Fraunhofer IPK Reference Model from Knowledge Management [264]. . . 62
5.6 Framework for Organizations Knowledge Management. 63
5.7 Core activities identifies in a Product creation. 64
5.8 Knowledge Processes. 65
5.9 Proposed concept for Knowledge Management. 66
5.10 Business Processes. 68
5.11 Requirements Management Positioning in a product of service life-cycle. . 69
5.12 External Processes Processes. 72

6.1 Framework for Knowledge Management using wiki-based front-end modules. 77
6.2 Methodology for wiki-based front-end contents formalization. 78
6.3 FinESPedia Main Page. 81
6.4 Knowledge Base's taxonomy establishment. 82
6.5 New Publication demonstration Scenario. 83

7.1 ALTER-NATIVA's knowledge base (an excerpt). 90
7.2 ALTER-NATIVA's Knowledge Management approach. 91
7.3 Thesaurus Building methodology [336]. 92
7.4 Form to build a concept page. 94
7.5 Delete a concept with children associated. 94
7.6 Creation of a concept that points to an existing concept master. 95
7.7 Creation of a concept that points to a non-existing concept master. 96
7.8 The created concept is referred as a master concept already. 97
7.9 ALTER-NATIVA's WebProtégé management interface. 98
7.10 ALTER-NATIVA's WebProtégé management interface. 100

8.1 Requirements Engineering Methodology. 108
8.2 Requirements Management Framework. 109
8.3 Methodology for wiki-based front-end contents formalization. 111
8.4 Creation and edition of User Requirements form. 114
8.5 User requirements form developed for user stories (behavior) gathering. . . 115
8.6 Requirements pages tagging (folksonomy). 116
8.7 Knowledge Base's taxonomy establishment. 117
8.8 Direct correspondence between front-end and back-end establishment. . . . 118
8.9 Groups of common requirements identified. 121
8.10 Example of architectural components recommendation. 122
8.11 Execution and validation results. 123

9.1 OSMOSE Process Manager Modelling Architecture. 126
9.2 Waterfall Approach to instantiate OSMOSE processes modelling and execution. 128
9.3 Osmosis process selection in tool interface. 131
9.4 Osmosis Process Manager Workspace. 132
9.5 Creating an OBM project. 134
9.6 Modelling an OBM process with EA*. 134
9.7 BPMN derived from an EA* process. 135
9.8 Complementing an existing BPMN process. 136
9.9 EA* (OBM) model of Monitoring of the manufacturing tool. 137
9.10 OTIM model of the Example. 139

10.1 Approach for the interoperability of the information exchange. 143
10.2 Method for Horizontal Transformation. 147
10.3 Mapping Tool Use Case. 149
10.4 Mapping Tool working area. 150
10.5 OWL 2.0 to LIMM Transformation [391]. 151
10.6 WSDL 1.0 Meta-model. 152
10.7 WSDL 1.0 LIMM Graphical representation. 153
10.8 Mapping Tool demonstration. 153

11.1 Framework for customers' sentiment analysis. 156
11.2 Methodology for Glossary Building and Enrichment. 158
11.3 Synonyms gathering cycle. 160
11.4 Term placing in the domain thesaurus: one of the terms of the synonyms list is already in the domain thesaurus. 160
11.5 Term placing in the domain thesaurus: none of the terms of the synonyms list is yet in the domain thesaurus. 161
11.6 Methodology for Portuguese hospitality industry customers' comments analysis. 162
11.7 Example of phrase syntactic tree. 163
11.8 Excerpt of Booking room description. 164
11.9 Portuguese hospitality industry glossary building demonstration scenario. . 165

List of Tables

2.1 Ontological dimension categories. 13

4.1 Example of market basket transactions. 42
4.2 Comparative study between several OL approaches from database schemas (adapted from [227]). 52
4.3 Comparative study between several OL approaches from dictionaries (adapted from [227]). 53
4.4 Comparative study between several OL approaches from Document type definitions. 53
4.5 Ontology Learning from text (adapted from [227]). 54
4.6 Ontology Enrichment from text (adapted from [227]). 55

7.1 Use case of concept creation 95

8.1 Applicability for Elicitation Approaches following the work of [338] and [339]. 104
8.2 Requirement Management tools comparative study (based on the work of INCOSE [341]). 105
8.3 Technical Requirements Tests and Validation of Results (illustration). . . . 123
8.4 OSMOSE Outcomes in relation to User Requirements Elicitation. 124

9.1 Summary of changes made in EA* to BPMN 2.0 transformation. 130
9.2 Osmosis Processes changes in the EA* to BPMN 2.0 transformation. 131

12.1 Outcomes related to the Research Questions and Sub-Questions. 168
12.2 Publications by Research Area. 174
12.3 Publications made in the Control and Decision Area. 174

Acronyms

AHP Analytic Hierarchy Process.

ANN Artificial Neural Networks.

BI Business Intelligence.

BSM Business Service Model.

CM Communication Mediator.

CRM Customer Relationship Management.

DM Decision-Making.

DRD Digital Resource Description.

DSS Decision Support Systems.

DW Digital World.

EDMS Electronic Document Management Systems.

EK Explicit Knowledge.

FL Fuzzy Logic.

GA Genetic Algorithms.

H Hypothesis.

IS Information Systems.

KA Knowledge Acquisition.

KB Knowledge Base.

KDD Knowledge Discovery in Data.

KM Knowledge Management.

KMa Knowledge Maintenance.

KMS Knowledge Management Systems.

KNN K-Nearest Neighbors.

KPS Knowledge Portal Systems.

KR Knowledge Representation.

LIMM Language Independent Meta-Model.

LOM Learning Object Metadata.

LSE Liquid Sensing Enterprise.

MDA Model Driven Architecture.

MDI Model Driven Interoperability.

MDSEA Model Driven Service Engineering Architecture.

ML Machine Learning.

MO Mediator Ontology.

MoMo Morphism.

NBS Native Business Services.

NLP Natural Language Processing.

OBM OSMOSE Business Model.

OER Open Educational Resources.

OL Ontology Learning.

OTIM OSMOSE Technology Independent Model.

OTSM OSMOSE Technical Model.

P2P Peer to Peer.

PNP Personal Needs and Preferences.

RQ Research Questions.

RW Real World.

SBS Supplier Business Services.

SCM Supply Chain Management.

SOA Service Oriented Architecture.

TIM Technology Independent Model.

TK Tacit Knowledge.

TSM Technology Specific Model.

VLOs Virtual Learning Objects.

VW Virtual World.

W3C World Wide Web Consortium.

WSDL Web Services Description Language.

XSD Schema Definition Language.

CHAPTER 1

INTRODUCTION

Knowledge can be defined as information that has been understood and embedded in the brain. Thus, it is difficult to transfer between individuals due its individual oriented nature [1]. In this context, researchers consider Tacit Knowledge (TK) as the background knowledge that a person uses when trying to understand anything that is presented to him [2]. Explicit Knowledge (EK) is another type of knowledge, which can be expressed in words and numbers, and can be easily communicated and shared in the form of hard data, scientific formulae, codified procedures or universal principles [3]. By transforming TK into EK, it can be consulted and used by a full community, instead of being locked in a single community's element. However, the transformation of tacit knowledge into explicit knowledge can be considered one of the most challenging steps under Knowledge Management (KM).

The more communication, involvement, and interaction of people, more is the chance for organizations to expose TK residing in individuals' heads. Thus, the importance of developing services or mechanisms to gather knowledge from domain experts and documents (structured, unstructured, and semi-structured) has increased. The domain experts, as main actors, are who better know how to characterize their domain. The result of involving them directly in the knowledge acquisition process and transformation into EK is that tacit knowledge can be managed through communities' knowledge processes [4]. It is proven that TK has a crucial influence on the success of innovation processes in companies and plays a vital role as a company resource and success factor [5]. Thus, mechanisms and tools to support domain experts KM are required.

Besides domain experts' tacit knowledge, the usage of domain documents and other sources (e.g. web pages) can and be considered as an advantage when applied to organizations' Decision-Making (DM). However, it means that somehow domain sources need to be formalized into useful knowledge so they can they can actually be used to support entities DM

Decision-making is defined as the cognitive process resulting in the selection of a specific resource or action among several alternative scenarios. Every DM process produces a final choice [6]. Making context based decisions will reduce the effort to coordinate tasks and resources by providing a context in which to interpret utterances and to anticipate actions. Thus, gathered knowledge is used to achieve fact-based reasoning and inference, agility, reactiveness, innovation, context awareness, decision intelligence and competitive intelligence [7].

1.1 Motivation

This research work recognizes knowledge management as an essential asset for organizations success. Thus, the author's motivation is naturally focused on its integration on organizations' processes (knowledge, business and external). This section outline consists in a short background to identify some benefits of knowledge management and inherent challenges. Even tough knowledge acquisition and sharing are part of the knowledge management, a brief study about these concepts is also present as part of the motivation for this work.

1.1.1 Knowledge Management

Knowledge Management can defined as the process of applying a systematic approach to capture, structuring, management, and dissemination of knowledge throughout an organization to work faster, reuse best practices, and reduce costly rework from project to project [3, 8, 9]. In the following summary are identified some benefits related to knowledge management integration in organizations and also some of its challenges [10–12].

- **Benefits**
  - Enabling better and faster decision making;
  - Reusing ideas, documents, and expertise;
  - Taking advantage of existing expertise and experience;
  - Communicating important information widely and quickly;
  - Promoting standard, repeatable processes and procedures;

  - Providing methods, tools, templates, techniques, and examples;
  - Making scarce expertise widely available;
  - Showing customers how knowledge is used for their benefit;
  - Customer focus services and Targeted marketing;
  - Improves staff engagement and communication.

- **Challenges**
  - Information
    - Transforming vast amounts of data into usable form;
    - Avoiding overloading users with unnecessary data;
    - Eliminating wrong/old data;
    - Ensuring customer confidentiality;
    - Keeping the information up to date;
  - Management
    - Getting individuals to volunteer knowledge;
    - Getting business units to share knowledge;
    - Demonstrating business value;
    - Bringing together the many people from various units;
    - Determining responsibility for managing the knowledge.
  - Technology
    - Determining infrastructure requirements;
    - Keeping up with new technologies;
    - Security of data access.

1.1.2 Knowledge Acquisition

In contrast to older times, when finding sources of information was the key problem to companies and individuals, today's information society challenges companies and individuals to create and employ mechanisms to search and retrieve relevant data from the huge quantity of available information and mine it into knowledge, which can be used to take the most suitable decisions [13].

Data is the storage of intrinsic meaning, a mere representation. Is defined as symbols that represent properties of objects and their environment [14]. Because digital data are so easily shared and replicated, it enables a tremendous reuse opportunity, accelerating investigations already under way and taking advantage of past investments [15]. As a result, the amount of data and information available for organizations analysis is exploding [16]. As a response, the big data initiative seeks to glean intelligence from data and translate that into business advantage. On the top of the benefits are: (1) Better, fact-based decision-making (22%); and (2) Improve customers experience (22%) [17]. That means that organizations use big data platforms to give them competitive answers to important questions, namely the ones related to betters DM approaches and customers' satisfaction. Thus, the key added value is to speed up the time-to-answer period, allowing an increase in the pace of DM at both the operational and tactical levels [18][19].

Similarly to the knowledge management presented study, some benefits and challenges associated to knowledge acquisition were also identified [17, 20–23].

- **Benefits**
  - Better fact-based decision-making;
  - Improvement of customers' experience;
  - Increased sales;
  - Product innovations;
  - Reduced risks in decision-making;
  - More efficient operations;
  - Higher quality of products and services.
- **Challenges**
  - System builders tend to collect knowledge from one source, but the relevant knowledge may be scattered across several sources;
  - It is difficult to recognize specific knowledge when it is mixed with irrelevant data;
  - Builders may attempt to collect already documented knowledge ignoring other sources - the knowledge collected may be incomplete;
  - Automated acquisition of knowledge from data along with the representation of this knowledge;
  - Identify different sources for different needs.

1.1.3 Knowledge sharing

Nowadays, it can be considered that knowledge sharing is a big concern for organizations' managers. Organizations are *suffering* because its members are having trouble accessing the knowledge required to do their job. Furthermore, even if this knowledge exists in the organization, sometimes the people who have it just refuse to share. Thus, in the short and medium term, much of knowledge manager's efforts need to be focused on ways to promote knowledge-sharing behaviour [24]. The following summary consists in a background study about knowledge sharing and identifies related key challenges and benefits related with knowledge sharing [10, 25–30].

- **Benefits**
  - Speed up response time - Organization's member don't waste time searching for the knowledge;
  - Increases efficiency - Avoids having several members to work on the same thing (waist of time) when it could be shared;
  - Increases creativity and innovation - Knowledge sharing has a positive effect on creativity and ideation.
  - Better decision making - Decisions are made based on (shared facts) instead of speculation.
  - Preserving of existing knowledge - If knowledge is shared among organization's members then it will not be lost when a members decides to leave
- **Challenges**
  - Cultural - Culture boundaries may often restrict the flow of information and knowledge among employees;
  - Who should share what - How should know what; to what level of detail; and how can organizations support these processes of Knowledge Sharing;
  - Technology - It can make the exchange of information and knowledge to become faster, easier, and smoothly;
  - Leadership - Knowledge sharing depends strongly on leaderships assignment and how the role is performed;
  - Security, privacy - The act of knowledge sharing my consist in a security and privacy issue itself;
  - Standards - It includes agents' communication; meta-data representation; business integration success; portals; and advanced collaboration.

- Management - Management involves having the right leader doing the right activities to provide the appropriate changes that will improve the systems and yield good results;
- Economic - Knowledge creation and usage could return is progressively higher economic returns;
- Implementation - Implementation is this case is mostly seen as acceptance;
- Influencing Factors
  * Information technology - ICT tools, ICT infrastructure and ICT know-how;
  * Organization - Structure, culture, reward and recognition, work processes and office layout;
  * People - Awareness, trust, personality, job satisfaction.

The author focus is related with the creation of mechanisms that allow organization's members to share knowledge. The ease of sharing knowledge is likely to influence people's willing to share (see Figure 1.1). This is consistent with research on recycling behaviour and IT usage, which has shown that the harder the task, the more important is the quality of motivation for knowledge share [31, 32].

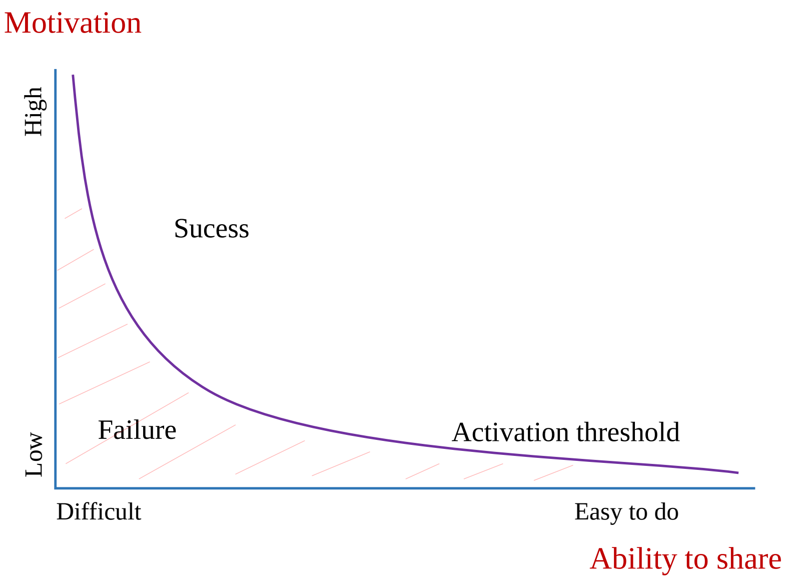


Figure 1.1: Motivation versus easiness to share.

1.2 Vision

The research work envisions the proposal of a framework focused on the knowledge management integration with the other organizational processes (e.g. business and external). The framework also aims to identify which are the internal (e.g. organization's experts) and external (e.g. customers and established networks) enablers for knowledge management.

In pursuing to the need of ease the knowledge sharing, the author also considered the development of an approach for knowledge sharing. The purpose of this approach is to facilitate tacit to explicit knowledge transference, and consequently make knowledge available for the organization instead of being locked in a single person.

Being the communication between teams with different background (e.g. business and technical) sometimes a hard task, the author also envisions an approach which allows different teams to collaborate using the same tool. This approach should also support the knowledge sharing between teams starting from the business goals definition to the processes implementations.

As part of the research, the author also wants to address the semantic interoperability of systems. By addressing this issues, the author intends to enable communication between domain entities (networks establishment), interpretation of external stimuli (e.g. customers' feedback) and awareness in resources search (e.g. people to interact, objects and services).

1.3 Adopted Research Method

The research method used in this dissertation is based on the classical research method [33] and it is composed by 7 steps. It starts by the definition of problem, and ends with the interpretation of the obtained results. It is possible to seen in Figure 8.3 that the first 6 steps are repeated cyclically until prove/show the studied theory. When the results show a positive result, they should be published and/or transferred to industry. By the last, the referred cycle must be carried out the required number of times in order to mature all the small research questions that compose the research overall objective. Each step of the adopted method is described in the topics bellow:

- **Research Question/ Problem** - A Research Questions (RQ) is an inquiry that is asked for the purpose of gaining knowledge or useful information on an area of interest to which the author is intended to participate and contribute for. Research questions are used to determine possibilities and gain valuable insight. The author RQ are presented in section 1.4.1;

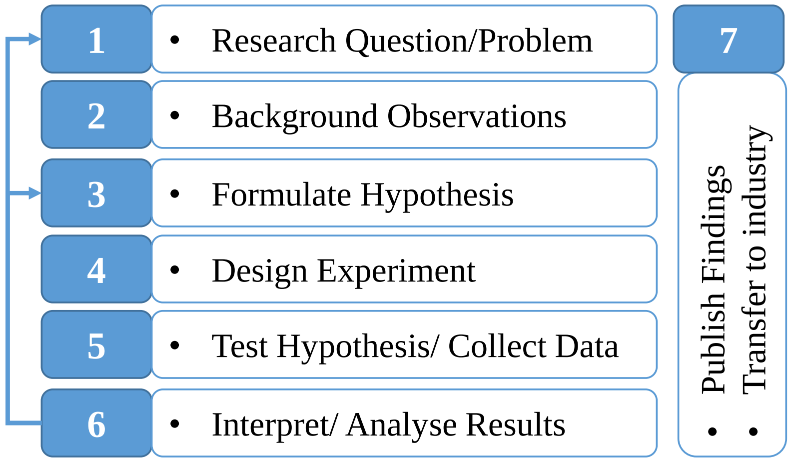


Figure 1.2: Adopted Research Method.

- **Background Observations** - The literature review is made in the background observations. It consists in the study of past similar works, presenting literature review and similar projects and works with the intention of support the dissertation start: studying existing ideas of other authors will create and open new solutions, to be used in the development for this dissertation;

- **Formulate Hypotheses** - Here is where are managed the predicted results of the research work. Important characteristics for a Hypothesis (H) to have are: 1) simple to understand; 2) specific; 3) conceptually clear; and 4) measurable. The hypotheses is presented in section 1.4.2;

- **Design Experiments** - Where are detailed the plan of the experimental phase steps. It is often composed by the design of a prototype or even a system architecture;

- **Test Hypotheses / Collect Data** - Consists in the evaluation of the system architecture and/or prototype is made. This is made by testing and simulating different scenarios.

- **Interpret / Analyse Results** - Further analysis upon each collected test data is made in order to validate the previously formulated H. If for some reason the results are not satisfactory, it is possible to try a different approach, and return to step 1. Moreover, when positive results are achieved, it is possible to look back to past steps giving some ideas for further research;

- **Publish Findings & Transfer to Industry** - When good results are achieved and a good contribution to the scientific community are made, the author should share these results with the community and in some cases transfer them to the industry. The presentation of the results to the scientific community is made through scientific papers (both in conferences and Journals) as an instance. The transference of results to the industry (e.g. methodologies and prototypes) required a previous acceptance to show that they can be used.

1.4 Research Problem and Hypotheses

1.4.1 Research Questions

Having in consideration the challenges presented in the motivation, the main RQ defined is:

RQ: How to increase knowledge availability and usage in organizations' processes?

The main research question can be decomposed in several small, and more focused ones:

- **RQ 1:** How to facilitate tacit to explicit knowledge transference?
- **RQ 2:** How to ease knowledge sharing between organizations multi-disciplinary teams?
- **RQ 3:** How can knowledge be effectively transferred, within and among organizations, facing the high number of semantic representation of the same segment of reality?

1.4.2 Hypotheses

The proposed Hypotheses (H) to address the defined research questions are?

- **H 1:** Since the ease to share is a key enabler for domain experts to expose their knowledge, **if** a proper approach which support them in this task is implemented **then** the motivation for knowledge share would increase contributing to organizations knowledge quality - as main actors, domain experts are o better know how to characterize their own features (e.g. resources and strategies).
- **H 2: If** a proper approach is implemented for multi-disciplinary teams to collaborate in the same tool, starting from business goals definition to processes implementation, **then** the knowledge sharing between them could be facilitated
- **H 3: If** mechanisms which are able to use text analysis techniques to extract knowledge from domain sources (users and documents) are implemented, **then** a domain

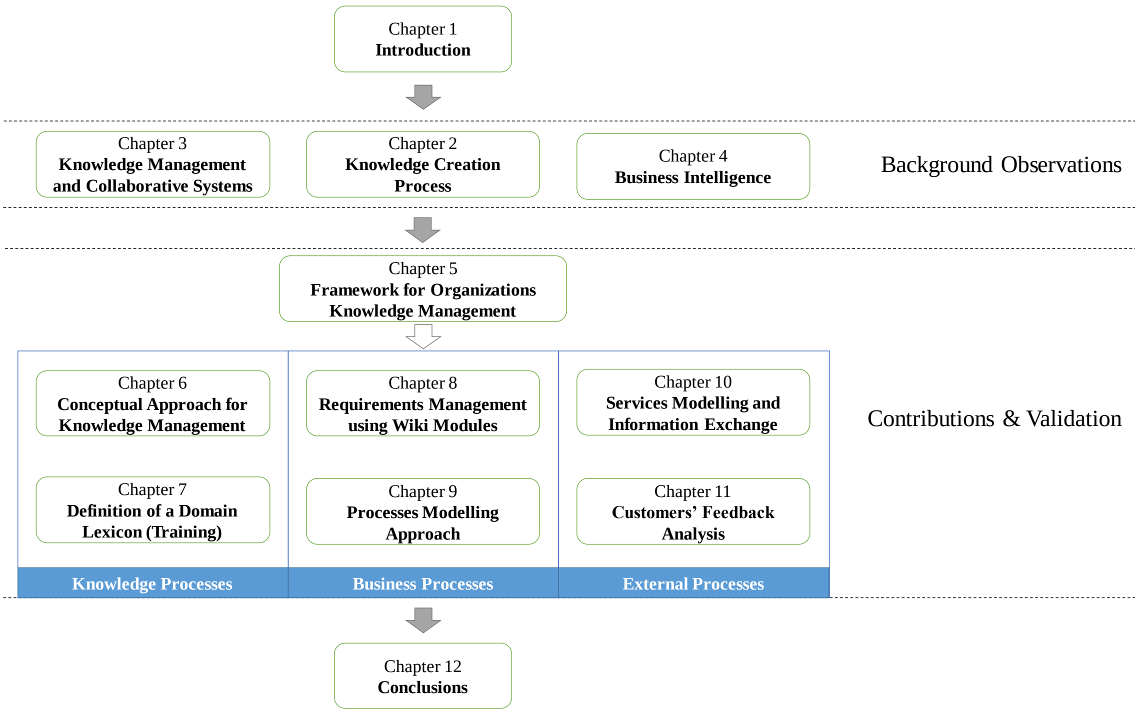


Figure 1.3: Thesis Plan Outline.

glossary, which goal is to unify the language used to represent a domain's knowledge, can be built to increase the semantic interoperability between domain participants (e.g. organizations, services, products).

1.5 Thesis Plan Outline

This dissertation is organized in four major sections that compose the overall contents like represented in Figure 1.3. It starts with the Introduction in **Section I**, with all the relevant information about the nature of the work. Then, in **Section II**, the background studies for the research work developed are conducted. **Section III** has a report on the contributions. It includes the description of the outcomes of the research and exploitation of results. Finally, **Section IV** makes the closure of the research and the document, establishing the conclusions and future work.

1.5.1 Section I - Introduction

This section is composed by Chapter 1 - Introduction - that established the guidelines of the presented research work. At first, some brief background studies are presented as part of the author's motivation and the consequent vision on this dissertation contribution.

Then, the research method is presented followed by the author research questions and hypotheses.

1.5.2 Section II - Background Observations

This section is organized in three chapters:

- **Knowledge Creation Process** - is the second chapter of this thesis. In this chapter, the author presentes the adopted model for knowledge creation process. The knowledge life cycle process and its sub-processes are also identified and explained in detail. Under the knowledge formalization and storage (section 2.3.2) several knowledge formalization degrees are identified and described.

- **Knowledge Management and Collaborative Systems** - in this chapter the KM concept is presented together with the several criteria and sub-criteria that lead to the selection of the most adequate knowledge management system in accordance with a certain organization profile. The several Knowledge management typologies identified by the author in the literature are also depicted. In this chapter, the collaboration concept is also addresses and several collaborative systems typologies are presented.

- **Business Intelligence** - Intelligent systems are present in almost every steps of knowledge management (acquisition, formalization and storage, use, and maintenance). Thus, chapter 5 is dedicated to the state of the art of the machine learning typologies algorithms (focused on data-mining). The Ontology Learning concept is also presented in this chapter together with several techniques and tool that allow ontology learning from structured, unstructured and semi-structured sources.

1.5.3 Section III - Contributions & Validation

- In the first chapter of this section - Framework for Organizations Knowledge Management - first is made a state of the are on already existing frameworks. Then, the proposed framework for organizations knowledge management is presented and three kinds if organization processes identified: Knowledge Processes; Business Processes; and External Processes. The validation chapters are organized to meet each of this organization processes Requirements.

- **Knowledge Processes**:
  - Conceptual approach for organizations knowledge management - This framework uses simple wiki-based front-end modules where tacit knowledge can

be expressed in a form of explicit knowledge directly by the different domain experts.

  - Definition of a Domain Glossary for Training Support - In this chapter, a glossary building approach is presented. Then, it will be used to support users in the search of other users to interact or learning content to access.

- **Business Processes**

  - Requirements Management Using Wiki Modules - In this chapter, a requirements management framework is proposed. It intends to allow both business and development requirements to coexist in a single environment promoting the communication between multi-disciplinary teams.
  - Processes Modelling approach - The proposed approach aims at separating the business points of view from the technical and physical means to realize it, hence promoting the cooperation and coordination between teams.

- **External Processes**

  - Services Modelling and Information Exchange - The aim of this work is to provide interoperability integration on the information exchanged between domain suppliers and a specific platform, thus providing support to organizations' external processes.
  - Customers' Feedback analysis - The proposed Framework establishes a set of components that aim to guide companies to customer's sentiment analysis. Its main characteristic is the ability of gather customer's data from web software and analyse the associated sentiments. This allow companies to make fact-based decision making to improve the services and products provided.

1.5.4 Section IV - Conclusions

In chapter 12 the main conclusions of the developed work are established. Then, some prospective on what can be done using the presented work to promote further developments are presented.

CHAPTER 2

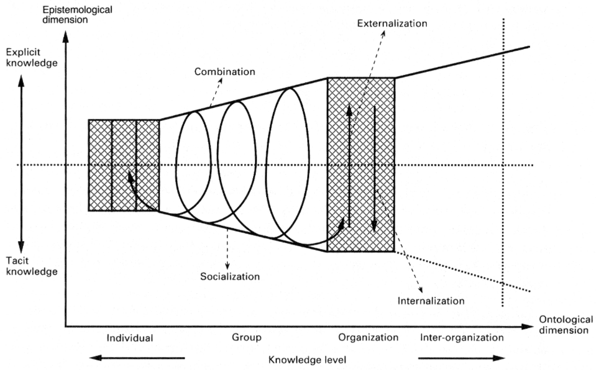
KNOWLEDGE CREATION PROCESS

The success on an organization in the twenty first century will be determined by the extent to which an organization's members can develop their intellectual capabilities through knowledge creation [34].

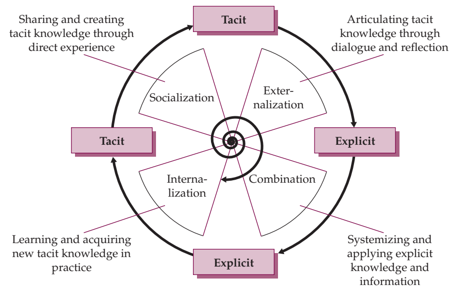
Table 2.1: Ontological dimension categories.

Individual Knowledge	Information believed by an individual as justified truth and stored in memory in a cognitive structure through a cognitive process called learning
Group Knowledge	Knowledge held by a group of individuals (e.g. organization departments)
Organizational Knowledge	Knowledge held by an organization
Inter-Organizational Knowledge	Knowledge held at an inter-organizational level (e.g. knowledge held between an organization and its suppliers)

In the pursuit of organize and describe the knowledge creation process in a company, Nonaka and Takeuchi [3] defined two dimensions of organizational knowledge creation/-conversion - epistemological and ontological. On the epistemological dimension the existence of two types of knowledge is recognized - tacit and explicit. On the other side the ontological dimension ranges from the individual and moves from the group, organization and beyond (e.g. inter-organizational) (see descriptions in Table 2.1). Given these definition, *A spiral emerges when the interaction between tacit and explicit knowledge is elevated dynamically from a lower ontological level to higher levels* (see Figure 2.1 (a) [35]). The spiral is created by four modules of knowledge conversion, in which the knowledge is converted from one type to another. The types of knowledge conversion are described by Takeuchi [36] and represented according with Figure 2.1 (b)). Moving through the spiral, the interaction between tacit and explicit is amplified. That means, the spiral becomes larger in scales as it moves up to ontological levels (Figure 2.1 (a)). The four modules of Knowledge creation/conversion represented in Figure 2.1 are:



(a) Spiral of Organizational Knowledge Creation [36].



(b) Epistemological dimension of Knowledge Creation [35].

Figure 2.1: Socialization, Externalization, Combination and Internalization Process of Knowledge Spiral.

- **Socialization (tacit to tacit)** - Process of sharing TK through observation, imitation, practice, and participation in a formal and informal community [36]. Some used methods are [37]:
  - Face to face communication;
  - Video conference tools;
  - Virtual reality tools.
- **Externalization (tacit to explicit)** - Process of articulating tacit TK into EK through dialog and reflection [35], [36]. Some of the used methods are:
  - Process capture tools;
  - Traceability;
  - Expert systems;
  - Discussion Platforms.
- **Combination (explicit to explicit)** - Process of integrating concepts into a knowledge system [36]. It implies Systematizing and applying EK and information [35]. It can be done through:
  - Systemic knowledge tools;
  - Web forums;

  - Best practice databases.

- **Internalization (explicit to tacit)** - Process of embodying EK into TK [3]. Examples are:
  - Collective knowledge;
  - Notes database.

2.1 Knowledge Creation Requirements

Although several studies focused on the knowledge creation and transference can be found in the literature, most of them target the source and state of knowledge [38] and not much effort has been paid to explore conditions that might facilitate the knowledge creation and transference. That means, even though there are many means for an organization to facilitate and support the knowledge creation process, not much have been said about the organizations role in the knowledge creation process.

Mechanisms and strategies such as organizational structural styles, organizational strategies, communication, trust, motivations, learning and training are considered factors that influence creating and sharing knowledge culture [39]. However, taking Takeuchi and Nonaka [37, 40] as the reference, the adopted conditions in fostering the knowledge creation spiral are:

- **Intention** - The level or organizational aspiration to its goal is considered the driver of the knowledge spiral. Business settings within the efforts to achieve the goals usually take the form of the strategy. From the organizational knowledge creation perspective, the essence of strategy relies in the developing of the capability to assure, create, accumulate and exploit knowledge. Thus, the most critical element of corporate strategy is to create a clear vision about the knowledge that should be developed and implement that vision in practical terms.

- **Autonomy** - It increases the motivation of individuals to create new knowledge or original ideas. By allowing individuals and groups to act autonomously the organization may increase the possibility of introducing unexpected opportunities. Autonomy gives individuals freedom to absorb new knowledge.

- **Fluctuation and creative chaos** - fluctuation (breakdown of routines, habits etc.) and 'creative chaos' increase tension and focus attention on defining problems and resolving crisis. They promote the knowledge spiral by strengthening the subjective commitment of individuals as well as stimulation of interaction with the external environment. Fluctuation and creative chaos act as a trigger for individual members to change their fundamental ways of thinking and challenge existing concepts.

- **Redundancy** - There are some ways to build redundancy into organizations such as: introduction of overlapping approaches for example when different departments work together, having strategic rotation and frequent meetings both on regular and irregular bases, or facilitating formal or informal networks - such as after office get-togethers [41].

- **Required variety** - An internal diversity organizations should match the variety and complexity of the environment. Providing equal access to information within the organization supports the exchange of different viewpoints and interpretations of new information. Organizational members can cope with many unexpected events if they have a variety of information and experience. This variety can be enhanced by combining information differently, flexibly and quickly.

2.2 Role of Managers to Promote Knowledge Creation and Transfer in organizations

People at management positions are a powerful source of influence on follower's work behaviours [42]. Subordinates' creative and innovative behaviours in the workplace are no exception. Indeed, these behaviours besides being influenced by factors like knowledge and skills [43, 44], have been discussed to be considered a motivational issue [45], which makes them of significant interest to researchers in the field of leadership.

Thus, managers are considered to play an important role in the promotion the knowledge creation and transfer in organizations. In order to sustain competitive advantage, managers' understanding of knowledge creation and transfer is vital since the success of a company might be determined by managers' intellectual capital. There are several leadership behaviors that can promote the knowledge acquisition. The ones adopted by the author go towards De Jong and Den Hartog research [46] and include innovative role modeling, intellectual stimulation, stimulating knowledge, diffusion, providing vision, consulting, delegating, support for innovation, recognition, monitoring, and task assignment.

2.3 Knowledge Life Cycle

The knowledge life cycle is the process that knowledge passes through an organization since it is identified, created, captured, shared, transferred, and utilized [47]. The author adopted the representation of knowledge life cycle through its composition into four main stages, corresponding to the internal cycle of Figure 2.2. The four main stages considered are: 1) acquisition; 2) formalization and storage; 3) use; and 4) maintenance. This stages define four quadrants, which encompass smaller steps from the data collection to the DM support and new strategies definition (change of point of care).

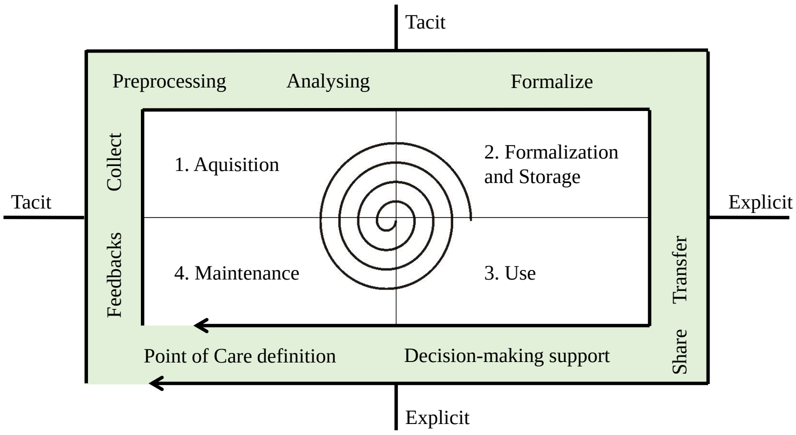


Figure 2.2: Knowledge Life Cycle.

The knowledge life cycle processes adopted can be correlated with the Nonaka and Takeuchi's Spiral of Organizational Knowledge Creation [35] represented in Figure 2.1. In this way, an entire knowledge life cycle is synchronized to the spiral conversion of knowledge cycle. The system knowledge is acquired, modeled, used and improved on its way from a smaller view (individual or group) to a wider view (inter-organizational). Moreover, for each flow from identification to capture, store, share and maintain, knowledge is converted from tacit to explicit and then to tacit again being constantly enriched with new feedback both from organizations' internal and external sources. The steps considered and illustrated in the figure will be described in detail on the following sections.

2.3.1 Knowledge Acquisition

Knowledge Acquisition (KA) is one of the main issues in KM. No doubt that KA plays an important task in building knowledge base systems. Acquisition refers to the process of getting the knowledge into the organization from external sources through using possible means [48].

2.3.1.1 Tacit Knowledge Acquisition

Extracting human tacit knowledge is one of the main challenges that most organization faces because extracting this type of knowledge is expensive and needs highly qualified people to develop some applicable methodologies which could be used successfully [49].

It needs high qualified efforts since the most of domain experts find it difficult to explain what they do or even why they do it. Also extracting a justification from domain experts about their decisions, practices, and sequence of steps in a specific task execution is complicated in some situations. This is because the nature of the tacit knowledge that is mostly composed by insights, intuitions, hunches, inherent talents, skills, experience, know-who, know-why, and working experience (see Figure 2.3) which are embedded within domain experts mind [50–55].

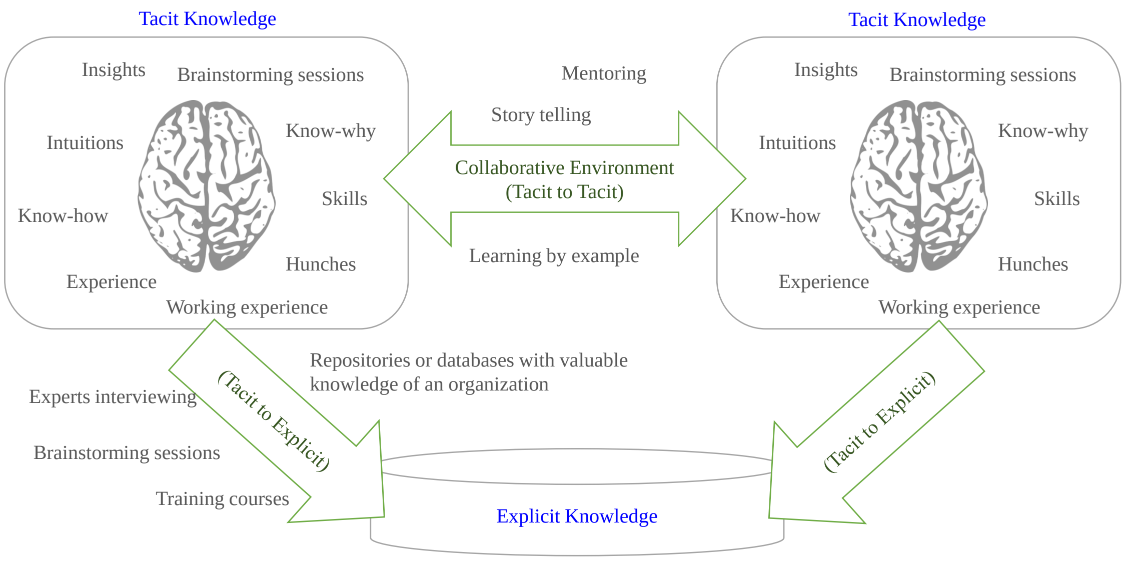


Figure 2.3: Tacit and Explicit Knowledge Acquisition.

Tacit to Tacit knowledge transference

Since the more communication, involvement, and interaction of people, more is the chance for organizations to expose tacit knowledge residing in individuals' heads the first step in tacit to tacit knowledge transference is to create a collaborative environment [56]. From here, several approaches in the literature can be found:

- **Learning by observation and learning by example:** Observing the behavior of another person, when he/she is applying its knowledge in daily situations, is a source of learning of tacit knowledge [57]. Learning by example is defined as *perceiving, reflecting and imitating existing procedures* and takes place when one person in the organization is using the example given by other to incorporate others tacit knowledge to its own knowledge. The intended effects of this type of tacit knowledge transfer are *passing on implicit routines, standardizing procedures and socialization effects* [58].

- **Mentoring:** A mentor program may enhance the availability of a leaving employees' know-how more than written reports. Usually the predecessor shows him the ropes for some time, for example for a period of three months. They are working in parallel, so that the successor gets an insight in the working method, benefiting from the advantage of asking questions right away when problems are arising [59].
- **Storytelling:** People learn more experience-based knowledge from stories than either rules of thumb or lectures. Stories provide the context together with vivid details that lodge in the mind longer than straight lecture or generalities and so enhance explicit knowledge with implicit. Socratic questioning such as *why? How do you explain? What then… ?* further engages a persons' brain being quizzed and arouses interest. Stories stir his emotions thus increasing his memories' absorption capacity [59].

Tacit to Explicit knowledge transference

There are also several approaches in the literature to explicit knowledge acquisition from domain experts' tacit knowledge:

- **Brainstorming sessions:** Brainstorming can be considered as a process for generating creative ideas and solutions through intensive and freewheeling group discussion. Every participant is encouraged to think aloud and suggest as many ideas as possible, no matter seemingly how outlandish or bizarre. Analysis, discussion, or criticism of the aired ideas is allowed only when the session is over and evaluation session begins [60]. This technique is the most widely adopted process for generate and share creative ideas within organizations.
- **Training courses:** One example is the organization of seminars to preserve, at least, some of a retiring experts know-how. One may create some sort of topically condensed report containing the most important items and "rules of thumb". Thus the younger generation/new-comers are helped better grasp the essence of available knowledge (where should I look first? Who is it worth to ask and which experts are reluctant?). Organizations can also organize training courses inside organization's working hours to enable workers to get benefit from experts. This allows organization to create a second level of experts gradually [50]
- **Experts interviewing:** Interviews allow capturing explicit and tacit knowledge when being conducted by trained interviewers [59]. They elicit knowledge, experience and personal perceptions relating to work processes and organization and its

aim is the knowledge transference to a successor worker and organizational learning.

- **Repositories or databases with valuable knowledge of an organization:** It consists in traditional forms of explicit information and knowledge repositories. By using this repositories, knowledge can be shared formally and systematically in the form of data, specifications, manuals, drawings, audio and video tapes, computer programs, and patents [38].

2.3.1.2 Databases Knowledge Acquisition

Discovering, or better to say, acquiring knowledge in databases as a discipline of business intelligence in its broader sense refers to the process of analysing large amounts of data with the aim of discovering new information and knowledge, and its application in resolving business problems. In its more specific tense, it refers to mining data as the initial step of this process, using data-driven learning algorithms, with the aim of discovering correlations between them [61].

The concept of Knowledge Discovery in Data (KDD) emerged from urgent need for a new generation of computational theories and tools to assist humans in extracting useful information (knowledge) from the rapidly growing volumes of digital data. At an abstract level, the KDD field is concerned with the development of methods and techniques for making sense of data [62]. Thus, KDD is the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [63]. Accordingly with Fayyad [62], KDD process comprises six steps, represented in Figure 2.4.

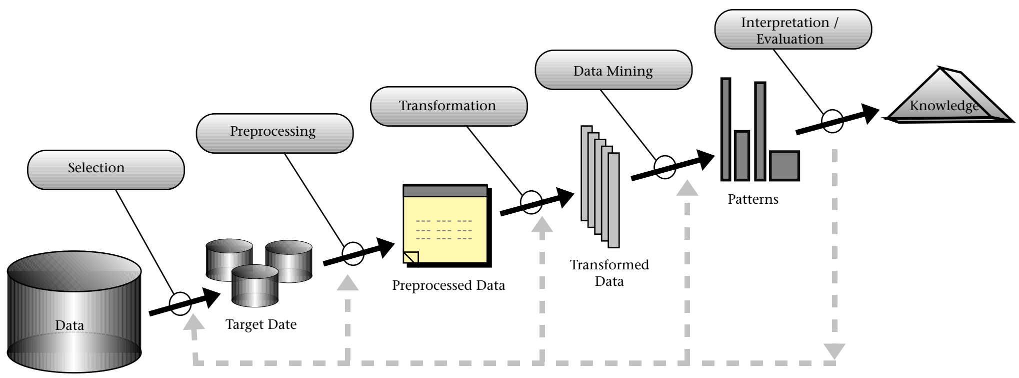


Figure 2.4: An Overview of the Steps That Compose the KDD Process (from [62]).

Based of Fayyad's steps, the external knowledge life cycle of Figure 2.2, starts at the first quadrant (acquisition), where is made the establishment of the data to be collected in order to add value to the knowledge usage (e.g. decision-making support). A formal data collection process is necessary as it ensures that data gathered is both defined and accurate and that subsequent decisions based on arguments embodied in the findings are valid [64].

In the acquisition quadrant is also made the collection, pre-processing, transformation and analysis of the data (data mining) in order to be retrieved as useful knowledge. Between others, data analysis is applied to find hidden patterns in data, evaluates significant data and interpretation of results [65]. The acquired knowledge needs, then, to be formalized and represented.

2.3.2 Knowledge Representation and Formalization Degrees

The representation and formalization of the acquired knowledge is made accordingly to a specific knowledge formalization degree and the output is formalized useful knowledge to be used (e.g. transfer, share, decision-making support). Knowledge Representation (KR) studies knowledge formalization and its processing within machines. Techniques of automated reasoning allow a computer system to draw conclusions from knowledge represented in a machine-interpretable form [66]. Accordingly, with [67], the notion of KR can be understood in terms of five distinct roles that it plays:

- **A KR is a surrogate**, that means, it enables an entity to determine consequences by thinking rather than acting, that is, by reasoning about the world rather than taking action in it;
- **A KR is a set of ontological commitments**, that means, all representations are imperfect approximations of the reality. Representing a set of things and ignoring others, we are unavoidably making a set of decisions about how and what to see in the world. Thus, we are making a commitment;
- **A KR is a fragment theory of intelligent reasoning**. This role comes about the initial conception of a representation is typically motivated by some insight indicating how people reason intelligently or by some belief about what means to reason intelligently at all;
- **A KR is a Medium for efficient computation**. It is because, reasoning in machines is a computational process. Simply explained, to use a representation, we must compute with it. As a result, questions about computational efficiency are inevitably central to the notion of KR.
- **A KR is a medium of human expression**. It is a mean by which we express things about the world, the medium of expression and communication in which we tell the machine about the world.

Due to its relevance, in the recent years the development of models and methods to knowledge formalization are studied and analysed. In Figure 2.5 are represented the

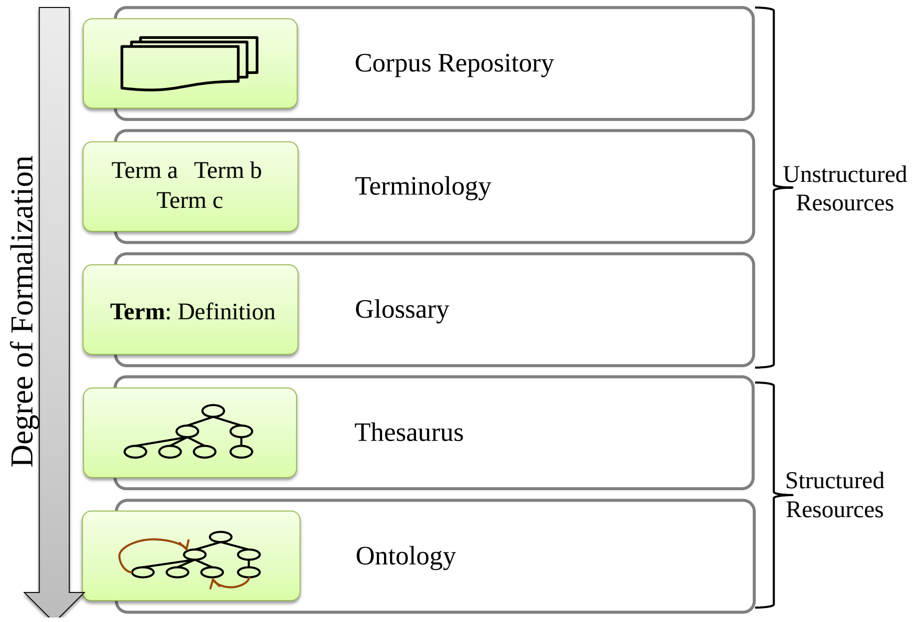


Figure 2.5: Knowledge Formalization Degrees.

adopted knowledge representation elements and its formalization degree. They range from unstructured corpus repositories to terminologies, glossaries, thesauri, taxonomies, ontologies, and logic rules, being the unstructured corpus repositories the less formalized and logic rules the ones with the most formalization degree. The author also categorized the different knowledge representation elements in **structured** and **unstructured** resources. In contrast with textual (or unstructured) resources, structured ones are used to explicitly represent domains and generic knowledge, making their inherent knowledge directly usable in applications [68].

Corpus repositories include linguistic text collections, ranging from large generic repositories such as the Web to specify domain text such as collections of texts or books on a specific subject. These repositories contain a large and ever growing amount of information expressed implicitly in natural language texts [68].

Accordingly, with the european association for terminology, **Terminology** is a set of terms representing a system concepts of a particular subject fields and the discipline dealing with it. They defend that the existence of numerous data banks or various that the use of good terminology is imperative if systems intend to functions efficiently [69]. Corpus-based methods can learn specialized terminology directly from a domain-specific corpus repository, but accuracy can be a problem because most of that repositories are relatively small [70]. Automatic term recognition (also known as term extraction) is a crucial component of many knowledge-based applications such as automatic indexing, knowledge discovery, terminology mining and monitoring, knowledge management and so on [71].

Glossary is a list of specialized terms, mostly in alphabetic order, that sometimes are unique to a specific subject. Each term is composed by its corresponding description. It includes descriptive comments and explanatory notes, such as definitions, synonyms, references, etc. [72]. Given the relevance of glossaries to better understand domains lexicon, several works were already presented using different methods and techniques. While some authors opt for rule-based systems usage to do the text-mining and glossary building [73], [74], others presented solutions based on machine learning [74], [75], and statistical techniques to simulate human consensus [76].

A **Thesaurus** is a structure that manages the complexities of terminology and provides conceptual relationships. The thesaurus can be represented by a taxonomy of domain reference terms with associated meaning. Taxonomies can be considered as a tree of categories. The organization of domain's terms in a taxonomic form allows the classification results from deeper nodes of the taxonomy to be propagated up to help classification of their ancestor. In the example given in [77], it was used a hierarchical representation of the hospitality industry terms to support customer's sentiment analysis. They considered that when some customer makes a positive or negative comment related to one of the deeper terms, the polarity of the opinion could be extended to their ancestors.

It is now frequently assumed that knowledge is modelled and stored in structures called **Ontologies** which are defined by Gruber as a formal and explicit specification of a shared conceptualization [78] and may be used as a unifying framework to facilitate knowledge sharing and interoperability between independently developed systems. Ontologies are computer implementations of human-like knowledge, for the purpose of describing domains of the world and sharing this knowledge between application programs (and also between people) [79]. Its recognised capacity to represent knowledge, to facilitate reasoning, use and exchange knowledge between systems contribute to increase the computational intelligence [80].

2.3.3 Knowledge Use

In this thesis plan, knowledge use is related to decision-making, share, and transference of the acquired and formalized knowledge. Simon at 1960 described the DM process as consisting of three phases: intelligence, design and choice [81]. Intelligence is used in the military sense to mean searching the environment for problems, that is, the need to make a decision. That means, to have context awareness. Making context based decisions will reduce the effort to coordinate tasks and resources by providing a context in which to interpret utterances and to anticipate actions [7]. Thus, knowledge provided by customers is used to achieve fact-based reasoning and inference, agility, reactiveness, innovation, context awareness, decision intelligence and competitive intelligence. Increasingly, organizations will differentiate themselves and gain sustainable advances from what they collectively known.

2.3.4 Knowledge Maintenance

There are many reasons for knowledge changes: the continual evolution of the modelled domain, the refinement of the ontological conceptualization, the modification of the application by adding features according to new end-user requirements and the reuse of the ontology for others tasks or applications. To take into account all these evolving aspects, ontologies have to be adapted to change requirements [82] in a formal dynamic Knowledge Maintenance (KMa) establishment. KMa is focused on the Knowledge Base (KB) improvement to actively be updated, monitored accordingly to the knowledge evolution of its related domain. When using the formalized knowledge, a company can conclude that the knowledge provided by the previous data analysis does not add relevant value to the DM process. Thus, new data sets must be defined to enhance further decision-making precision. This is reflected in a new point of care definition and a new data collection process. Since the useful knowledge acquirement is oriented to the findings, there is no fixed cycles of phases to obtain an optimal DM action.

2.4 Discussion

This chapter provides the basis to understand the knowledge creation process and life cycle. Under the knowledge life cycle concept, the knowledge acquisition (both from tacit and domain documents) is described. Domain experts are the ones which better know its environment, thus, a mechanism that allows to acquire their knowledge could be of great use to increase organizations products and services value and quality. The data acquisition from domain documents concept arose from the need to extract useful knowledge from the rapidly growing volumes of digital data. The capability of interpret customers' feedback from social data, marketing conditions, etc., present in domain documents, can be considered a pillar for organizations competitiveness. There are several methods and techniques to acquire knowledge from domain documents. These will be explained in detail in chapter 4. After KA, the knowledge needs to be stored in an appropriate structure. Thus, several knowledge formalization degrees are also addressed in this chapter. Finally, knowledge is ready to be used to support organizations' DM. However, it needs to be maintained in order to keep pace the domain changes. The KMa process should be validated by domain experts in order to guarantee the quality of the previously acquired knowledge (the right point of care and learning sources establishment).

CHAPTER 3

Knowledge Management and Collaborative Systems

Knowledge is referred to as the sum of information - facts, procedures, concepts, interpretations, ideas, observations and/or judgments - that human beings can process and store in their minds. However, this definition encompasses not only the knowledge contained in individual's minds, but also the information existing inside single and networked organizations [83]. It leads to the fact that, increasingly, organizations will differentiate themselves and gain sustainable advances from what they collectively known, how efficiently they use what they know and, and how quickly they acquire and use knew knowledge [84]. Thus, a deliberate and systematic approach to manage a company's KB, one populated with valid and valuable lessons learned and best practices, is required. It leads to the KM concept. Knowledge Management can be achieved through the use of an adequate Knowledge Management Systems (KMS).

3.1 Knowledge Management Systems

KMS are the IT-based systems developed to support and enhance the organizational processes of knowledge life cycle addressed in the previous chapter (creation, formalization, use, and manage) [38]. KM tools are also referred as enablers of business processes that create, store, maintain and disseminate knowledge [85]. However, some elements have to be taken into consideration during the process of selecting the most appropriate KMS [86]. A preliminary analysis is necessary to characterize the business activities that the system is intended to serve. Then, it is necessary to focus and completely understand the requirements that the developed system should fulfil. Finally, a KMS typology should be selected in such a way that the previously defined requirements are met. This decision should be done using an appropriate technique.

Although the selection of the most appropriate knowledge management systems can be supported by multiple criteria techniques [83], the Analytic Hierarchy Process (AHP) is the method that best reflects judgments based on opinion and emotion, and that best prioritize preferences for different alternatives by expressing their ranking [88, 89]. Moreover, the structure and modality of AHP ensures that all the desired specifications are included in the decision process according to the decision maker's perceptive [83]. The AHP, originally introduced by Saaty [90], is a flexible, structured technique for dealing with complex decisions. It is aimed at breaking down different choices arising within a hierarchical structure consisting of a goal, criteria, sub-criteria and alternatives [83]. It has been applied to a wide variety of decisions such as car purchasing [91], vendor selection [92], information system project selection [93], and software selection [94]. It was also already applied to support the selection of KMS [83] [87].

The Diagram presented in Figure 3.1 is based on the combination of both the works related to KMS selection works mentioned before. The decision hierarchy was structured from the left to the right, where the goal of the decision maker is placed (choice of suitable KMS), passing through the the intermediate levels (criteria and sub-criteria on which the subsequent elements depend) to the selection of the Knowledge Management System Typology. Advanced information technologies (i.e. internet, intranet, data mining) can be used to support the several parts of knowledge life cycle [95]. Thus, they should be part of the selected KMS. The following sections consist in the definition of the represented criteria (sub-criteria) and knowledge management systems typology.

3.1.1 Knowledge Management systems selection criteria

The criteria and sub-criteria have been selected on the bases of the literature review and they can be classified into four groups: cost, functionality, vendor, and stakeholders' satisfaction.

3.1.1.1 Cost

Cost is a common factor influencing the purchaser to choose the software. The available budget reserved for purchasing a KMS includes maintenance, long term operating expenses and costs for user training [96]. Technically, these costs can be grouped under two major sub-criteria: 1) Capital Expenditure - non-recurring costs (e.g. product, license and training costs); and 2) Operating Expenditure - costs involved in operating KMS (e.g. maintenance costs, software subscriptions).

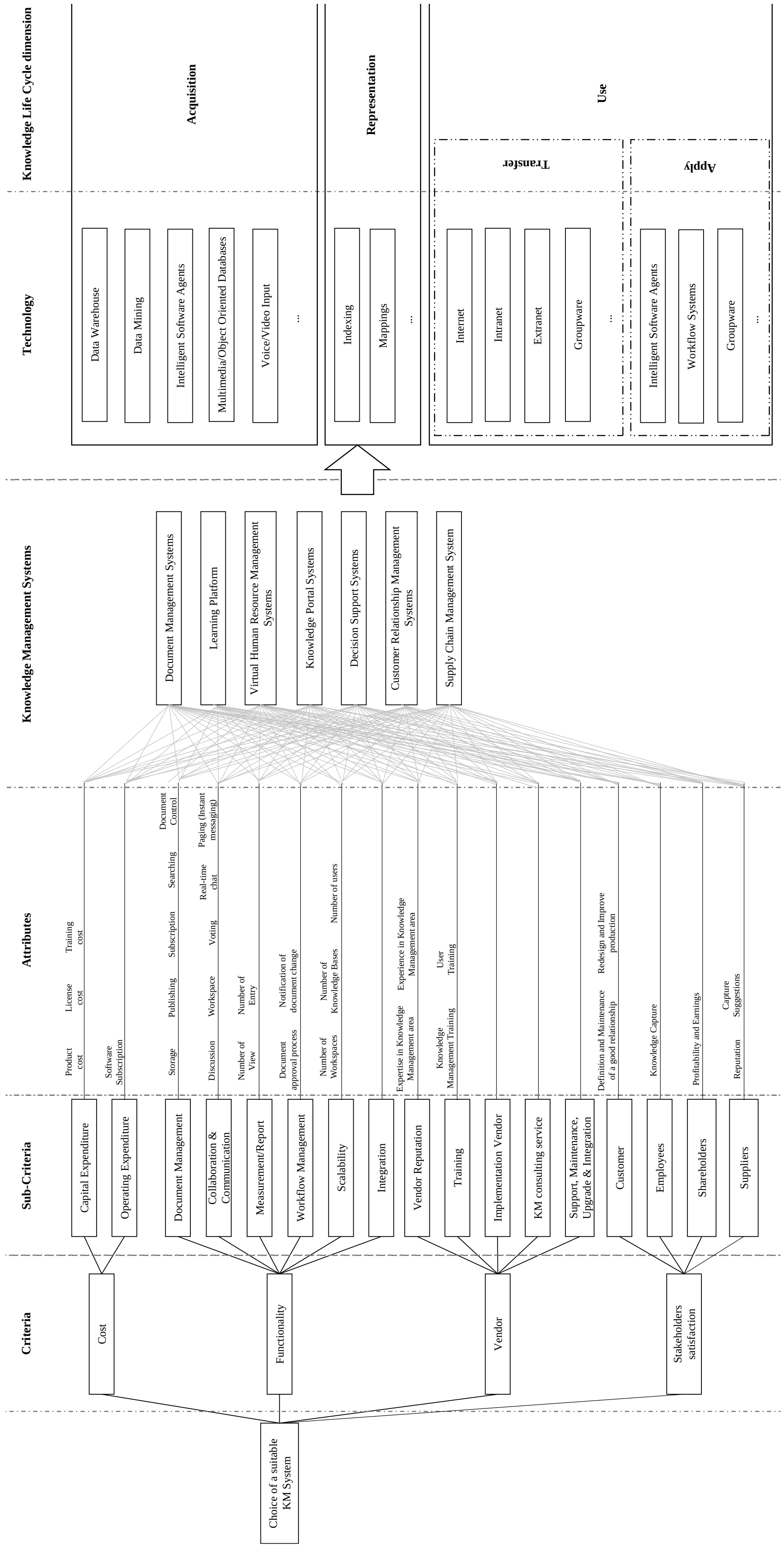


Figure 3.1: Hierarchical structure for Knowledge Management System typology selection (bases on [83] [87])

3.1.1.2 Functionality

The Functionality criteria refers to features that the KMS performs and how well the software can meet the user's needs and requirements. Thus, functionality is usually considered when selecting KMS typologies [97]. Based on the literature, the 6 key functionality of a KMS are:

- **Document Management** - In many organizations, knowledge is handled by documents [98]. Thus, it affects organizations efficiency. Document manage involves documents storage, publishing, subscription, searching and control [99].
- **Collaboration & Communication** - Within organizations, collaborating in solving problems, sharing knowledge, discussion and teamwork create a significant amount of knowledge assets [83]. In addition, it can be done through discussion, workspace sharing, voting, real-time chats and video conferences.
- **Measurement** - Is the keeping of records on activities and changes in managed knowledge. It consists, as an instance, in the knowledge of the number of views and/or entries of a product/service presented as reports.
- **Workflow Management** - allows the movement of documents in information processes among individuals and applications to be specified accordingly to a predefined process. It can be used in communicating, cooperating, coordinating solving problems and negotiation (e.g. document approval process, notification of document changes).
- **Scalability** - Refers to the ability to scale up without degradation in performance when the number of workspaces, knowledge bases and users grow [100]
- **Integration** - Ability to integrate and use different KMS giving internal and external users the means to manage an organization knowledge [101].

3.1.1.3 Vendor

The quality of KMS vendor support and its characteristics are of major importance in the selection of software [102]. It is also critical for the successful installation and maintenance of the software. The important factors affecting the decision to select a KMS are [87]:

- **Vendor reputation** - The reputation of the vendor can reflect the quality of the services and support given to the customer. Related to vendor reputation is: (a) the vendor's expertise and (b) the vendor's experience in the area of KM.

- **Training** - The availability of training courses offered by vendor companies was considered along with the actual effectiveness of these courses as a sub-criterion for KMS selection.
- **Implementation partner** - Most companies overlook the importance of finding the right implementation partner.
- **KM consulting services** - KM consulting services are important and required if a company does not have experience in KM. These services can offer help/advice to the company before, during and after the implementation of the KM tool.
- **Support, maintenance, upgrade and integration** - A vendor does not only sell a KM tool but is also responsible for the ongoing support, maintenance, upgrading and integration of the tool.

3.1.1.4 Stakeholders Satisfaction

Stakeholders satisfaction should also be considered as an important part of decisions management. This criterion can be subdivided in the following sub-criteria [83]:

- **Customers** - The definition of a good relationship with customers is one of the most important strategies for every organization. After customer opinion and suggestions, the company can redesign and improve production and sales.
- **Employees** - In the context of KM, the employees are considered the key players in creating, sharing and capturing knowledge inside the organization.
- **Shareholders** - The use of existing knowledge within an enterprise gives the company's activities added value in terms of cost reduction, time management, human resources development, new product development and the sharing of knowledge among workers.
- **Suppliers** - Reputation, service and support orientation play a vital role in selecting a KMS and are considered important factors in guiding the decision maker during the selection process of a KMS provider.

3.1.2 Knowledge Management systems typologies

In this section the most widely diffused typologies of KMS are addressed. These are: document management systems, learning platforms, virtual human resource management, knowledge portal systems, decision support systems, customer relationship management systems, and supply chain management systems.

3.1.2.1 Document Management Systems

Traditional paper-based documents have caused many related problems such as maintaining, customizing, sharing, reusing, tracking and accessing documents [103]. Recently, computer and information technologies have provided varying facilities to manage, store and retrieve documents. Thus, many Electronic Document Management Systems (EDMS) emerge in recent years, e.g. Mayan EDMS[1], Synergis software[2], MaxxVault[3], OpenKM[4], docStar[5], Master Control software package[6], and iSO-Pro[7]. EDMS has automated the routine aspects of creating and maintaining quality systems documentation bringing several benefits to organizations [103]. Some of these benefits are: [104]: 1) directly manipulate documents; 2) index and store to retrieve the documents; 3) communicate through the exchange of documents; 4) collaborate around documents; 5) model and automate the flow of documents.

From the description of ISO9000 requirements [105] and ISO10013 guidelines [106], EDMS processes fall into three functional categories: 1) document creation/release; 2) document change; and 3) document access. Each functional category is further sub-divided into several activities [107], as illustrated in Figure 3.2.

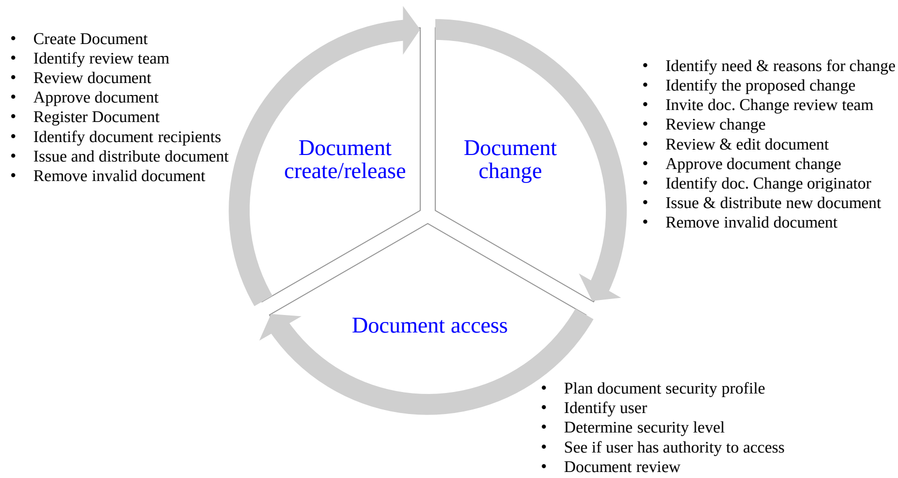


Figure 3.2: Function view of document management.

One the most challenging steps of EDMS is the retrieval of documents (information) from the previously stored ones [104]. The search algorithms are usually based in the three Boolean operators *and*, *or*, and *not* which together with a selection of search terms

[1] http://www.mayan-edms.com/
[2] http://www.synergissoftware.com/
[3] http://www.maxxvault.com/
[4] http://www.openkm.com/
[5] http://www.docstar.com/
[6] http://www.mastercontrol.com/
[7] http://isoprosoftware.com/document-control-software-module-2/

identify the documents with the required characteristics [108, 109]. More advanced systems implement automatic text analysers which produce a richer semantic representation of an input text. This representation is then used to answer questions related to the text in natural language rather than a restricted query with Boolean operators. Some of this intelligent systems are addressed latter in chapter 4.

3.1.2.2 Learning Platform

In recent years, the learning environment has changed and shifted from traditional classrooms to online learning, mobile learning and ubiquitous learning via employing various new learning devices and technologies [110]. In many countries, the internet and the web have now been a part of education for long enough that it is hard to imagine what we would do without them. Many universities offer online courses, many professors post their syllabi and conduct some discussions online, educational software proliferates, and learning management systems are commonplace [110, 111]:

- **TANGO (Tag Added learNinG Objects)** - A ubiquitous learning system proposed for tackling the issues of learning at the right time and in the right place within a ubiquitous environment [112];
- **Musex** - A collaborative learning system for supporting children in learning and exploring collaboratively in a museum with two PDAs [113];
- **PERKAM (PERsonaliz ed Knowledge Awareness Map)** - A learning system proposed to tackle the issues of seeking knowledge in a ubiquitous environment [114]. It allows the learners to seek knowledge, to share knowledge, and to exchange individual experiences among peers;
- **PPM** - A ubiquitous learning platform that supports learning indoors with RFID readers and outdoors with GPS, personal annotation management, and real-time group discussion [115];
- **PMDS** - The picture mail database system (PMDS) for improving the quality of university lectures, especially communication between the teacher and students during mass lectures [110]. It allows students to submit, search for, extract and view pictures using mobile phones during lectures.

3.1.2.3 Virtual Human Resource Management Systems

Virtual Human Resource can be considered as a network-based structure built on partnerships and typically mediated by information technologies to help the organization acquire, develop, and deploy intellectual capital. Firms compete less on products and markets and more on competencies, relationships, and new ideas. Thus, people are organization's most important asset [116]. Innovative practices that are aimed at eliciting optimal performance, and their common objectives include the following [117]:

- **Problem solving teams** - are aimed at involving production workers or other non-managerial workers in solving production problems;
- **Rotation of workers across jobs** - is used to increase worker flexibility and increase teamwork;
- **Careful screening and selection of workers** - is required to identify those who have both high-level job and task-related skills and also "team skills" to work together to solve problems;
- **Job security** - is used to assure workers that improvements in production performance will not result in the direct loss of jobs;
- **Information sharing** - is important to provide the information and motivation for greater involvement and decision making;
- **Training** - is needed to do problem solving, to increase knowledge for better decision making and to introduce workers to the skills needed for more job tasks;
- **Incentive pay** - in a wide variety of forms, is introduced to provide the incentive for greater employee effort and employee involvement in decision making.

3.1.2.4 Knowledge Portal Systems

Portals enable e-business by providing a unified application access, information management and knowledge management both within enterprises, and between enterprises and their trading partners, channel partners and customers [118]. A portal can be viewed as a way to access and disseminated information within a company since information chunks can be stored in various systems using different formats. It consists in a single point access to Internet resources, an integration platform focusing on unification oriented towards the business processes of the company. Therefore, portals synchronize knowledge and applications, creating a single view into organization's intellectual capital [119]. Knowledge Portal Systems (KPS) optimize knowledge distribution within entire organizations and can be regarded as an extension of an enterprise information portal to knowledge management [120, 121]. An extensive study regarding KPS applied to several industry sectors (e.g. automobile, microelectronics, pharmaceutical) can be found in [119].

3.1.2.5 Decision Support Systems

Decision Support Systems (DSS) are computer technology solutions that can be used to support decision-making and problem solving [122]. Current DSS facilitate a wide variety of decision tasks including information gathering and analysis, model building, sensitive analysis, collaboration, alternative evaluation, and decision implementation [123]. A state-of practice for data and model-driven DSS can be found in [123]. DDS tool design encompass several components like: 1) sophisticated database management capabilities with access to internal and external data, information and knowledge; 2) powerful modelling functions accessed by a model management system; and 3) powerful, yet simple user interface designs that enable interactive queries, reporting, and graphing functions [122].

3.1.2.6 Customer Relationship Management Systems

Customer Relationship Management (CRM) comprises a set of processes and enable systems to support a business strategy to build long term, profitable relationship with specific customers [124]. One of the most accepted definitions of CMR is the one provided by Swift: "Enterprise approach to understanding and influencing customer behaviour through meaningful communications in order to improve customer acquisition, retention, loyalty, and profitability" [125]. CRM consists of four dimensions [125–127]: 1) customer identification; 2) customer attraction; 3) customer retention; and 4) customer development. Thus, CMR often allow: 1) capture and analyse information about customer purchase behaviour; 2) long-term success through deeper and closer customer relationships; 3) precise matching of marketing offers due to detailed customers profiles analysis; and 4) track effectiveness of marketing programs [128].

3.1.2.7 Supply Chain Management System

Supply Chain Management (SCM) is the management of material and information flows both in and between facilities, such as vendors, manufacturing and assembly plants and distribution centres [129]. SCMS plays a role in creating profitability and a competitive advantage. It emphasizes the value of knowledge within the supply chain and enhances the strategic importance of efficient data, information or knowledge among members of the SCM network, such as suppliers, manufacturers, distributors and retailers. The SCM affects the knowledge receivers - designers, decision-makers and peer agents - by supporting their decisions and their future market strategies in a more efficient way [130].

3.2 Collaborative Systems

According to Schrage [131], collaboration "... *is the process of shared creation: two or more individuals with complementary skills interacting to create a shared understanding that none had previously possessed or could have come to on their own. Collaboration creates a shared meaning about a process, a product, or an event. (...) It can occur by mail, over the phone lines, and in person."*

Thus, collaborative work involves considerable explicit and tacit communication between collaborators to be successful. Individuals need to negotiate shared understanding of tasks goals, of tasks decomposition and sub-tasks allocation and of tasks/sub-tasks progress. Collaborators need to know what is currently being done and what has been done in the contexts of the task goals [132]. Several collaborative models and frameworks can be found in the literature [133–137]. In this document, going towards the adopted definition of collaboration, the model considered is the one presented in [133]. This model consists in six principal phases necessary for interaction and knowledge sharing in intelligence gathering (see Figure 3.3):

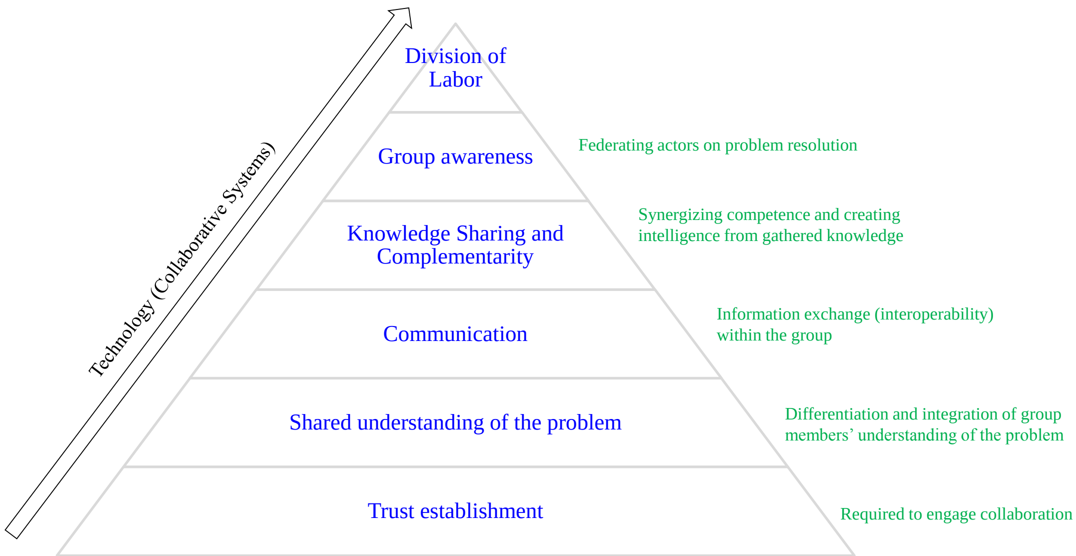


Figure 3.3: A pyramid collaborative model (based on [133]).

- **Trust phase** - For collaboration among group of individuals exist, a minimum trust among them is required. Trust in this context can be related to the beliefs about the partners and conceived ideas about the qualities that partners must possess so as to judge them credible. Collaboration cannot not take place without a minimum level of trust which enables to or more people to jointly solve a problem.

- **Shared understanding phase** - When a group of individuals decide to gather information on a particular problem, they need to define the problem itself and clarify it. What the problem represents to each member differs because of their individual interpretation. In order to engage in a collaborative way, the members of the group must differentiate and integrate their individual understanding of the problem in order to produce a shared representation of the problem.

- **Communication phase** - Communication is not an isolated phase but rather interwoven with all other phases.

- **Knowledge Sharing phase** - This phase allows group members to synergise their competences and to collectively produce actionable knowledge for decision-making. This is where intelligence is generated as users combine their innate or previously acquired knowledge with newly acquired knowledge from collected information in order to enhance clarity in decision-making. This phase allows members to share both tacit and explicit knowledge trough socialization, externalization, combination and internalization (see chapter 2).

- **Group awareness phase** - The possibility of receiving signals from one to another in a group (traditional or vistual) provides understanding of the actions and intentions of the group.

- **Division of Labour phase** - This phase allows members of a group to divide tasks among themselves in order to reduce redundancy in their activities ans to increase the rapidity of their work.

The need of technology, or collaborative systems in the adopted model increases when going from the bottom to the top of the collaborative pyramid. The worldwide web technology and electronic networks have created an environment, where place no longer matters. Collaborative systems bring geographically disperse teams together, supporting communication, coordination and collaboration. This results in a tremendous time and cost saving, greatly decreases travel requirements, faster and better DM and improved communications flow through organizations [138]. Thus, collaborative systems can be considered as a basis for groupware and KM [139]. Collaborative systems (tools) allow groups of users to communicate and cooperate on common tasks. They cover a wide range of applications such as audio/video conferencing, collaborative document sharing/editing, distance learning, work-flow management systems, etc [140], and can be classified in four categories [138] : 1) Group file and document handling; 2) Computer conferencing; 3) Electronic meeting systems; and 4) Electronic workspace.

- **Group file and document handling** - The core function of group file and document handling involve documents and files management. As described in section 3.1.2.1, it involves maintaining, customizing, sharing, reusing, tracking and access documents. Collaborative **synchronous** work on documents can also be a part of a group document handling requirement together with basic communication capabilities, such as e-mail notification and e-mail.

- **Computer conferencing** - Space for **asynchronous** and threaded discussions as well for real time discussions is provided under computer and conferencing collaborative tools category. It enables the possibility for users to see and work on documents simultaneously, on each other's screen, or whiteboard [141] and full mailing capabilities are also provided. Audio and video conferencing are quite common as well [138].

- **Electronic meeting systems** - Meetings can either be regular (same time, same place), synchronous (same time, different place), or asynchronous (different time, different place). Meeting participants have the possibility to chat, conduct real-time discussions, use audio and video conferencing facilities, write or draw in real-time on a blank slide, participate in surveys-anonymously if preferred and make group decisions, share documents and files, show and annotate PowerPoint slides, share live software applications and even work simultaneously on documents. Apart from the work-centred activities, the team also engages in team-centred activities including greeting, seeking additional participants, introduction and parting [138]. Finally, meeting-centred activities support the meeting process including its set-up, maintenance of the agenda and minutes, and distribution of the minutes after the meeting [142].

- **Electronic workspace** - The aim of this category of collaborative tools is to provide teams a common space to coordinate and organize their work. Groups can centrally store documents and files, work with them, solve problems through discussion, keep to-do lists and address books with information about group contacts, and even track projects milestones and interactions. There are workspaces for different groups, and the possibility for users to be member of several workspaces accordingly with the number of projects in which they are involved is also provided [143].

3.3 Interoperability as a Collaboration enabler

In today's economy, strategic business partnerships and outsourcing have become dominant business paradigms evidencing a tremendous increase in trade and investments between nations. Accordingly to Friedman [144], the world is becoming a *tiny flat place* with information exchanged and applied informatively across continents, independently of races, cultures, languages or systems; where mass-customization has become a major business hub replacing mass-productions; and with trends changing businesses from technology and product driven to market and customer-driven thus increasing trade and information exchange, as well as the need for **interoperability** [145, 146].

To succeed in this collaborative but at the same time competitive environment, organizations need to be interoperable, sharing technical and business information seamlessly within and across organizations [147, 148]. In this sense, being defined as the ability that two or more systems have to exchange information and use it accurately [149], interoperability, more precisely the lack of it, could disturb the creation of new markets, networks, and can diminish innovation and competitiveness of business groups.

The C^4IF, presented in [150] is a framework that uses the basic linguistic concepts of the Information Systems (IS) communication domain. In this work, four types of interoperability were defined and make the path from the physical connection between systems (organizations) and their capacity to understand the exchanged information and knowledge and use it to perform actions together (collaborate):

- **Connection** - Refers to the ability of information systems to exchange signals between them. To succeed, a physical contact/connection must be established between those (two or more) systems;

- **Communication** - Refers to the ability of data exchange in IS; To succeed, a pre-defined data format and/or schema needs to be established and accepted by the interlocutors.

- **Consolidation** - Refers to the ability of understand the exchanged data. For this, a commonly accepted meaning (domain lexicon) for the data needs to be established;

- **Collaboration** - Refers to the ability of systems to act together. A commonly accepted understanding for performing functions/services/processes/actions needs to be established.

However, for collaboration can exist, it is necessary to understand the semantics (lexicon) of the exchanged information and knowledge. The IEEE defines semantic as relationships of symbols or groups of symbols to their meanings in a given language [151]. The collaboration between systems through the understanding of the information that is being exchanged is addressed by the semantic interoperability concept. It emerged as the capability of organizations to 1) Discover required information; 2) Explicitly describe the meanings of the data they wish to share; and 3) Process received information in a manner consistent with the intended purpose of such information [152].

Thus, Semantic interoperability enables organizations to process information from external sources in a meaningful manner controlling the exchange of information and knowledge, and ensuring that it is understood and preserved through the different exchanges between domain entities (e.g. individuals and organizations).

3.4 Discussion

Knowledge is an organization important asset, thus it should be properly managed. This chapter provides the guidelines to understand how to select the most appropriate knowledge management system. The selection criterion is where an organization profile and goals are described, being of great importance to match the organization with the more adequate knowledge management system typology. Widely diffused knowledge management system typologies are also presented in this chapter. The creation of a shared context is crucial to knowledge management. Thus, the knowledge transference and management within organizations requires a supportive and **collaborative** culture and the elimination of common collaboration barriers. Between others, collaborative systems are capable to bring together geographically disperse teams to work together supporting their communication, coordination and collaboration. However, for collaboration could exist, (semantic) interoperability is required, so organization can successfully understand the exchanged knowledge and together (collaboratively) act upon it.

Chapter 4

Business Intelligence

The value of information increases with the number of users who can access that information, multiplied by the number of business areas in which the user works. Thus, organizations desperately need to timely gather relevant information and knowledge from all available sources. For this purpose, organizations are currently accumulating vast amount of data from disparate internal and external sources [153].

According to [154], Business Intelligence (BI) '*is the process of collection, treatment and diffusion of information that has an objective, the reduction on uncertainty in the making of all strategic decisions*'. Moreover, Business Intelligence can be defined as an organization's ability to gather all its capabilities and skills and transform them into knowledge [155]. Thus, it can be stated that the very nature of Business Intelligence encourages organizations to have more access to, and control over, the data.

Whether knowledge is stored in ontologies, prepositional knowledge based or simple databases, it must be maintained and kept up to date. There are many reasons for ontology changes: the continual evolution of the modelled domain, the refinement of the ontology conceptualization, the modification of the application by adding functionalities according to new end-user requirements and the reuse of the ontology for others tasks or applications. Thus, Ontology Learning (OL) techniques are introduced in section 4.4 since together with some machine learning techniques could be used to facilitate knowledge systems maintenance.

BI refers to various software solutions, including technologies such as Data warehouse, Data Mining, On-Line Analytical Processing (OLAP), Extraction Transformation Load (ETL) and other reporting applications needed to acquire the right information (or knowledge) for the business decision-making with the major purpose of enhancing the overall business performance [156]:

- **Data Warehouse** - Data warehouse is an integrated collection of summarized and historic data, which is collected from internal and external sources [157], It collects relevant data into a repository, where it is organized and validated so in can serve decision-making objectives [156]. Many organizations use data warehouses to bridge the gap of turning data into knowledge. Thus, data warehouse forms the backbone of the information supply chain to a decision support system, and consequently provides to organizations a way of turning knowledge into tangible results.

- **Data Marts** - Data marts are small sized data warehouses, tipically created by individual departments or divisions to facilitate their own decision support activities e.g. customers management, marketing, finance [158].

- **Data Mining** - Data mining is the extraction of implicit, previously unknown, and potentially useful information from data. the idea is to build computer programs that sift through databases automatically, seeking regularities or patterns [159]. Some of the more well known data modelling types for whose data mining techniques are used for are presented in section 4.2.

- **Extraction Transform Load (ETL)** - ETL is a set of actions by which data is extracted from numerous databases, applications and systems, transformed as capture, and loaded into a target database [160]. The most prominent tasks include [161]:

  1. <u>Identification</u> of relevant information at the source side;
  2. <u>Extraction</u> of information;
  3. <u>Transportation</u> of this formation to the Data Staging Area;
  4. <u>Transformation</u> (i.e. customization and integration) of the information coming from multiple sources into a common format;
  5. <u>Cleansing</u> of the resulting data set, on the basis of database and business rules;
  6. <u>Propagation</u> and loading of the data to the data warehouse and the refreshment of data marts.

- **On-Line Analytical Processing (OLAP)** - OLAP refers to the technology that enables the users to interact and present data in the Data Warehouse [162]. OLAP tools are a combination of analytical processing procedures and graphical user interface. The key features enabled by it are: multidimensional view of data, calculation intensive capabilities, and time intelligence [159].

4.1 Knowledge Management and Business Intelligence Integration

Both Business Intelligence and Knowledge Management have shown significant value in promoting decision-making from available knowledge. While Business intelligence focuses on explicit knowledge available in Data Warehouses, Knowledge Management encompasses both tacit and explicit knowledge. Since Business Intelligence and Knowledge Management concepts promote learning, decision-making, and understanding [163], in order to improve the efficiency of decision making and to adapt to changing environment and markets, the integration of Business Intelligence and Knowledge Management is needed [163, 164]. The benefits of integrating of BI with KM are 1) Ensure a real support in deploying successful business across the organization by smoothly managing multicultural teams of employees in providing highest quality products and global services to multicultural customers. 2) End-user preference and experience can be included in BI implementation, 3) provide better understanding on business context, interpretation results and training to the end-user.

Even though both of them differ in their objectives and technologies used to develop them, together BI and KM can improve the organizational performance. BI and KM integration assists today's managers for improved/optimized decision making process by sharing data and information across the organization, getting the details from internal and external sources, forecasting the future trend and taking better decision [156].

4.2 Data Mining

Each data mining technique can perform one or more of data modelling types [165]: 1) Association; 2) Classification; 3) Clustering; 4) Forecasting; 5) Regression; 6) Sequence discovery; and 7) visualization. These data modelling types will be explained in the following subsections.

Machine Learning (ML) provides the technical basis of data mining. it is used to extract information from the raw data in databases - information that is expressed in a comprehensible format and can be used for several purposes [159]. The machine learning concept will be defined in detail together with its most widely used algorithms in section 4.3.

4.2.1 Association

The idea of mining association rules originates from the analysis of market-basket data where rules like "A customer who buys products x_1 and x_2 will also buy product y with probability c%." In brief, an association rule is an expression $X \Rightarrow Y$, where X and Y are sets of items. The meaning of such rules is quite intuitive: Given a database D of transactions - where each transaction $T \in D$ is a set of items -, $X \Rightarrow Y$ expresses that whenever a transaction T contains X then T probably contains Y also as well [166]. For example, the following rule (see equation 4.1) can be extracted from the data set shown in Table 4.1 [167]. The rule suggests that exists a strong relationship between the sale of diapers and beer because many customers who buy diapers also buy beer. It can be used by retailers to identify new opportunities to cross-selling their products to the customers.

Table 4.1: Example of market basket transactions.

Id	**Items**
1	{Bread, Milk}
2	{Bread, Diapers, Beer, Eggs}
3	{Milk, Diapers, Beer, Cola}
4	{Bread, Milk, Diapers, Beer}
5	{Bread, Milk, Diapers, Cola}

$$\{Beer\} \Rightarrow \{Diapers\} \tag{4.1}$$

Association learning can be applied to Ontology learning. As an instance, in [168], the authors implemented a system that is able to ontology learning through the establishment of associations between words applied to a search engine. If two words are searched together very often, there is a strong relation between them. Thus, the next time a user looks for one of these words, the other will be part of the search output.

4.2.2 Classification

Systems that construct classifiers are one of the commonly used tools in data mining. Such systems take as input a collection of cases, each belonging to one of a small number of classes and described by its values for a fixed set of attributes, and output a classifier that can accurately predict the class to which a new case belongs. Some common application domains in which the classification problem arises, are the following [169]:

- **Customer Target Marketing** - Since the classification problem relates feature variables to target classes, this method is extremely popular for the problem of customer target marketing. In such cases, feature variables describing the customer may be used to predict their buying interests on the basis of previous training examples.

- **Medical Disease Diagnosis** - The features may be extracted from the medical records, and the class labels correspond to whether or not a patient may pick up a disease in the future.

- **Supervised Event Detection** - In many temporal scenarios, class labels may be associated with time stamps corresponding to unusual events. For example, an intrusion activity may be represented as a class label. In such cases, time-series classification methods can be very useful.

- **Multimedia Data Analysis** - It is often desirable to perform classification of large volumes of multimedia data such as photos, videos, audio or other more complex multimedia data.

- **Document Categorization and Filtering** - Many applications, such as news wire services, require the classification of large numbers of documents in real time. This application is referred to as document categorization.

- **Social Network Analysis** - Many forms of social network analysis, such as collective classification, associate labels with the underlying nodes. These are then used in order to predict the labels of other nodes. Such applications are very useful for predicting useful properties of actors in a social network.

4.2.3 Clustering

Clustering means the act of partitioning an unlabelled dataset into groups of similar objects. Each group, called a 'Cluster', consists of objects that are similar between themselves and dissimilar to objects of other groups. Thus, data items are grouped according to logical relationships, for example in consumer preferences, data can be mined to identify market segments or consumer affinities. Clustering can be performed in two different modes: crisp and fuzzy. In crisp clustering, the clusters are disjoint and non-overlapping in nature (Any pattern may belong to one and only one class). In case of fuzzy clustering, a pattern may belong to all the classes with a certain fuzzy membership grade [170]. In the following is a list of clustering and ontology learning works [171]:

- **Discovery of ontology learning resources using word-based-clustering [172]** - The author proposed a data mining process using machine clustering mining approach for learning ontologies from resource repositories in XMLs or RDFs. The generated ontologies provided a controlled vocabulary of concepts which contribute to the sense disambiguation in seeking interesting and appropriate knowledge in the domain of education e-markets.

- **Using ontology in hierarchical information clustering [173]** - The author proposed an approach of using hierarchical clustering system modified for ontological indexing and topic-centric test collection of documents. The authors state that ontologies can impose a complete interpretation of subjective clustering onto a document set.

- **Modeling user interests by conceptual clustering [174]** - In this work is presented a document clustering algorithm, named WebDCC (Web Document Conceptual Clustering), to carry out incremental, unsupervised concept learning. Afterward, the semantics extracted from Web pages can be integrated into ontology.

- **Learning expressive ontologies [175]** - The authors presented two conceptual approaches for learning expressive ontologies: 1) lexical approach to generate complex class descriptions from definitory sentences; and 2) logical approach to generate general purpose ontology such as disjointness axioms.

4.2.4 Forecasting

In today's economic environment there is ample opportunity to leverage the numerous sources of time series data that are readily available to the savvy decision maker. This time series data can be used for business gain if the data is converted first to information and then to knowledge - knowing what to make when for whom, knowing when resource costs (raw material, logistics, labour, and so on) are changing or what the drivers of demand are and when they will be changing. All this knowledge leads to advantages to the bottom line for the decision maker when times series trends are captured in an appropriate mathematical form. Thus, there is a significant value in the interdisciplinary notion of data mining for forecasting [176].

4.2.5 Regression

In the example provided in Figure 4.1, the horizontal line is called the X axis and the vertical line the Y axis. Regression analysis looks for a relationship between the X variable and the Y variable (graph line). Based on the established relationship, the algorithm is capable to use existing values to predict what other values will be. Forecast uses the regression line assuming that the relation which existed in the past between two variables will continue to exist in the future [177]. In the provided example, the value of Y_1 based on the given value of X_1.

4.2.6 Sequence discovery

Sequence discovery is the identification of associations or patterns over time [178, 179]. Its goal is to model the states of the process generating the sequence or to extract and report deviation and trends over time [179].

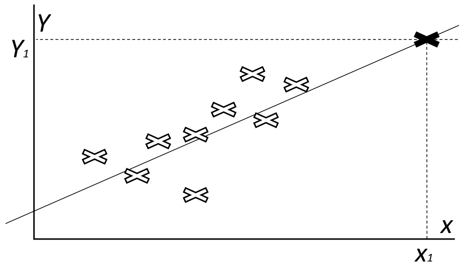


Figure 4.1: Regression applied to Forecasting.

4.2.7 Visualization

Visualization refers to the presentation of data so that users can view complex patterns [180]. It is used in conjunction with other data mining models to provide a clearer understanding of the discovered patterns or relationships [181]. Examples of visualization model are 3D graphs, "Hygraphs" and "SeeNet" [180].

4.3 Machine Learning

Crucial systems to understand are those involved in memory, but in addition, learning mechanisms are at the heart of how the brain processed information. Patterson [182] states that is by modifying the synaptic connection strengths (or weights) between neurons that useful neuronal information processors for most brain functions, including perception, emotion, motivation, and motor function, are built. One example is the study made by Patterson in [183]. This study used basic emotions as a facilitator for learning. Thus, emotions were defined in this case, due to the animal's use, as states elicited by rewards and punishments. A reward is anything for which an animal will work. A punisher is anything an animal will work to escape or avoid. Rewards and punishments can be more formally defined as instrumental reinforces, i.e. stimuli or events which, if their occurrence, termination, or omission is made contingent upon the making of a response, alter the probability of the future emission of that response [184].

The way the brain works have been inspired several intelligent systems development. Currently, it is widely accepted that systems that possess knowledge and are capable of decision making and reasoning are regarded as 'intelligent' [185]. There are recognized techniques, such as fuzzy logic, artificial neural networks, machine learning and evolutionary algorithms that contribute to increase a system's *machine intelligence quotient* [186]. The rationale behind the intelligent label of those techniques is their ability to represent and deal with knowledge [187]. In the following some of the more well-known of these techniques will be addressed.

Learning process include the acquisition of new declarative knowledge, the development of motor and cognitive skills through instruction or practice, the organization of new knowledge into general, effective representations, and the discovery of new facts and theories through observation and experimentation. This has been the most challenging and fascinating goal in artificial intelligence. The study and computer modeling of learning process in their multiple manifestations constitutes the subject matter of machine learning [188].

Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more effectively the next time [189]. Thus, ML paradigm can be viewed as "programming by example", learning to do better in the future, based on what was experienced in the past [190]. There are countless applications of ML . Accordingly, with Isabelle Guyon [191], typical applications of machine learning include:

- **Retail :** It is aligned with the interest of identifying prospective customers, dissatisfied customers, good customers, bad payers, etc.;
- **Biomedical and Biometric :** in order to identify people of risk to possess certain disease, to make a diagnosis or to predict the outcome of a treatments (prognosis);
- **Security :** face recognition, iris verification, fingerprint, signature and also DNA validation.
- **Computation and Internet:** computer interfaces designing (troubleshooting wizards, handwriting and speech recognition interfaces, and brain waves (handle computer machines through the brain))
- **Internet applications :** Hit ranking, spam filtering, text categorization, text translation and recommendation services.

Some of the most widely well-known algorithms to machine learning will be explained in the following sections.

4.3.1 Artificial Neural Networks

An Artificial Neural Networks (ANN) is an information-processing paradigm that is inspired by the way biological nervous system, such as brain, process information. Neural networks can be useful learning from existing data even when humans find it difficult to identify rules. Such as humans, ANN learn from experience and are able to adapt the KB when facing new data. Focus on the application of ANN to ontology learning, one application to consider is [192] where is proposed a method consisting of projective adaptive resonance theory neural network and Bayesian network probability theorem to automatically construct ontology. One problem related to ANN is if the neural network is implemented as a 'black box', then any information 'learned' by the network during the training is unavailable. Previous researchers, such as [193, 194] developed design techniques that allow network operation to be decoded after training. This researches made possible the automatic learning and adaptability of ANN with user's feedback related to the information learned.

4.3.2 Fuzzy Logic

In order to lead with uncertainty knowledge one solution is the application of Fuzzy Logic (FL) in OL. FL is a multivalued logic able to absorb vague information, usually described in natural language, and convert it into a numerical format for easy computational manipulation, searching for shaping or emulate the human reasoning. In [195] is presented a fuzzy temporal model integrated with an ontology model to allow annotating ontology definitions with time specifications. Another successful application of FL to emulate human behaviour is its usage to measure knowledge sharing, namely the confidence and knowledge complexity level [196].

4.3.3 Decision Tree

A decision tree compactly encoded a sequence of tests on the values of certain features. Each inner node corresponds to a feature; the edges represent decisions for one of the feature's possible values. A leaf represents the predicated value of the target variable. A decision tree is constructed in an iterative way. In each step, the learning algorithm chooses one feature and created a new branch for each of the possible values. At each branch, one of the remaining features is chosen. Thus, the hypothesis space is subsequently divided [197]. The decision trees have several advantages comparing with others machine learning methods, as [198]: 1) Easy to understand and interpret; 2) Necessary little data preparation; 3) Can interpret numerical and categorical data; 4) Can be interpreted by a white box model; 5) It is possible to validate the model by statistical tests; 6) It is robust; and their performance when using large datasets is good.

Some works have emerged using decision trees to knowledge learning. In [199] is proposed a generic interactive procedure, relying on ontology to model qualitative knowledge and on decision trees as a data-driven learning methods. In [200] the problem of inductive learning using ontologies and data is formalized. They describe an ontology-driven decision tree learning algorithm to learn classification rules at multiple levels of abstraction. Finally, in [201], a concept learning framework for terminology representation is introduced. It is grounded on a method for inducing logic decision trees as an adaptation of the classic tree induction methods to the description logics representations adopted in the Semantic Web context.

4.3.4 Bayesian Learner

Statistical modelling is often also called Bayesian learning because it is based on Bayes Theorem [1]. Bayesian learners are probabilistic models making the simplifying assumption that all features are independent and have the same relevance. However, despite their simple design, Bayesian learners outperform many more complex approaches when applied to real-world problems [197].

The Bayesian learner estimate hypothesis based on their data, then the hypothesis is used for prediction purpose. Then is used the posterior probabilities of the hypothesis to weight the predictions [202]. Assuming that X denotes an unlabelled example. let it have only the feature *color* with the value *red*. Now the classification into the categories *apple* or *banana* is intended. H is the hypothesis that this example belongs to a certain category C, for example that it is an *apple*. The conditional probability $P(H|X)$ is the probability that our example is an apple given that is red. With Bayes Theorem we can calculate this probability if we know the probability that an example is an apple $P(H)$, that an example has the colour red $P(X)$, and that an example has the colour red if it is an apple $P(X|H)$ [197].

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \tag{4.2}$$

In [203] is presented an effort on developing a principled methodology for automatic ontology mapping based on BaysOWL, a probabilistic framework developed by the authors for modelling uncertainty in semantic web, in [204] is proposed a methodology to extract concept relations from unstructured text using a syntactic and semantic probability-based Naive Bayes classifier, finally, in [205] the improvements related to the introduction of ontologies formalism in the e-learning field are discussed, and a novel algorithm for ontology building through the use of Bayesian networks is shown.

4.3.5 K-Nearest Neighbor Learning

The K-Nearest Neighbors (KNN) has been broadly used in ML applications due to its conceptual simplicity, and general applicability [206]. A KNN classifier is trained by storing all training patterns presented to it. During the test stage, the K stored entity pairs closest to the test entity pair are found using a distance function. A vote is then taken

[1] https://en.wikipedia.org/wiki/Bayes%27_theorem

amongst those K neighbours, and the most frequent class is assigned to that test entity pair. This assignment minimizes the probability of the test entity pair in question being wrongly classified [207].

In [208], the authors present a method to combine similarity measures of different categories without having ontology instances or any user feedback in regard with alignment of two given ontologies. For that, KNN is one of the classifiers investigated. In [209], a KNN search procedure is presented, for retrieving resources in knowledge bases expressed in OWL. The procedure exploits a semi distance for annotated resources, that is based on a number of dimensions corresponding to a committee of features represented by OWL concept descriptions. This procedure can retrieve resources belonging to query concepts expressed in OWL.

4.3.6 Markov Chains

In 1907, Markov began the study of an important new type of chance process. In this process, the outcome of a given experiment can affect the outcome of the next experiment. This type of process is called a Markov chain [210].

In [211] the authors propose a new model to automatically estimate weights on concepts within the ontology. This model initially maps the ontology to a Markov chain. Further, the transition probability matrix is used to compute the probability of steady stages based on the left eigenvectors. Finally, the importance is calculated for each ontology concept. Another interesting application of Markov networks is about to build a probabilistic scheme for ontology matching, one of this cases is iMatch, where a Markov network is constructed on the fly according with two input ontologies [212].

4.3.7 Genetic algorithm

There is no rigorous definition of Genetic Algorithms (GA) accepted by all the evolutionary-computation community. However, it can be said that all the GA have at least the following elements in common: 1) population of chromosomes; 2) selection according to fitness; 3) crossover to produce new offspring; and 4) random mutation of new offspring [213].

In computer world, genetic material is replaced by strings of bits and natural selections replaced by fitness function. Matting of parents is represented by crossover and mutation operations [213, 214].

- **Selection** - This operator selects chromosomes in the population for re-production. The fitter the chromosome, the more times it is likely to be selected to reproduce.
- **Crossover** - This operator randomly chooses a locus and exchanges the sub-sequences before and after that locus between two chromosomes to create two offspring.

For example, the strings 10000100 and 11111111 could be crosses over after the **third** locus in each to produce two offspring: 10011111 and 11100100. It is the crossover operation intention to mimics biological recombination between two single-chromosome organisms.

- **Mutation** - This operation randomly flips some of the bits in a chromosome. For example, the string 00000100 might be mutated in its second position to yield 01000100. Mutation can occur at each bit position in a string with some probability (usually very small).

Some of the applications of GA are: image enhancement and segmentation, software testing, parameter and system identification, control, robotics, pattern recognition, engineering designs, speech recognition, planning and scheduling, and classifier systems[214–216]. More focused in this thesis plan context, GA have been applied to knowledge discovery [217], financial forecasting (the algorithm for Forecasting is introduced to predict the values of financial statement variables. The mutation operation is guided by domain knowledge to make small or large changes in an organism) [218], and decision support [219].

4.4 Ontology Learning

Ontologies constitute a powerful tool to support natural language processing [220], information filtering [221] and information retrieval [222]. However, the manual construction of ontologies is an expensive and time consuming task because the professionals required (i.e. a domain specialist and a knowledge engineer) usually are highly specialized [223]. It becomes even more critique with the Semantic Web dream [224] and the explosion of information due to the Read/Write Web. Thus, a systematic body of study in large-scale extraction and representation of facts and patterns necessity became obvious [225].

Learning semantic resources from text instead of manually creating them might be dangerous in terms of correctness, but has undeniable advantages. Further, the cost per entry is greatly reduced, giving rise to much larger resources than an advocate of manual approach could ever afford [226]. That realization give rise to the research area known as OL. It is concerned with knowledge discovery in different data sources and with its representation through an ontologic structure and is a powerful approach for automating the knowledge acquisition process [223].

Ontology learning can also be defined as the set of methods and techniques used for building, semi-automatically or automatically, ontology from scratch, enriching, or adapting an existing ontology using several sources [227]. Compared with manually crafting ontologies, ontology learning is able to not only discover ontological knowledge at a large scale and faster pace, but also mitigate human-introduced biases and inconsistencies [228]. Benz [229] categorized OL based on its three types of inputs: 1) Structured data: database schemes; 2) Semi structured data: dictionaries; and 3) Unstructured data: natural language text documents, like the majority of the HTML based webpages. These categories defined the structure of this chapter structure.

4.4.1 Ontology learning from structured data

Ontology learning from structured data approaches extract parts of ontologies using the available structured data. Examples of structured data sources are database schemas, and existing ontologies & knowledge bases.

4.4.1.1 Databases Schema

Structured data sources like databases present several advantages against text, since they can be presented as simple domain models. Relational databases are a large percentage of existing structured data sources, and have been often used by companies and organizations to support their operational activities. These databases are centred on the notion of relation, and are composed by a set of tables, one for each relation in the database. Since, rules for creating a relational database from an E-R (entity-relation) model are almost standard [230], several methods and approaches can be found in the literature. The comparative study of Table 4.2 was adapted of the study presented in [227] and presents several approaches for OL from database schema.

4.4.1.2 Ontologies and Knowledge Bases

In this section, several successful OL from existing ontologies and knowledge bases works are presented. These works aim at refine (in accordance with users' configuration), merge and align existing ontologies to build one to serve specific purpose. As an instance, in [236] an approach for learning ontology from RDF annotations of Web resources was proposed. A universal similarity paradigm reflecting the implicit coherence among the ontologies is presented. Ontology alignment and ontology construction methods are used. The output follows users' configuration such as their preferred structured and filtering threshold. Other approach [237], consists in searching for online ontologies to represent a set of concepts. It ranks the retrieved ontologies accordingly with some criteria, then extracts relevant parts of the top ontologies, and merge them to acquire a richest domain representation. Finally, in [238, 239], the authors aim to elicit an ontology from a knowledge base of rules. Given a knowledge based system built with ripple down rules (a tree

Table 4.2: Comparative study between several OL approaches from database schemas (adapted from [227]).

Name	Main goal	Main Technique	Reuse of other ontologies	sources used for learning	Tool associated	Evaluation	Source
Johannesson's method	To map a relational schema with a conceptual schema	Mappings Techniques	No	Relational schemas	Information not available in source	User	[231]
Kashyap's method	To create and refine an ontology	Mappings Techniques and Reverse Engineering	Yes	Schemas of domain specific databases	EDEN	User	[232]
Rubin and colleagues' approach	To create ontological instances	Mappings Techniques	Yes	Relational schemas of a database	Information not available in source	User	[233]
Stojanovic and colleagues' approach	To create ontological instances from a database	Mappings Techniques and Reverse Engineering	No	Schemas of domain databases	OntoLift (KAON)	User	[234]
Man Li and colleagues' approach	To create ontology from database	Mappings Techniques	No	Relational schemas	Information not available in source	Information not available in source	[235]
PaM4OL	To create domain ontology from database	Mappings Techniques and Reverse Engineering	No	Schemas of domain databases	Information not available in source	User	[230]

structure where the nodes are rules), the authors propose an algorithm to extract the class taxonomy where a class is a set of different rule paths giving the same conclusion, and a rule path for node *n* consists of all conditions from all predecessors' rules plus conditions of the particular rule of node *n*. The experimental results are based on a large real-world medical ripple down rules knowledge based system [227].

4.4.2 Ontology learning from semi-structured data

Although the results of ontology learning procedures using structured data usually show better results, most of the available knowledge is in the form of semi-structured and unstructured text. For this reason, techniques for learning ontologies from semi-structured data (e.g. data composed by structural information plus free text) have been emerged.

4.4.2.1 Dictionary

Different methods and tools have been developed to build ontologies using machine readable dictionary (MRD). In Table 4.3 is presented a comparative study based on [227] in which is presented the main goal, the used technique, the input sources, the associated tools and evaluation method. Examples of semi-structured data are dictionaries like WordNet [245] or the Wiktionary.

Table 4.3: Comparative study between several OL approaches from dictionaries (adapted from [227]).

Name	Main goal	Main Technique	Reuse of other ontologies	sources used for learning	Tool associated	Evaluation	Source
Heart's method	To create a thesaurus, and also to enrich WordNet with new lexical syntactic relation	Linguistic patterns	WordNet	Text WordNet	Information not available in source	Comparing the relations discovered with WordNet	[240]
Rigau and collegues' method	To develop semi-automatically WordNet versions	Word-sense disambiguation Analysis dictionary definition	WordNet	MRD: monolingual and bilingual	SEISD	WSD procedures	[241]
Jannink and Wiederhold's approach	To build a taxonomy	Statistical approach PageRank Algorithm	No	MRD	Information not available in the source	Expert comparing the output with WordNet	[242]
Bozzato and collegues' approach	Construction of a domain ontology and a complete domain terminology	Clustering, Saturation, Relationship Identification	No	MRD	Information not available in the source	Information not available in source	[243]
Kietz and collegues's approach	Semi-automatic ontology acquisition from a corporate intranet	Usage of lexical and semantic annotations	GermaNet	MRD, GermanNet, NL texts	OntoEdit	Standard measures such as precision and recall	[244]

Table 4.4: Comparative study between several OL approaches from Document type definitions.

Name	Main goal	Main Technique	sources used for learning	Tool associated	Evaluation	Source
Karoui's and colleagues method	build a tourism ontology from related Web pages	Aussenac-Gilles methodology [246] and clustering techniques	Web pages structure (HTML script) related to the tourism	Information not available in source	Comparing the relations discovered with the ones manually defined	[247]
Davulcu's and colleagues method	To create a hierarchical semantic structure (taxonomy)	Frequency-based mining and tree mining algorithms	Web pages structure (HTML script)	Information not available in source	User	[248]
Hazman's and colleagues method	To build ontology	N-GRAM	structure of phrases appearing in the documents HTML headings and hierarchical structure the HTML headings	Information not available in source	Gold standard evaluation [2]	[249]

4.4.2.2 Document type definitions

OL from Document type definitions use the structure of documents (web documents) as source for learning. Some works to mention are the ones presented in Table 4.4

Table 4.5: Ontology Learning from text (adapted from [227]).

Name	Main goal	Main Technique	Reuse of other ontologies	sources used for learning	Tool associated	Evalua-tion	Source
Aussenac-Gilles and colleagues' approach	To learn concepts and relations among them	Linguistic analysis and Clustering techniques	Yes	Domain Text and Domain ontologiest	GEDITERM and TERMINAE	User	[250]
Bachi-mont's method	To build a taxonomy	NLP techniques	No	Domain text	DOE	Expert	[251]
Hwang's method	To elicit a taxonomy	NLP techniques and ML techniques and Statistical approach	No	Domain text	Information not available in papers	Expert	[252]
Khan and Luo's method	To learn concepts	Clustering techniques and Statistical approach	Yes	Domain text and	Information not available in papers	Expert	[253]
Kietz and colleagues' method	To learn concepts and relations among them to enrich an existing ontology	NLP and Statistical approach	Yes	Domain and nonspecific domain Text, Domain ontologies, and WordNet	Text-To-Onto	User	[244]
Missikoff and colleagues' method	To build taxonomies and to fuse with an existing ontology	NLP and Statistical approach and ML technique	Yes	Domain text and Wordnet	OntoLearn	Expert	[254]
Xu and colleagues' approach	To learn concept and relations among them	NLP techniques, Statistical approach, and Text-mining techniques	Yes	Annotated text corpus WordNet	TFIDF	Expert	[255]

4.4.3 Ontology learning from unstructured data

There is a great focus on acquisition of ontologies from unstructured text, a format that scores highest on availability but lowest on accessibility [226]. OL from texts is a specific case of OL and has been widely used in the community of engineering knowledge since texts are semantically richer than the other data source type [261]. These approaches are generally based on the use of textual corpora of the domain which will be used to build

[2] Gold standard evaluation is carried out by comparing the learnt ontology to a predefined ontology (golden standard) that is usually built manually from scratch by domain experts

Table 4.6: Ontology Enrichment from text (adapted from [227]).

Name	Main goal	Main Technique	Reuse of other ontologies	sources used for learning	Tool associated	Evalua-tion	Source
Aguirre and colleagues' method	To enrich concepts in existing ontologies	Statistical approach, Clustering, and Topic signatures	Yes	Domain Text and WordNet	Information not available in papers	User	[256]
Alfonseca and Man-andhar's method	To enrich an existing ontology with new concepts	Topic signatures and Semantic distance	Yes	Domain text and WordNet	Welkin	Expert	[257]
Faatz and Steinmetz approach	To enrich an existing ontology with new concepts	Statistical approach and Semantic distance	Yes	Domain corpus Domain ontology	Any ontology workbench	Expert	[258]
Lonsdale and colleagues' method	To discover new relationships in an existing ontology	NLP, Mappings, and Linguistic technique	Yes	Terminologi-cal databases, Domain ontology, WordNet, and Domain text	Information not available in papers	User/Ex-pert	[259]
Roux and colleagues' approach	To enrich a taxonomy with new concepts	Verb-patterns	Yes	Domain text Domain ontology	Information not available in papers	Expert	[260]

ontology (see Table 4.5). By applying a set of text mining techniques, granular ontologies can also are enriched with discovered concepts and relation from textual resources (see Table 4.6). Ontology learning from text is the process of identifying terms, concepts, relations, and optionally, axioms from textual information and using them to construct and maintain an ontology. Techniques from established fields, such as information retrieval, data mining, and natural language processing, have been fundamental in the development of ontology learning systems [225] (see chapter 4). OL from text techniques can be divided in constructing ontologies from scratch and extending existent ontologies [226]. Thus, the results presented in tables 4.5 and 4.6, based on the OntoWeb survey of ontology learning methods and techniques, [227] are categorized in accordance with its purpose.

4.5 Discussion

In this chapter, the Business Intelligence concept is presented. The importance of its integration with Knowledge Management and the advantages enabled are also explained. Then, some widely known machine learning techniques are depicted. Machine learning techniques are used to support several data-mining applications like: association, classification, clustering, forecasting, sequence discovery, and visualization. Since ontologies are computer implementations of human-like knowledge, for the purpose of describing domains of the world and sharing this knowledge between application programs (and also between people) [79], ontology learning emerged as a key topic to accomplish this thesis purpose. OL is concerned with knowledge discovery in different data sources and with its representation through an ontological structure and is a powerful approach for automating the knowledge acquisition process [223]. Consequently, in this chapter, several approaches to enrich or build ontologies from scratch, both from structured and unstructured sources were presented. Their goals range from build concepts taxonomies to build full ontologies trough the identification of several relations between concepts.

CHAPTER 5

Framework for Organizations Knowledge Management

This chapter intends to present the author's main conceptual contribution, which is a framework for organizations knowledge management. But before the framework presentation, a background study about previous related work that supported the presented one is done within the following sections.

5.1 Frameworks and Models for Knowledge Management

In the knowledge management domain, frameworks, (reference) models, and architectures are widely used to describe components, design aspects or technical architectures and their interdependence [262–264]. In this context, knowledge management frameworks are created to achieve a common understanding of the domain [262, 265, 266], to structure approaches and practices [267] and to identify research gaps [38, 267].

Heisig [264] analysed around 160 frameworks to identify the success factors and most important components. As a result of his work, the aspects identified as critical success factors were: (1) human-oriented factors (culture, people, leadership), (2) organization (processes and structures), (3) technology (infrastructure and applications) and (4) management (strategy, goals and measurement). Within this section, some of the widely known knowledge management frameworks and models were studied.

The first framework considered consists in the European perspective for Knowledge Management (see Figure 5.1) [262]. It is better explained to the definition of its three core layers:

- **Business focus** - It should be in the centre of any knowledge management initiative and represents the value-adding processes of an organization, which may, typically include strategy development, product/service innovation and development, manufacturing and service delivery, sales and customers support. These processes represent the organizational contexts in which critical knowledge is created and applied.

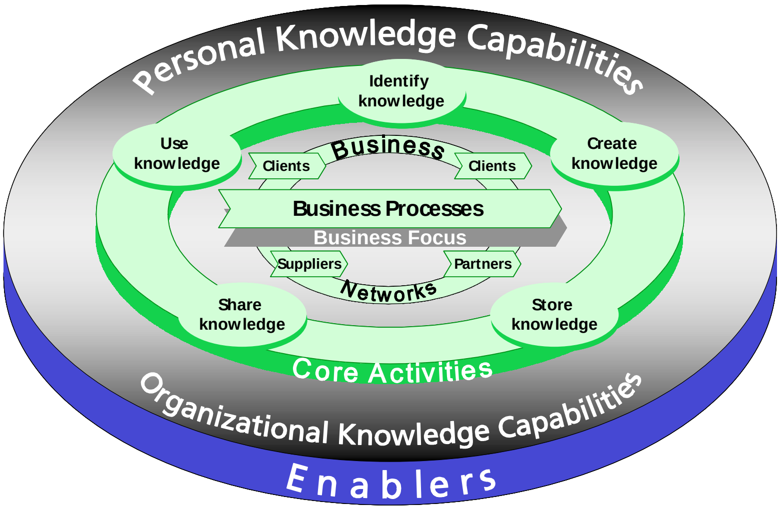


Figure 5.1: Framework for Knowledge Management: An European Perspective.

- The **five core knowledge activities** - have been identified as most widely used by organizations in Europe: identity; create; store; share; and use and consist in the second layer of the framework. These activities are typically performed to support business processes. Their integration and performance within an organization have to be supported by the right KM methods and tools (see chapter 3).

- The **enablers** are the third layer and comprise two main categories:

  – Personal Knowledge - Includes capabilities such as ambition, skills, behaviour, experience, tools and time management which have to be developed at the personal and group level in order to generate improvements from knowledge handling.

  – Organizational Knowledge - capabilities are those that leaders have to establish in order to facilitate knowledge handling within the value-adding processes, by both internal stakeholders (such as managers and employees) and external partners (such as suppliers and clients). These capabilities include:

mission; vision and strategy; the design of processes and organizational structures; understating of the cultures.

Maier's framework [268] is organized in three different levels (strategic, design, organizational) and by knowledge types which are connected by generic knowledge activities (see Figure 5.2). The architecture identifies key aspects of knowledge management as well as potential tools and methods around those (e.g., ontologies, technical architectures, or roles). It is based on clear, research-based classifications and categorizations and identifies influence factors and solutions for different purposes.

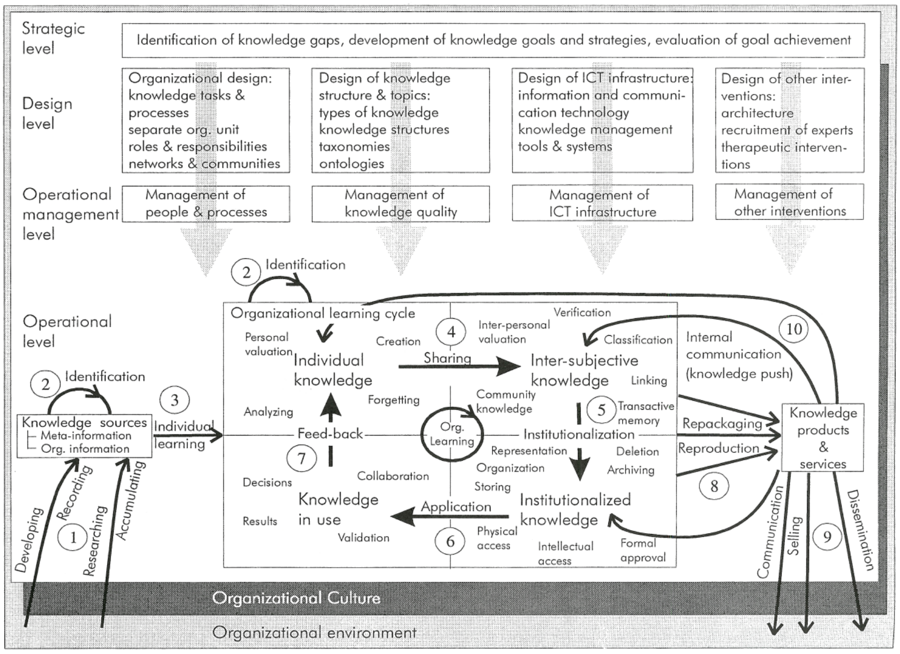


Figure 5.2: Maier's Framework for Knowledge Management [268].

Other frameworks like [269, 270] are based on the so considered pillars for Knowledge Management. Wiig [269] illustrated his framework by visualizing KM as being supported by three pillars of methods and approaches. These pillar on which comprehensive KM rests are:

- **I. Explore the Knowledge and its Adequacy**:
  - Survey knowledge
  - Categorize Knowledge (i.e. describe and characterize)
  - Analyze knowledge and knowledge-related activities
  - Elicit and codify knowledge
  - Organize Knowledge

- **II. Establish the Value of Knowledge**
  - Appraise and evaluate the value of knowledge
  - Establish the value of knowledge-related actions
- **III. Manage Knowledge Explicitly**
  - Synthesize knowledge-related activities
  - Handle, use, and control knowledge
  - Leverage, distribute, and automate knowledge
  - Implement and monitor knowledge-related activities

The foundation of this KM framework reflects Wiig's general understanding of knowledge, that is, how is created, manifested, used, and transferred.

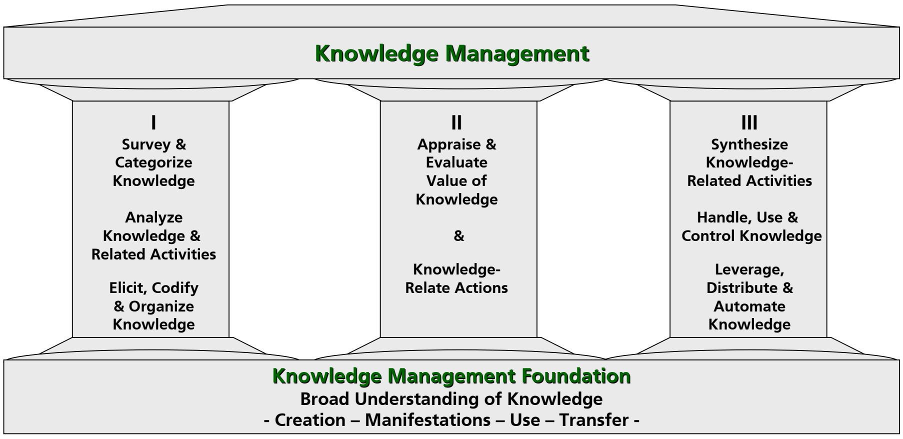


Figure 5.3: Wiig's Three Pillars of Knowledge Management [269].

Proposed by Stankosky [270, 271], the Four Pillar Framework suggests that there are four fundamental elements or pillars to KM:

- **Leadership** - Leads with the environmental, strategic, and enterprise level decision-making processes involving the values, objectives, knowledge requirements, knowledge sources, prioritization, and resource allocation of the organization's knowledge assets.
- **Organization** - Deals with the operational aspects of knowledge assets, including functions, processes, formal and informal organizational structures, control, measures and metrics, process improvement, and business process re-engineering.
- **Technology** - Deals with the various information technologies peculiar to supporting and/or enabling KM strategies and operations.

- **Learning** - Deals with organizational behavioural aspects and social engineering. The learning pillar focuses on the principles and practices to ensure that individuals collaborate and share knowledge to the maximum.

The foundations of the four pillar frameworks are multiple disciples like Systems Engineering, Organizational Development, Systems Management, and Organizational Behaviour.

Figure 5.4: Stankosky's Four Pillars of Knowledge Management.

The last approach presented is the Fraunhofer IPK Reference Model from KM 5.5 proposed by Heisig [264]. The core of the Fraunhofer reference model consists is the business processes as application fields of knowledge. They integrate the knowledge domains and provide the context. The core KM activities (Apply, Generate, Distribute, and Store) relate (and support) the specific business processes. At the last, the six core design areas are related to the importance of the key enablers 'culture', 'organization and roles', 'strategy and leadership', 'skills and motivation', 'controlling and measurement', and 'information technology'. Essential it is defended that an organization's success in Knowledge Management practices is *substantially determined by an organization's ability to handle valuable knowledge resources*.

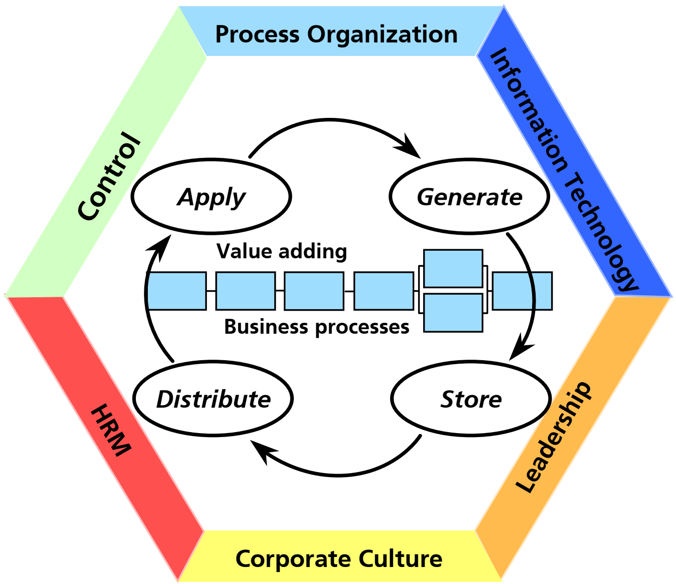


Figure 5.5: Fraunhofer IPK Reference Model from Knowledge Management [264].

5.2 Proposed Framework for Organizations Knowledge Management

The author's presented framework is based on the previous background study. From the frameworks and models presented, is possible to conclude that some of them are strongly focused on the knowledge management itself and not so much on the knowledge integration with the other organizations processes [269, 270]. There is not also a distinction between the organization's business and external processes, being the business related with the organization core activities and the external processes related with communication/collaboration with external networks (e.g. partners, suppliers). The value of each enabler is not also clearly identified in the previous.

To overcome the identified issues, the author propose the framework of Figure 5.6. The core of the proposed framework is described by a three level processes categorization, being the knowledge processes the support of both business and external processes. The characterization of the considered processes is the following:

- **Knowledge Processes** - Consisting on the knowledge related activities of an organization[264, 268, 272]. Accordingly with chapters 2 and 3, examples are: Knowledge identification, acquisition, development, distribution/sharing, preservation and use.

- **Business Processes** - Consisting on the core business processes of an organization [273, 274]. Examples are: Human resources, manufacturing and marketing sales.

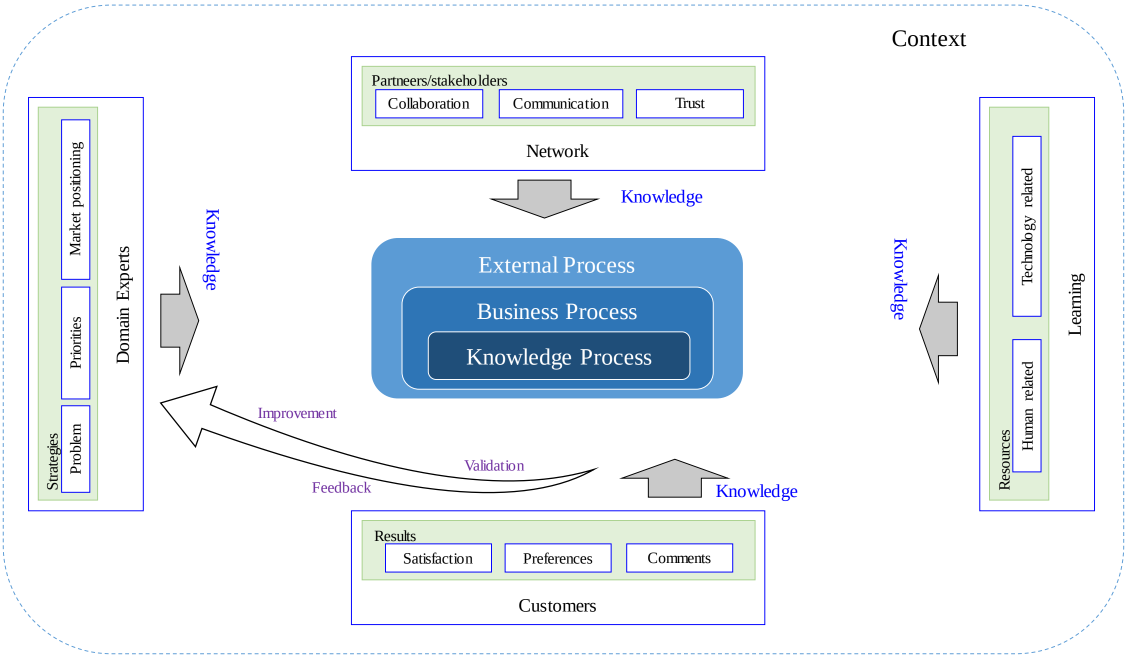


Figure 5.6: Framework for Organizations Knowledge Management.

- **External Processes** - Processes with external agents (e.g. stakeholders and customers) [275]. Examples are: cooperation establishment, awareness building, negotiation (e.g. partners, strategic alliances), cooperation agreement, culture exchange, customers relationship management.

The enablers considered can be divided into four main groups: Organization domain experts, Customers, Infrastructure, Network.

- Going towards to what was described in chapters 2 and 3, **domain experts** knowledge play an important role on organizations strategies, namely problems solution, definition of priorities and decision of markets positioning. Domain experts contribute, as an example, with their know why, know how, working experience, intuitions.

- **Customers** play an important role in products or services validation. They provide the feedback required to improve. To sustain a long term, profitable relationship with specific customers, the knowledge about their preferences and satisfaction levels can be considered essential. Moreover, the knowledge of the comments that customers write in blogs and social networks allows the organization to detect and act upon possible rumours about their services and products.

- **Networks** establishment is essential to create profitability and competitive advantage. Thus, it is necessary to empathize the value of collaboration, communication and trust among the members of the network such as suppliers, manufacturers, distributors and retailers.

- The **Learning** is composed by two main groups: Human-based and Technology-based. The human based is essentially related to learning by example and observation, mentoring, storytelling, training courses, experts interviewing, and others (see chapter 2). Technology-based encompass, as examples, document management systems, learning platforms, knowledge portal systems, supply chain management systems.

As clearly stated in the proposed framework, knowledge management should be core of all organization's processes 5.7: 1) In the knowledge processes itself (e.g. glossary establishment, domain experts' knowledge management), business processes (e.g. requirements management, business processes modeling) and external (e.g. partnerships establishment and customer's relationship management). Thus, to validate the proposed framework the modules corresponding to the next sections were implemented.

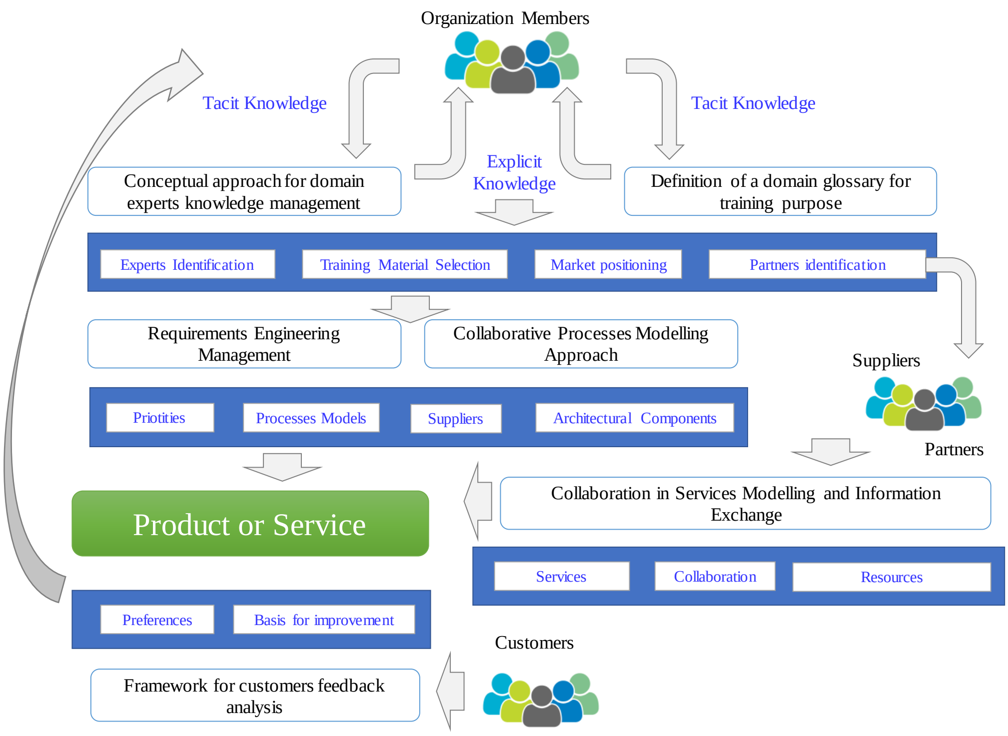


Figure 5.7: Core activities identifies in a Product creation.

5.2.1 Knowledge Processes

Knowledge Processes consist in planing, organizing, motivating, and management of people, processes and systems in the organization to ensure that its knowledge-related assets are improved and effectively employed. Like described in chapter 2, knowledge processes involve knowledge acquisition/creation, formalization, storage, refinement, transfer, sharing and utilization. Its goal is to improve organization's assets to effectuate better knowledge practices, improve organization's behaviors, support better decisions and, in general, improve organization's performance. Mechanisms and strategies such as communication promotion, trust and motivation building, learning and training are considered factors that influence creating and sharing knowledge culture [39].

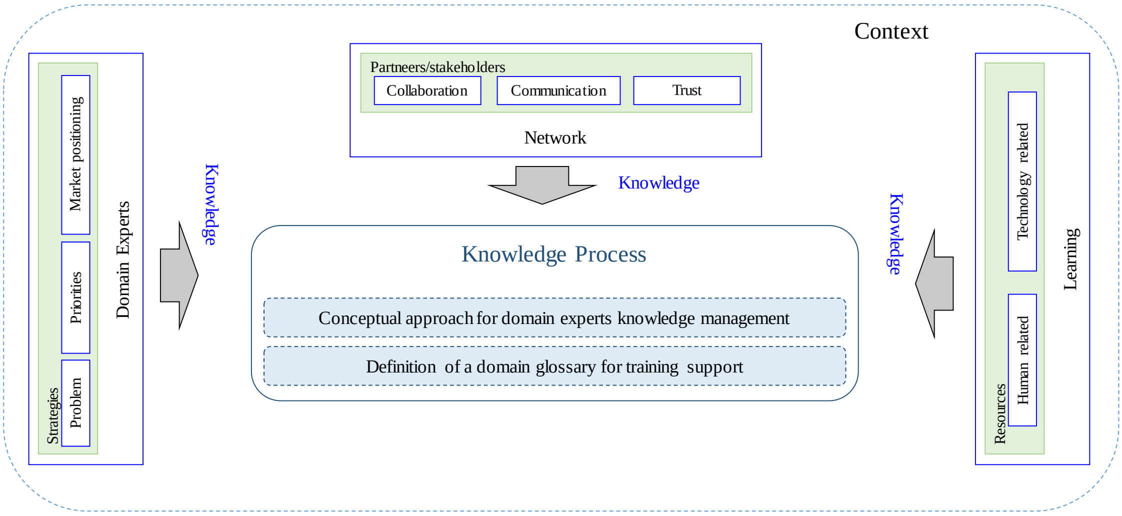


Figure 5.8: Knowledge Processes.

Contributing to this topic, the author proposes a novel conceptual approach for domain experts knowledge management. It is grounded in the idea that the tacit knowledge gathering from domain experts and its posterior transformation into explicit knowledge (ready to be used by a full community) should be facilitated. It is a fact that when an information system intends to represent a domain's knowledge it needs to be aligned to the community that it represents. Consequently, it is required to have a solution where community members could present their knowledge about the domain and discuss it with their peers. Additionally, such knowledge must be available and dynamically maintained by all the involved actors. It can be done through an explicit information front-end.

The author also proposed the definition of a domain glossary to facilitate resources sharing and understanding. The proposed solution, explained in detail in chapter 7, consists in the taxonomic representation of reference concepts. It will allow trainees to use a reference lexicon to communicate between them, and at the same time, interpret network semantic sources. Consequently, knowledge assets can be better transmitted and shared between community's members.

5.2.1.1 Conceptual approach for Domain Experts Knowledge Management

The proposed KB establishment concept is presented in Figure 5.9. As can be observed, one of its components is an explicit information front-end (e.g social software), where the knowledge is kept in a format that allows domain experts to consult and utilize it. In turn, domain experts need to be able to use the explicit information to turn it into their own personal knowledge in order to create and share additional (explicit) knowledge from it. This corresponds to the right cycle of Figure 5.9, which is aggregated through automatic synchronization with the left cycle of the figure, in such way that if there is new knowledge added by a domain user, it would smoothly be available in the knowledge base for any further application (e.g. enhanced searching or reasoning services). The result is an ontology, which model is constantly refined accordingly with the explicit information front-end module in order to better handle the knowledge provided by the domain experts.

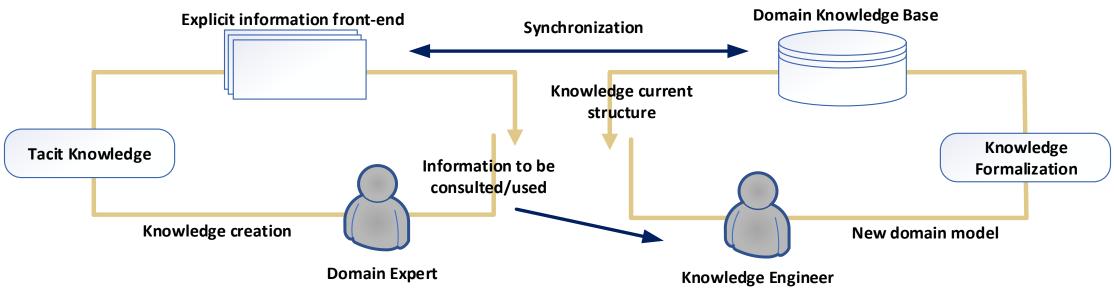


Figure 5.9: Proposed concept for Knowledge Management.

5.2.1.2 Definition of a Domain Glossary for Training Support

Learning applications that can benefit from domain glossaries usage are search engines and plagiarism detection. Plagiarism is commonly defined as literary theft, i.e. stealing words or ideas from other authors [276]. The growth of Internet usage increased the availability of information, not only for legitimate purposes but also for plagiarism or similar. One type of plagiarism consists in the copying and merging of text segments with slight modifications, e.g. changing words order, replacing words with synonyms, or by entering or deleting filling words [277]. For that reason, many researchers recognise the need to incorporate semantic information into similarity checks to allow detecting plagiarism [278, 279].

One example of glossaries applications is Google's search motor in which the exact words in the query do not need to be anywhere in a page to be ranked. This is fundament by Matt Cuttsin in an interview, which he stated: 'Key phrases don't have to be in their original form. We do a lot of synonyms work so that we can find good pages that don't happen to use the same words as the user typed.' [280]. Thus, glossaries application to communities resources search (by providing an understanding of shared resources) could increase the compatibility between users suggested learning objects or other users to interact (e.g. teachers and fellow students).

5.2.2 Business Processes

There are several definitions of business processes in the literature. Jacobson [281] describes a business process as the *set of internal activities performed to serve a customer*. In the first chapter of Ould's book [282], the business process concept is defined through a set of key features: purposeful activity; carried out collaboratively by a group; it often crosses functional boundaries; it is invariably driven by outside agents or customers.

A more informal definitions is the one adopted by PNMSoft that refers to business processes as a series of steps performed by a group of stakeholders to achieve a concrete goal. A business process is a series of steps performed by a group of stakeholders to achieve a concrete goal. These steps are often repeated many times, sometimes by multiple users and ideally in a standardized and optimized way [283].

The progressive adoption of Business Process Management (BPM) paradigm by enterprises puts the spotlight on the business process life-cycle and on tools and technologies to support each stage of process modelling [284]. A business process can be defined as a set of activities performed in coordination in an organizational environment to reach a business objective [285]. Therefore, to better align the implementation and support of a process life-cycle, a separation of the business points of view from the technical and physical means to realize it is required [286], hence **promoting the coordination and cooperation between teams**. This is why a proper solution for both business processes requirements elicitation and modelling should be implemented.

5.2.2.1 Requirements Engineering Management

Requirements Engineering is the science and discipline concerned with analysing and documenting requirements [287], and follows a process that leads to a set of well-formulated requirements. The classical process, composed by 4 semi-processes, starts by the requirements elicitation and goes to its analysis, specification, and validation [288]. The defined processes are known to produce quality based on the knowledge or strategy of the company [289]. Thus, **techniques that allow the requirements elicitation from all the company sectors (e.g. marketing, management, and operators) should be implemented**. But before the elicitation process, a preparation stage should be followed in order to analyse the impact of several well-known elicitation techniques. While some elicitation techniques like brainstorming and brainwriting might work with the marketing and management sectors, some constrains, like physical location, might limit its application to the development

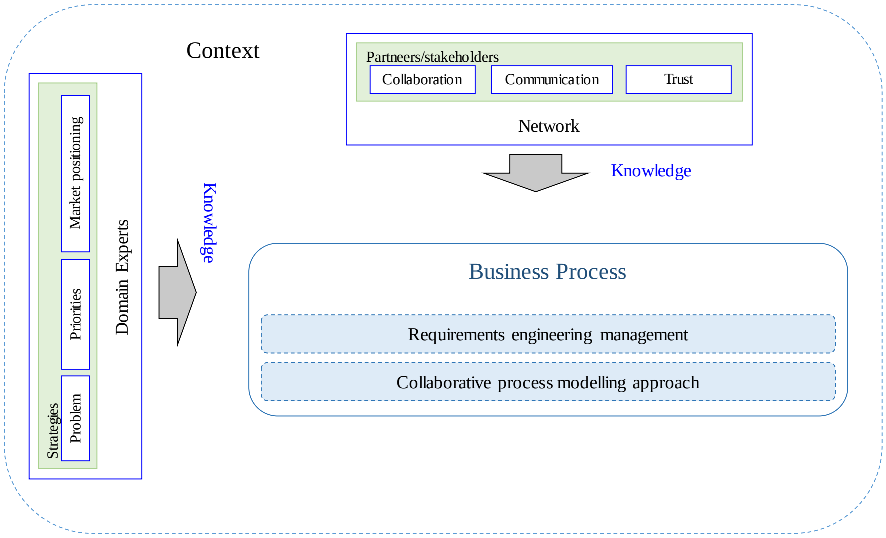


Figure 5.10: Business Processes.

sector. In these cases, questionnaires can show better results.

An important precondition to requirements elicitation is the type of requirements. Experts have long suggested that inconsistent market requirements can adversely affect manufacturing performance [290]. An often encountered scenario involves the consideration of both marketing requirements and manufacturing goals. Traditionally, sales are focused in serving the varied customers' needs. This situation goes against the effective utilization of manufacturing resources and contributed to poor cash flow [291]. Thus, it is important to differentiate the types of requirements to avoid conflict situations. Even more important than conflicts avoidance, a proper approach that allows both business and development requirements to coexist and 'communicate' is required. That is the reason why the author considers that Behaviour Driven Development approaches should be used.

Behaviour Driven Development

Behaviour Driven Development is a specification technique that '*automatically certifies that all functional requirements are treated properly by source code, through the connection of the textual description of these requirements to automated tests*'. Similarly, [292] focus on the implication of BDD as a design technique and states that BDD is used to integrate products verification and validation in the design phase using an outside-in style, which implies thinking early on what is the client acceptance criteria before going into design of each part that composes the functionality.

Other works like [293] argue that BDD is relevant in the whole product life-cycle, especially in the interaction between the business and software development. In addition,

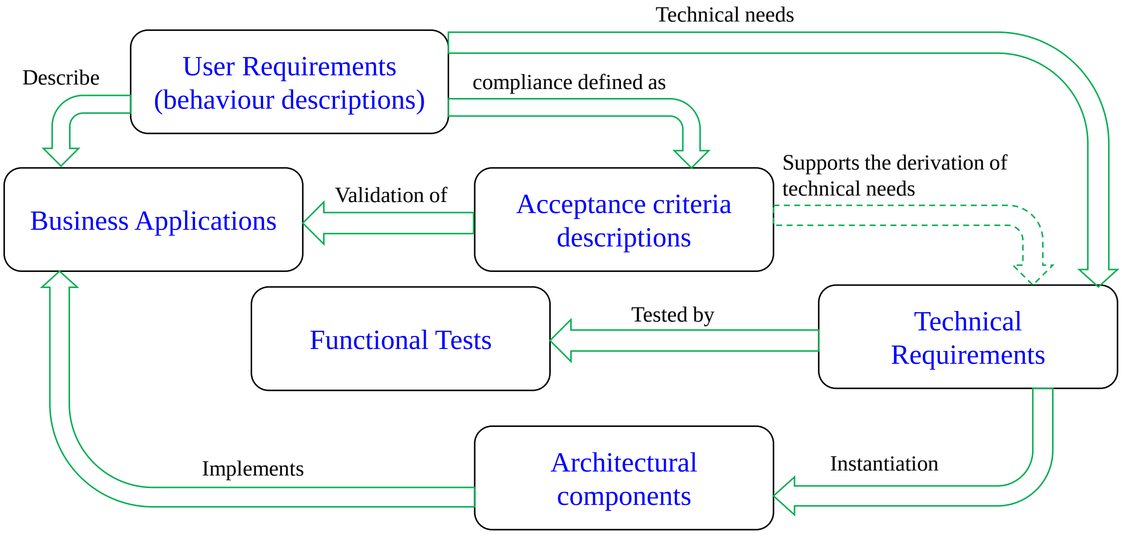


Figure 5.11: Requirements Management Positioning in a product of service life-cycle.

this author argues that BDD permits to deliver value by defining behavior, and it is focuses on **learning** by **encouraging questions, conversations, creative explorations and feedback**. In [294], it is highlighted the value of BDD for business domain and the **interaction of business and developers**, claiming that BDD **allows developers and domain experts to speak the same language** and encourages collaboration between all project participants.

The methodology represented in Figure 5.11 illustrates how the communication between the business and developing (technical) teams can be achieved. At first, the user requirements are elicited and managed. This is done in such a way that each requirement is written in a form of a behaviour description. Thus, its structure intends to specify:

- **Who** or **which** business or project role is the driver or primary stakeholder of the story (the actor o derive business benefit from the story);
- **What effect** the stakeholder wants the story to have;
- **What business value** the stakeholder will derive from this effect.

To accomplish this, the requirement (user story) template should take the form:

- **As a** [X]
- **I want** [Y]
- **So that** [Z]

Where Y is some feature, Z is the benefit or value of the feature, and X is the person (or role) who will take benefit of the feature.

The defined user requirements, are then used to derive the validation criteria for a specified business application. Since a story's behaviour is simply a business application acceptance criterion, if the system fulfils all the acceptance criteria, it's behaving correctly. Following the proposed methodology, the description of the acceptance criteria in terms of functional tests should take the following form:

- **Given** some initial context
- **When** an event occurs
- **Then** ensure some outcomes

Since user stories are written using an unambiguous language, the technical team is able to understand the purpose of the requirement and define some technical requirements from them. Unlike user requirements, that are essentially descriptions of how the system is supposed to do, technical requirements are directives of how the system should be built, or which components are demanded. Typically, each user requirement means one desired behaviour and should originate one or more technical requirements. On the other hand, one technical requirements can be used to accomplish several user requirements. Thus, a N:N relationship exists between user requirements and technical requirements. The same happens with the relation between technical requirements and architectural components. A technical requirement can be linked to one or more architectural components, as an architectural component can be used to implement one or more technical requirements.

5.2.2.2 Collaborative Process Modelling approach

Research in Enterprise Interoperability suggests that organizations can seamlessly interoperate with others at all stages of development, as long as they keep their business objectives aligned, software applications communicating, and the knowledge and understanding of the domain harmonized [295]. However, business process modelling is frequently overcomprehensive and hard to accomplish in a collaborative way, due to the extended domains and activities/services each enterprise is providing. With the advances and integration of IoT (intelligent) devices providing services for the network, that gap is becoming more obvious. The community is in need for an approach that puts in practice the separation of concerns to shorten the domain of analysis of business objectives, bringing together teams and enterprises with different know how, to collaborate.

Part of the problem can be tackled following a model driven approach, which instead of writing the code directly, enables software products to be firstly modelled with a high level of abstraction in a platform independent way. It provides many advantages like the improvement of the portability, interoperability and re-usability through the architectural separation of concerns [296]. In fact, Ducq et al. [286] adapted this concept to manufacturing services design and development, with the Model Driven Service Engineering Architecture (MDSEA). It follows the Model Driven Architecture (MDA) and Model Driven Interoperability (MDI) principles [297], supporting the modelling stage and guiding the transformation from the business requirements (Business Service Model (BSM)) into detailed specification of components that need to be implemented (Technology Specific Model (TSM)).

The MDSEA approach implies that the different models, obtained via model transformation from the upper-level ones, should use dedicated service modelling languages that represent the system with different concerns, i.e., ICT, Human and Physical levels [298]. Indeed, MDSEA provides the building blocks for enterprise system development in the scope of an ecosystem of collaborating teams (enterprises), providing: 1) The capability to transform a business specific model into a functional one so it can be perfected by a system architect detailing the necessary resources; 2) The capability to transform a functional model into a technology specific one envisaging the generation of concrete services.

5.2.3 External Processes

A partnership is where two or more people need to work together to accomplish a goal while building trust and a mutual beneficial relationship. This means the partnership is voluntarily agreed upon, built on the desired to have trust, and based on agreed-upon mutual benefits [299].

One of the benefits that can be achieved by partnerships establishment is the capability to react to specific requirements of customers, and to take position on the new market. Even if the success of these partnerships is related to the reach of the objectives of enterprises (customer satisfaction, positioning on new market), it is important to keep in mind that this success is also related to the quality of communication and interaction between partners [300]. Thus, the interoperability between the information exchanged and customers satisfaction through feedback analysis are issues that the author intend to address under the External Processes concept

5.2.3.1 Enable Interoperability in Services Modeling and Information Exchange

In the emerging society, competitive markets are becoming increasingly dynamic, and consequently complex, with companies not prospering and surviving through their own

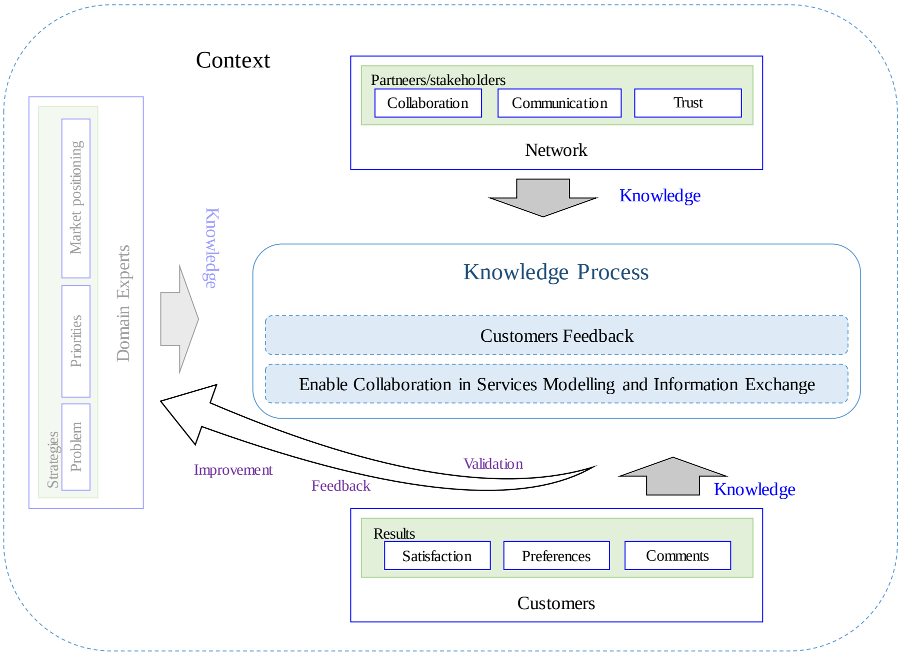


Figure 5.12: External Processes Processes.

individual efforts [144]. Thus, global markets are willing to improve their competitiveness through collaboration and partnerships, motivating companies to look for enhanced interoperability between systems and applications in industry. Therefore, enterprises need to be able to dynamically adapt themselves to take advantage of market opportunities, establishing collaborative business practices to compete with big enterprises [301]. Moreover, the ability of an enterprise to interoperate with others is not only a recognized quality and advantage to obtain competitiveness in today's market, but it also becomes a matter of survival for many companies.

Interoperability can be defined as the ability of two or more systems or components to exchange information and to use the information that has been exchanged [149]. It means that enterprises need their systems/products to work with the surrounding systems/products without great efforts to be interoperable. However, one of the main difficulties regarding the interoperability between systems and applications is related to the high number of semantic representations of the same segment of reality (e.g. systems and products) which are not semantically coincident (even inside the same domain) [184], as a consequence of the heterogeneity of communities and enterprises. It results in a difficult semantic interoperability achievement. Semantic Interoperability, defined as the ability of systems/components to share and understand information at the level of formally defined and mutually accepted domain concepts [302], traditionally is achieved through peer-to-peer mappings where each participant tends to use its own data format and business rules, handling as many mappings as the number or partners to interoperate. Here, another interoperability issue emerges, one related to systems dynamics. Commonly, systems are time-variant, and even if we are able to find a 'good' model to describe it, when facing the dynamics of the environment, such model will become obsolete in time. Moreover, a model is just a representation of how an entity see the world. It can (and should) be constantly refined in order to adapt to new requirements. As a consequence, all this dynamics and heterogeneity leads in most cases, the network to experience interoperability problems because if just one of the network members adapts to a new requirement, the harmony is broken, and the network begins experiencing interoperability failure [303]. This is even more evident in multi-domain networks (e.g. collaborative product design) where information is dispersed and frequently replicated in many Information Systems through Web Services usage.

5.2.3.2 Customers Feedback

Social software tools typically handle the capturing, storing and presentation of communication and focus on establishing and maintaining a connection among users. Due to its current increase of usage (e.g. facebook, twiter), it has become a big data source. Big data is a term applied to data sets whose size is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time [304]. Social software focuses on human communication, thus the vast majority of the information is in human readable format only. As a result, researchers are actively contributing to the appearance and enhancement of techno-logical solutions to handle such data sets with the goal to supply new mechanisms for enterprise's management [305]. This results in the increase of companies' use of web software to promote services and products, which is also often used by their customers to post their comments expressing sentiments as response [306]. Consequently, web software is commonly accepted as a communication bridge between customers and companies, where users' opinion become a major criterion for the improvement of the quality of services. Blogs, review sites and micro blogs provide a good under-standing of the reception level of the products and services [307]. Consequently, a topic that is currently getting much attention is how to use electronic customers comments to increase the quality of companies' services and products in order to increase organizations competitiveness level [306, 308]. Additionally, a relevant challenge is how to handle real-time data analysis. Thus, it is desirable to automatically provide to a company manager the formal knowledge about the following questions: *What is the comment about?*; *What is the polarity of the comment? Is it positive or negative?*.

The response to the previous questions and the capacity to manage the knowledge about customers' opinions would allow a company to improve the quality of the offered services and/or products by reacting to customers' sentiments, and avoiding to spend, for instance, a large amount of money on customers' satisfaction surveys. In addition, the use of surveys, if they are not in the form of open question, may narrow companies' range of understanding and actions. An example is the open question: *Which are the things that you dislike the most?*, and the more narrow question *Did you like the staff?*. Based on the second question, the manager can only make fact-based decisions relative to the staff. The free opinion sharing (first question) provided by customers' sentiments, besides helping to identify key improvement variables, can also be used as a rumour detection. As an instance, if several customers write something like: *The food is very bad and I advise everyone to never go to that place*.

A quick detection of this rumour will allow the company manager to react accordingly and avoid major losses. Another advantage resulting of such sentiment analysis is the potential for "spying" the market competition. Most of the web software allow access to the HTML source. Thus, by accessing to the comments posted in the competition page it is possible for a company to track customers' preferences and act in accordance to that background.

CHAPTER 6

Conceptual Approach for Domain Experts Knowledge Management

As stated before, a key factor to increase organizations competitiveness is their ability to capture their members' knowledge and experience and transform it into organizational knowledge. This need grounded the knowledge management establishment process presented in section 5.2.1.1 where both knowledge engineers and domain experts contribute to increase organizations' knowledge (explicit knowledge). Then, that approach supported the development of the framework for knowledge management presented in this chapter. This framework uses simple wiki-based front-end modules where tacit knowledge can be expressed in a form of explicit knowledge directly by the different employees.

6.1 Related Work

Knowledge management tools are pieces of software that enable the user to create, edit or perform other operations over explicit knowledge forms (e.g. ontologies). In [309], the authors consider that ontology tools can be applied in all the stages of the ontology life cycle (creation, population, validation, deployment, maintenance and evolution) . Some ontology management tools to consider are Ontopia[1], TM4L [310], and Protégé[2]. They are all very complete, since they all provide support to several types of ontology languages (OWL, RDF, XML) and graphic visualization methods. However, in what concerns domain experts' usage, they may be difficult to use without knowledge engineers support. For that reason, the suggested knowledge management approach relies in the collaborative

[1]http://www.ontopia.net/page.jsp?id=about
[2]http://protege.stanford.edu/

aspects of Semantic wikis to allow collaborative knowledge management in an iterative way by domain experts, which might not have the technical skill required for complex solutions. By using widespread and well-accepted wiki technology, domain experts are able to model and update their knowledge in a familiar environment by reusing externalized knowledge already stored in wikis.

Semantic wikis enrich wiki systems for collaborative content management with semantic technologies [311]. An overview of relevant research can be found in [312], where is possible to verify that prominent wikis like Semantic MediaWiki [313], ikeWiki [314], and SemperWiki [315], manage to disseminate semantic technologies and are used to support several semantic applications. Thus, domain experts and ontologies are able to cooperate in one system while wiki pages are presented in a human-readable format in parallel to the formal ontologies. Some works to consider are [316], and also [317]. In the first work, the authors gather wiki knowledge by defining a set of relations between Semantic MediaWiki annotations and OWL DL concepts [316]. In the latter, the authors also focus on Semantic MediaWiki annotations, but with some interactive assistance to support users in the knowledge representation process [317]. They provide functionality for collaboratively authoring, querying and browsing Semantic Web information. In both their works, explicit knowledge is achieved through a set of mappings that relate with ontological concepts. This is very powerful when one is aiming to build machine reasoning and intelligence capabilities. Nevertheless, in their proposal, all the textual and descriptive information is lost, which can be a major drawback when a feedback loop based on natural language needs to be maintained with Human users. The proposed work addresses this challenge complementing the state of the art by building a knowledge base where not only annotations are used to create ontological relations, but also content from wiki articles, gathering natural language descriptions in data properties, and consequently obtaining a richer representation of a domain.

6.2 Framework for Knowledge Management

The framework instantiates the knowledge management approach and uses ontologies and wiki front-end modules, able to facilitate the achievement of explicit knowledge from domain experts' tacit knowledge. Since the knowledge is constantly refined and updated by the domain experts' community, it would allow to make decisions based on individual's tacit knowledge. As can be observed in Figure 6.1, the proposed framework is composed by four modules: 1) wiki-based front-end; 2) Synchronization module; 3) Knowledge

Base; and 4) Reasoning and Decision Making module. The Framework input is the front-end users' knowledge, which is processed and consumed by the community that uses it to re-feed this cycle with more knowledge.

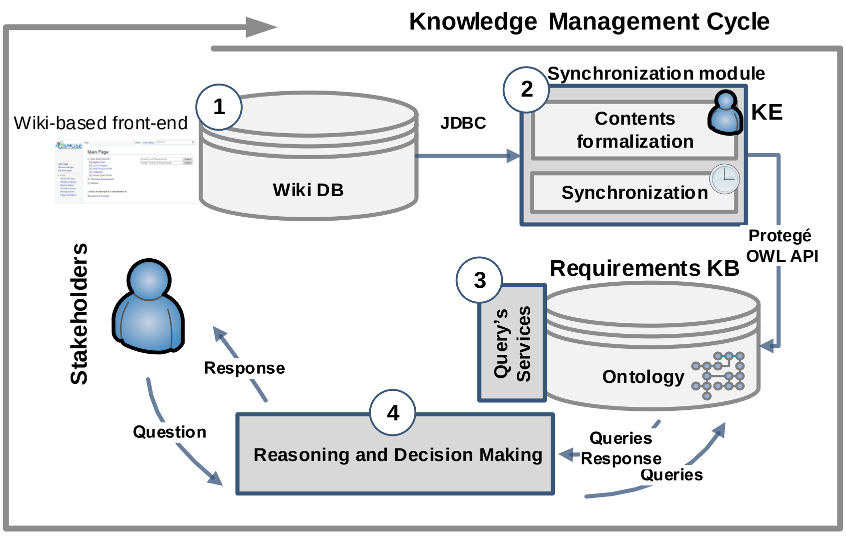


Figure 6.1: Framework for Knowledge Management using wiki-based front-end modules.

As illustrated in Figure 6.1, the first module is a wiki-based front-end which corresponds to the explicit information front-end of the knowledge base establishment process presented in Figure 5.9. It is characterized by being collaboratively edited by domain experts based on the knowledge consulted. However, the content of wiki-based front-ends is characterized by being human readable only. This means that its content is not formalized to facilitate computerized use (e.g. reasoning). To overcome this issue, a synchronization module, composed by two sub-modules: 1) Contents formalization; and 2) Synchronization is used for contents formalization and its storage in a knowledge base. Thus, in this context, the purpose of the ontology, modelled by knowledge engineers, is to hold the explicit knowledge about a domain in a formalized way so that it can be used by the community for reasoning purposes. Such functionality is performed by the module Reasoning and Decision making, which is able to provide structured and useful contextual information.

6.2.1 Wiki-based front-end contents formalization methodology

Wiki-based front-ends are encyclopaedias that are collaboratively edited by its users, which contribute with their (tacit) knowledge. A key factor to extract knowledge from wiki-based front-ends is that such pages often follow a global template that facilitates the retrieval of information. Such front-ends provide categories that are used to classify articles and other pages. These categories are implemented by MediaWiki[3]. They help readers to find, and navigate around, a subject area, to see pages sorted by title, and thus find articles relationships. One particularity is that the resulting category system can consist in a hierarchical representation of categories related, as an example, by the relation 'is a', as the classes in an ontology.

Organized using several body sections, wikis use their headings to clarify articles and break the text, organizing its content (e.g. article title, sections, subsections). Some sections of articles can contain hyperlinks, and they point to a whole category, article or specific element of an article. A hyperlink between several pages, can somehow, be compared to a relation between instances of an ontology. Therefore, the organization of an article can be seen as a characterization by properties of its content (object and data properties).

Based on that organization of wiki-based front-ends the methodology for contents formalization of Figure 6.2 is proposed. As can be observed, the step 0 of the methodology consists in the creation of a wiki root class in the ontology. It will handle the knowledge represented by the domain experts in the wiki-based front-end. The process of assigning categories to other categories, in the proposed methodologies (step 1), will be used by the knowledge engineers to build ontology's classification taxonomy, being the tagging between them handled as the ontological relation 'is a'. The classification of categories' contents can be facilitated if a classification taxonomy of those contents is defined (step 2). This will allow to better structure the gathered knowledge and visualize relations between knowledge base's instances.

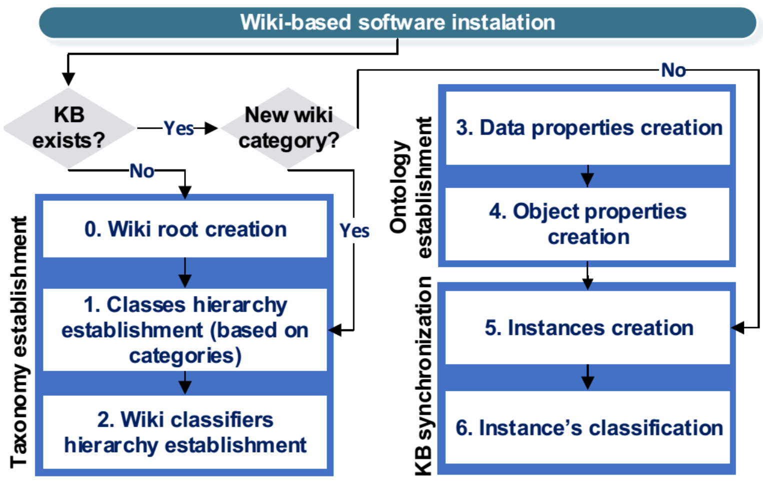


Figure 6.2: Methodology for wiki-based front-end contents formalization.

In this methodology it is assumed that the content of all pages under a specific category follows the same structure. With that assumption, it is possible to follow with the steps 3

[3]http://en.wikipedia.org/wiki/MediaWiki

and 4 of the methodology. In step 3 and 4, for each article section is created a data property or object property to represent that knowledge in the ontology. The object properties created will connect the classes under the wiki root class and those under the classifiers taxonomy previously defined. Data properties will represent knowledge that is not under that taxonomy.

The process of assigning articles to categories, in the proposed methodology (steps 5 and 6) will be used to instantiate the ontology. This is done by creating an instance under the class with the article's category name (step 5). Then, based on HTML analysis of articles' content, the knowledge of its sections can be represented in the data and object properties of the previously created instance (step 6).

The methodology also covers the creation of a new category on the front-end after the knowledge base is defined. It is aligned with the necessity of domain experts to share new kind of knowledge, which is not formalized yet.

6.2.2 Synchronization between wiki-based front-end modules and ontologies

The synchronization module runs periodically and starts by connecting to the wiki front-end database in order to verify if any changes occurred since its last run. JDBC (Java Database Connectivity) is used to querying the front-end database[4]. By querying the wikimedia table 'recentchanges', the authors have access to the set of changed pages, and its type: edition, creation, or removal. If the change is an edition or a creation, through the link to the table text (links to new & old page text) it is possible to have access to the current content of the front-end page.

After the collection of the recent changes the HTML of each article or category's page is processed in order to create/ populate the necessary instances, data properties and object properties in the knowledge base (steps 5 and 6 of the proposed methodology). In these steps of the execution flow it is also verified if the information remains consistent (e.g. the pages (articles) of the same category have the same structure). After the processing of all detected changes, the update of the ontology is made. This update is made using Jena OWL API. It provides the necessary classes and methods to load and save OWL files and to query and manipulate OWL data models.

[4]https://upload.wikimedia.org/wikipedia/commons/4/41/Mediawiki-database-schema.png

6.3 Supporting The EISB During the ENSEMBLE Project

ENSEMBLE (Envisioning, Supporting and Promoting Future Internet Enterprise Systems Research through Scientific Collaboration)[5], was a Support Action funded by the European Commission (EC) that coordinated and promoted research activities in the domain of Future Internet Enterprise Systems (FInES), providing a sustainable infrastructure for the FInES community to contribute and support the EISB (Enterprise Interoperability Science Base) initiative [318], as well the 2015 Roadmap [319]. The FInES cluster, now DBI community[6], has been supported by the EC in support of the Digital Agenda for Europe, a flagship initiative of the Europe 2020 strategy.

The following scenario is related to the gathering of tacit knowledge from the FInES community and transforming it into explicit knowledge, so that it could be available to the full community, and support knowledge intensive initiatives such as the EISB. To achieve that, a wiki-based front-end (explicit information front-end) has been used, the FInESPedia[7]. It provides explicit knowledge to the community users and allows them based on that, to create new tacit knowledge and post it in the front-end. Moreover, the knowledge provided by the users is formalized in the EISB reference ontology.

6.3.1 FinESPedia

FInESPedia aims at providing an overview of the state of the art in Future Internet Enterprises Systems. This source of knowledge, more focused on the collaborative gathering and sharing of information from domain experts, is accessible through the FInES cluster portal[8]. As can be observed in Figure 6.3, its homepage is dived into four main sections, namely: 1) FInES Research Roadmap 2025; 2) FInES Position Paper Towards Horizon 2020; 3) Enterprise Interoperability Science Base (EISB) where this use case is focused; and 4) FInES Task Forces.

Going into further detail (on the EISB), the FInESPedia is essentially composed by Scientific Areas, EISB Glossary and also the Neighbouring Domains Glossary, that are being synchronized with the EISB ontology, as explained next. To the formalization of FInESPedia front-end knowledge, the methodology of section 6.2.1 was followed.

[5]http://www.fines-cluster.eu/jm/ENSEMBLE-Public-Category/ensemble-objectives.html
[6]http://www.dbi-community.eu
[7]http://finespedia.epu.ntua.gr/
[8]http://www.fines-cluster.eu/jm/

Figure 6.3: FinESPedia Main Page.

6.3.2 Application of the methodology

The EISB knowledge base is a component that intends to capture ENSEMBLE community knowledge with precise and semantically meaningful definitions. As explained along the paper, it also serves as a facilitator for knowledge reasoning, allowing different views of the information gathered from the wiki. Having this kind of knowledge would facilitate the search of specific information, for instance papers or methods of a specific EISB area, or a specific set of tutorials related to a specific EISB topic, or even a set of expert researchers. Furthermore, this ontology can be a valuable asset for the scientific base itself, gathering meta-information relevant to both Enterprise Interoperability and the neighbouring domains [320].

6.3.2.1 Taxonomy establishment based on the methodology

Figure 6.4 represents the application of the methodology to establish the knowledge base taxonomy. As can be observed, step 0 of the methodology consists in the creation of the class 'EISB_Wiki', which will handle the categories' taxonomy represented in the wiki-based front-end.

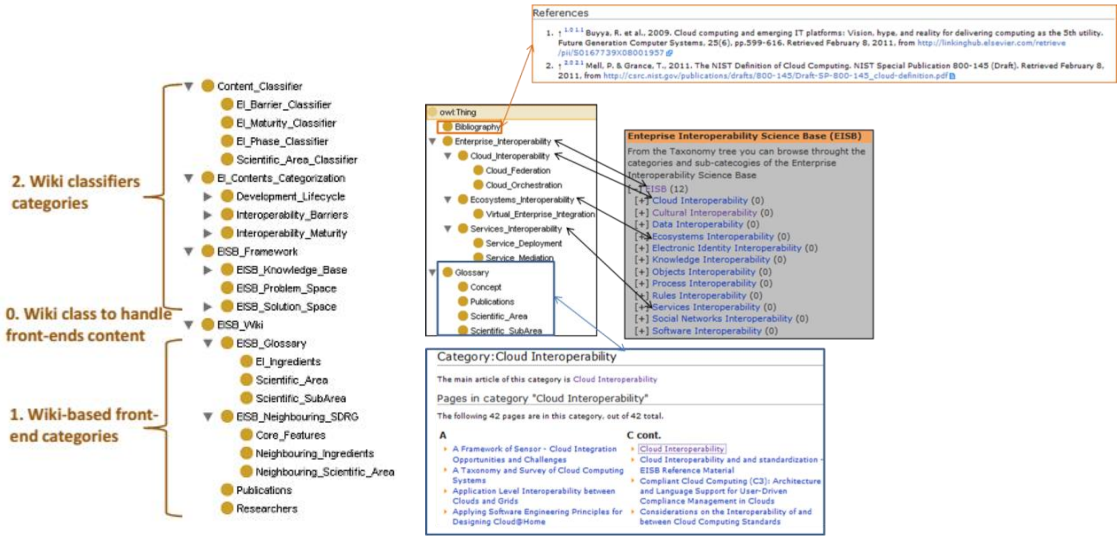


Figure 6.4: Knowledge Base's taxonomy establishment.

The step 1 of the methodology was accomplished by navigating in the front-end articles' category classification, as was explained in section 6.2.1. It results in the identification of four main classes to be handled under the class 'EISB_Wiki':

- EISB Glossary - Representation of the contents of the glossary page of FInESPedia, including: EI Ingredients, including the detailed information about the various EISB ingredients (e.g. methods, tools, experiments...); Scientific Area, regarding the EISB scientific areas represented in the wiki page; and Scientific SubAreas, regarding scientific sub areas represented in the FInESPedia [321];
- EISB Neighbouring SDRG - Serves the same purpose of the EISB Glossary, but refers to the Neighbouring domains instead; (see [322] for technical details on the neighbouring domains);
- Publication -Information regarding the publications presented in FInESPedia;
- Researchers -Information about the researchers acting in the EISB community.

Step 2 of the methodology consists in the categories contents' classifiers. This is a knowledge engineers' works in which they analyse the knowledge that the domain experts want to represent in order to create a classifiers taxonomy from it. The four main classes of the classifiers taxonomy are:

- EI Contents Categorization - that aims to represent the information about the different categories that the content of the wiki can take, namely: Interoperability Maturity, which holds the information about the various maturity models available;

Development Lifecycle, which houses the information about the different development phases of certain publication (Assessment, Design, Implementation); and Interoperability Barriers, Indicating which type of EI barrier is targeted accordingly with the image of the ISO standard 11354 [323];

- Content Classifier - which stores information relative to classifications of the EISB contents: EI Barrier Classifiers, which assigns (High-Low) relevance of a certain content regarding its interoperability barrier (e.g. Technical- High); EI Maturity Classifier, which has the information relative to the maturity of the wiki content (e.g. mature, infant, ...); Phase Classifier, which classifies publications relatively to its development lifecycle (e.g. Design-High); and Scientific Area Classifier, which classifies a wiki content with the relevance pertaining to a certain scientific area (e.g. Data Interoperability - Medium);

- EISB Framework -the purpose of this class is to hold information about the elements that compose the EISB universe. It handles the knowledge about the framework components: EISB Knowledge Base (the scope of the previous descriptions); EISB Problem Space; and EISB Solution Space (Hypothesis, Laws, etc.) [320].

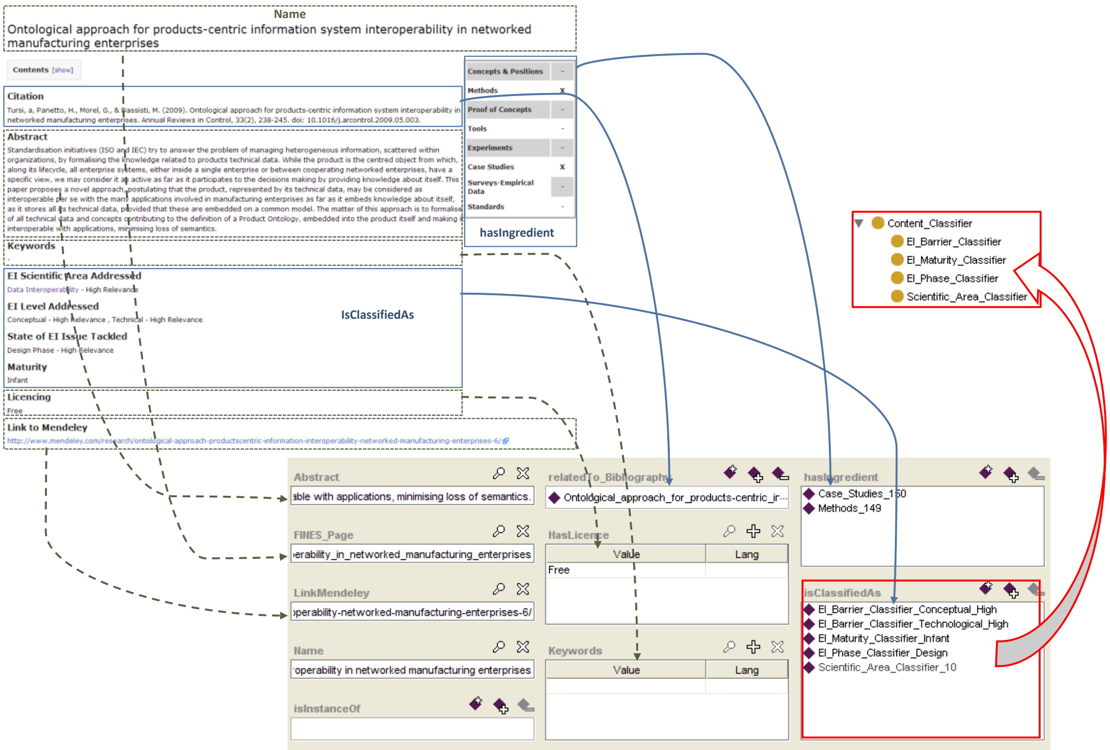


Figure 6.5: New Publication demonstration Scenario.

6.3.2.2 Ontology properties establishment based on the methodology

In this subsection, the steps 3 and 4 of the methodology are demonstrated. The type of pages that were selected to exemplify the methodology were those under the category 'Publications'. It was assumed that the articles under this category follow the same structure of the page illustrated in the top of Figure 6.5. Concerning its content and the classifiers taxonomy established on step 2, the data properties (green dotted areas of Figure 6.5) defined are: 'Abstract'; 'FINES_Page'; 'Keywords'; 'HasLicence'; 'Link_Mendeley'; and 'Name' (step 3 of the methodology). The object properties (blue line continued areas) defined were: 'hasIngredient'; 'IsClassifiedAs'; and 'related_to_Bibliography'.

6.3.3 Ensemble's Knowledge base synchronization

After structuring the information retrieved from the wiki front-end, and concerning the scenario of a new publication creation, it is possible to do the synchronization between the front-end and the ontology in order to populate the knowledge base with domain experts' knowledge. The synchronization tool is triggered by a 'cron job' that runs daily. Then, the recent changes are analysed in order to verify if there is any new publication in the FInESPedia. That verification is made by analysis of the HTML content of the pages to verify in which category the page belongs.

The wiki front-end to ontology synchronization of a new publication is illustrated in Figure 6.5. It is possible to verify that the various sections of the wiki page have a direct correspondence in the ontology (result of knowledge engineers work), and all the contents are therefore successfully migrated. It is also possible to verify that the object property 'IsClassifiedAs' relates the wiki pages' content with a taxonomy under the classifiers defined in step 2.

After contents formalization, the knowledge management framework is capable of handle articles' creation, edition and elimination without the intervention of knowledge engineers. However, if other non-modelled category occurs (other type of tacit knowledge), knowledge engineers need to re-follow the proposed knowledge structuring methodology. In such way, the new sub-domain knowledge inserted by the domain experts can be transformed into explicit knowledge to be presented to the community.

6.3.4 Enabling the integration with other works

By having explicit knowledge formalized, it also becomes much easier to integrate complementary knowledge. In the case of ENSEMBLE this situation became very clear with the example of the FInES 2025 roadmap. Like the EISB, also the roadmap has been supported by an ontology for knowledge management with a wiki front-end (in this case only to visualize information). Due to both ontologies links were easily defined between knowledge domains, enabling readers to navigate through the wiki between the roadmap and the EISB knowledge, increasing their awareness of the FInES and enterprise interoperability domains.

6.4 Discussion

To increase the competitiveness level, organizations must be able to keep its employees' knowledge inside the organization, even when they leave. The same happens with researchers and scientists so that the community can capitalize their knowledge. Hence, responding to those needs, a knowledge management framework based on wiki front-end modules was implemented. With the proposed framework, domain experts can actively contribute to the knowledge of their community through a simple and, considerably well-known interface, as wiki-based front-ends. This kind of front-ends are known by being easy to setup and use, tracking of changes, and on-the-fly publishing. Thus are being largely selected to share knowledge in several areas like teaching [324], collaborative modelling [325], process development [326], and others. However, beside the mentioned advantages, wikis are also characterized for being human readable only. This means that its content is not formalized to facilitate computerized use, an issue addressed by Semantic wikis. Most of the state of the art solutions are based on mappings between semantic annotations and ontological relations. The presented solution is able to complement that, handling all of the wiki articles content in natural language.

CHAPTER 7

DEFINITION OF A DOMAIN GLOSSARY FOR TRAINING SUPPORT

It is commonly accepted that for organizations to prosper they need to ensure that ensure that all its members have access to the knowledge, applications and services required to do their job. Thus, a system that delivers this knowledge to the right person (based on a specific context) at the right time, place and form should be implemented. Then, the practice of sharing such knowledge reveals and expands our individual and corporate competencies, which once understood can be reinforced and redistributed via e-Learning. In that case, a practical e-Learning development approach can facilitate and promote the development of competencies and knowledge in industry [327]. It follows the idea that training is one of the basic means of human resources development in business organizations, aiming to motivate employees, to develop their potential and to help them perform better. Consequently, having knowledge management and e-Learning working together is a powerful resource for users to meet their learning requirements in the form of correct and complete information access [328].

E-Learning makes use of Internet technologies to enhance knowledge transfer and performance in education and training. These technologies offer learners control over content, learning sequence, pace of learning, time, and often media, allowing them to tailor their experiences to meet their personal learning objectives [329]. Virtual education services should be accessible to everyone and for that reason is necessary to consider the specific needs of each user, and consequently adapt the process to solve those needs in a dynamic way. This can be done through the formal representation of knowledge that can be used to support adaptive e-Learning services, even when *diversity* characteristics are addressed.

Diversity is here addressed not only due to its relevance in today's policy, but also because industries would like to be better introduced in communities even acting as inclusion catalysts. Ignoring diversity issues costs time, money, and efficiency. Some of the consequences can include unhealthy tensions; loss of productivity because of increased conflict; inability to attract and retain talented people of all kinds; complaints and legal actions; and inability to retain valuable employees, resulting in lost investments in recruitment and training [330, 331].

7.1 ALTER-NATIVA

ALTER-NATIVA is an ALFA III project, which main goal is to promote higher Education in Latin America as a means to contribute to the economic and social development of the region [332]. The aims of the project were to: 1) provide education for everyone; 2) provide an environment of formation to professors when leading with persons with some disabilities (diversity condition); and 3) balance the inequalities of opportunities when accessing information. To meats these goals a platform to establish an international network of higher education institutions with recognized expertise in the areas of pedagogical education and development of information technology was implemented. Such platform is composed by four main elements: 1) ATutor, 2) COLABORA, 3) eLearning Repository, and 4) a reference KB.

Accessible Tutor known as ATutor[1] was the learning management system selected in the ALTER-NATIVA project to support two main and necessary process for achieving the projects objectives: 1) be the learning environment for delivering to teachers different courses in order to improve their abilities to attend the diversity in the learning process; 2) be the environment for creating learning experiences which include the creation of accessible open educational resources.

The creation of accessible Open Educational Resources (OER) in ATutor is also supported by the integration of TinyMCE[2]. TinyMCE is the selected web content authoring tool, which provides a friendly user interface developed completely in JavaScript language as well as a set of functionalities that allow teachers to create web pages in an intuitive way without worrying about HTML code, because the editor automatically creates HTML. TinyMCE was improved in order to provide a better support when the teachers create accessible and open educational resources; in particular, some accessibility issues were addressed in order to facilitate to teachers the attention of the Web Accessibility Guidelines 2.0 in the OER creation process.

[1] Atutor Learning Management Systems: http://www.atutor.ca/

[2] TinyMCE:https://www.tinymce.com/

COLABORA is an infrastructure that facilitates the collaboration activities of the ALTERNATIVA network. A community of practice is a group of people who share a concern or a set of problems of common interest about a topic and who deepen on their knowledge through ongoing interaction. Thus, the COLABORA platform has allowed the management of communities of practice in favor of attention to diversity. In COLABORA, tasks are elements that allow the achievement of activities, which are part of the interaction between collaborators of each community of practice.

ALTER-NATIVA has an e-Learning Repository of accessible learning objects, which aims to organize, store and retrieve educational resources produced by the members of the ALTERNATIVA network. The objects in the repository are organized into areas such as science, mathematics, language and communication. The repository has specific tools for labeling objects. However, it also allows import labeled object from the ATutor component. It uses the standard Learning Object Metadata (LOM) and IMS AccessForAll[3] to add accessibility information in learning objects. LOM is a multi-part standard that specifies Learning Object Metadata. The purpose of this multi-part standard is to facilitate search, evaluation, acquisition, and use of learning objects, for instance by learners or instructors or automated software processes [333].

The IMS AccessForAll intends to facilitate the formalization of Digital Resource Description (DRD) and Personal Needs and Preferences (PNP), through a meta-data specification. DRD is a lightweight metadata schema for describing and linking to digital objects. It is based on Qualified Dublin Core with local extensions. It is intended for use with simple digital objects as an alternative to more complex schemas [334]. PNP is a meta-data that intends to specify how resources are to be presented and structured; how resources are to be controlled and operated; and, what supplementary or alternative resources are to be supplied. All of this has the main goal to meet the needs (preferences) of learners with disabilities and of anyone in a disabling context, with the purpose of offering them an appropriate interaction with digital resources, including configuration of assistance technologies.

In a kind of conclusion both DRD and PNP are parts of the ISO/IEC 24751 standard, which derives from IMS[4] Learner Information Package Accessibility and IMS AccessForAll that intends to facilitate the matching of individual user needs and preferences with educational digital resources that meet those needs and preferences. [335].

[3]Web Content Accessibility Guidelines (WCAG) 2.0: https://www.w3.org/TR/WCAG20/
[4]https://www.imsglobal.org/

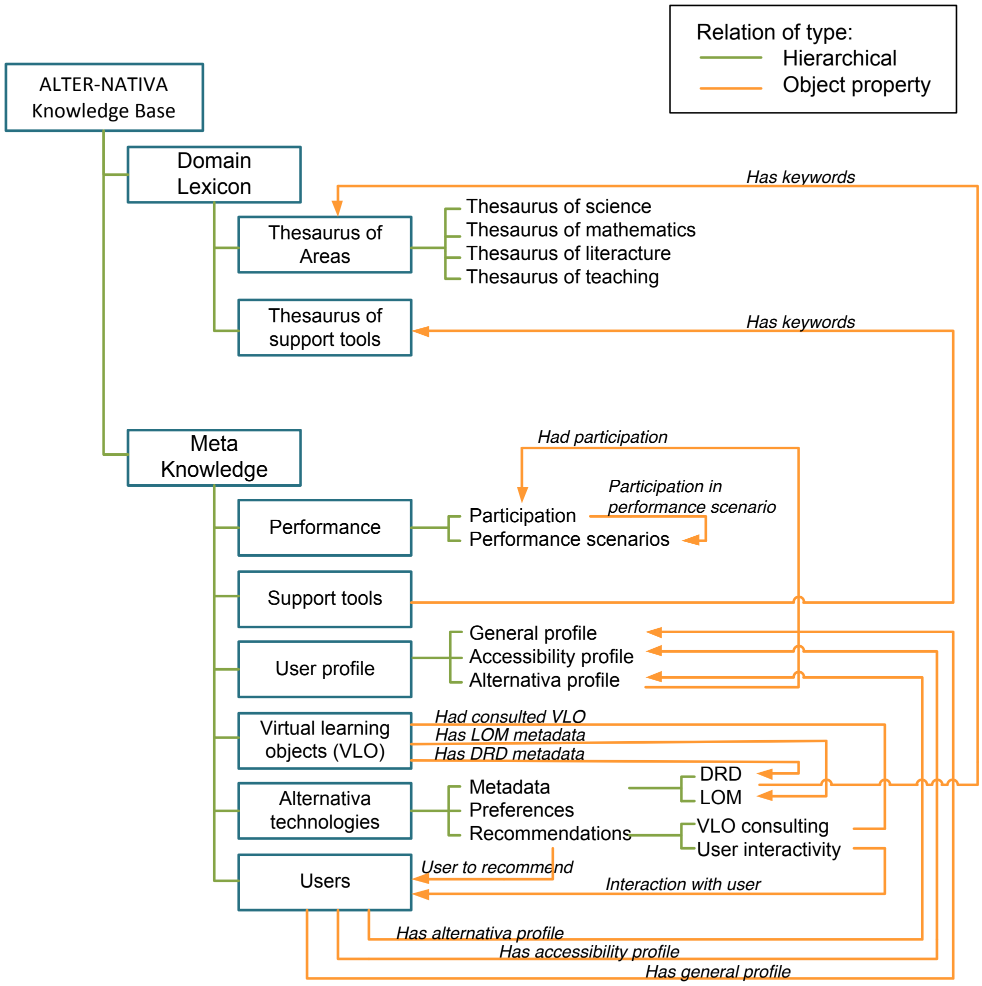


Figure 7.1: ALTER-NATIVA's knowledge base (an excerpt).

7.2 Knowledge Management Approach

The establishment of the ALTERNATIVA KB arises from the need to represent all the knowledge related with the project outcomes. It has three main objectives, the first is to enable professors to establish cooperations through COLABORA tool; the second is related to possibility of creating Virtual Learning Objects (VLOs) in ATutor, whose should be saved and categorized using a common vocabulary (domain lexicon) understood by all the community; and finally, the third is related to the searching and recommendations of people (e.g. professors) or VLOs available at the e-Learning Repository related to specific domain topics characteristics. To accomplish this, the KB is provided with specific web services and graphic user interfaces to facilitate knowledge management. The proposed knowledge management approach aims to facilitate community members to actively participate in the constant domain knowledge updating process. Two distinct parts compose the ALTER-NATIVA's KB: the 'Domain Lexicon' and the 'Meta Knowledge' 7.1 and will be explained in detail in the following sections.

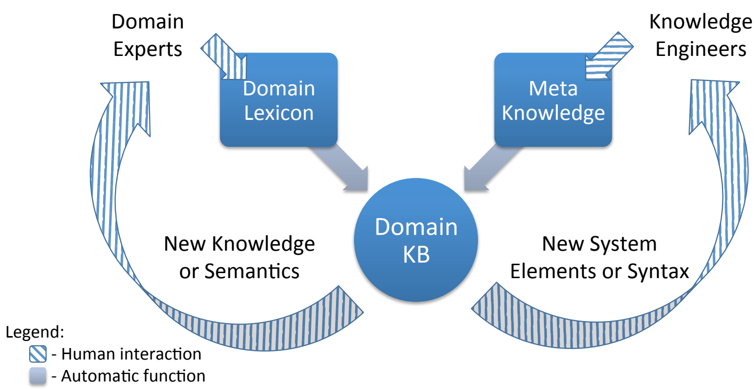


Figure 7.2: ALTER-NATIVA's Knowledge Management approach.

Figure 7.2 illustrates the adopted knowledge management approach. In this approach, the domain lexicon (thesaurus) is built directly by the domain experts whose contributed with concepts definitions (semantics). The knowledge engineers defined the meta- knowledge structure based on the standers mentioned in the previous section. The result is an ontology representing the system 'syntax' directly connected with the thesaurus provided by the domain experts (domain semantics). These two distinct branches compose the system Knowledge Base.

Both the Knowledge Base aggregate automatic synchronization. That means, if a domain experts contribute to a new concept, it will be smoothly available in the Knowledge Base for any further searching or reasoning services. In the same way, if a knowledge engineer contributes to a standard (meta-knowledge) update it will be also directly available to be used. Both branches of this approach follow the MENTOR[336] methodology as the procedure to build the reference knowledge from scratch. The domain lexicon branch follows it till the thesaurus and the meta-knowledge till the ontology building steps. The process followed for the thesaurus building has, three steps: (1) terminology gathering; (2) glossary building; and (3) thesaurus building 7.3.

The terminology-gathering step concerns the process of collecting all relevant terms or concepts in a specific domain previously defined. All the participants in the process should give their inputs. There is no rule from where the terms should come, since they are related to the domain established. All the concepts provided from contributors are acceptable in this step, as nobody has authority at the moment to erase another participant's term. Thus, the terms should be collected with reference to the contributor, in order to enable him/her to provide definitions during the next step.

The Glossary building step builds a glossary in the domain defined. It starts with annotations attribution to the terms collected in the step before (Terminology gathering). Then, each contributor should provide the annotations for his/her own terms. After having

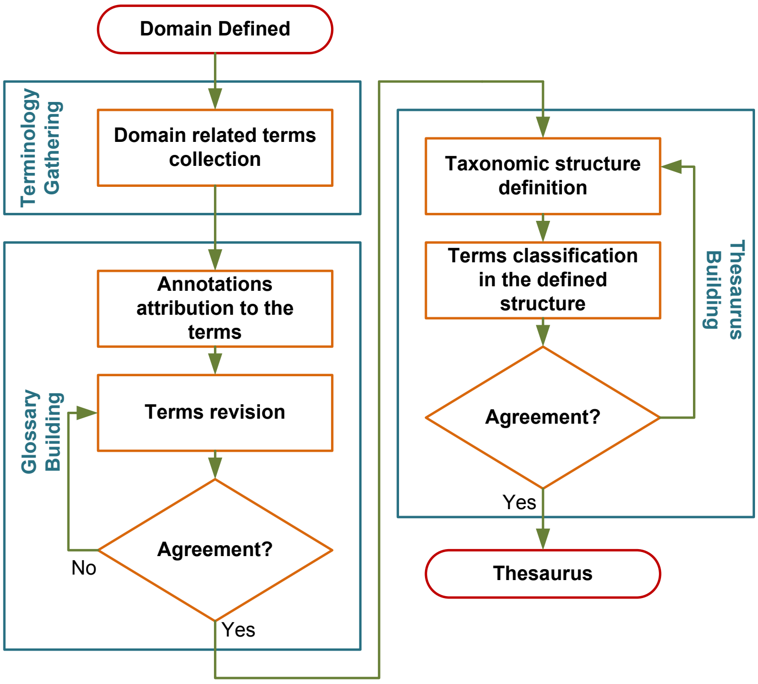


Figure 7.3: Thesaurus Building methodology [336].

all the terms provided with annotations, it proceeds to the terms revision cycle to reach a reference definition. The process for revision of terms can have four mismatches cases:

- **Existence of two syntactically different terms with the same meaning description** - the action is to adopt one of the terms for being the reference in such semantics meaning.
- **Existence of two syntactical equal terms with the same meaning description** - the action is to erase one of them.
- **Existence of syntactically different terms with two different meaning descriptions** - no action needed, both must be kept.
- **Existence of two syntactically equal terms with two different meaning descriptions** - the action is to consolidate all the provided descriptions together in one of them and erase the other. In such a case, a new term could be proposed to the list if there is no agreement in the conjunction of the input descriptions and if the term to be born is not present in the terminology list.

After a careful revision of all the terms with a successful agreement in their meaning consolidation, the glossary is defined from the terminology list in the domain specified. The thesaurus-building step is composed of a cycle where the knowledge engineers and the domain experts define a taxonomic structure from the glossary terms, establishing some as thesaurus node terms. Afterwards, the other terms are classified into semantic proper paths in the existent taxonomic structure, until reaching the thesaurus leafs. If there is an agreement in the structure and in the terms classified, the thesaurus is defined. If not, the cycle is repeated.

7.3 Thesaurus building by domain experts

The 'Domain Lexicon' part of the proposed KB is dedicated to represent the lexicon of ALTERNATIVA. Thus, its main purpose is to have a set of reference concepts and meanings about its domain represented in thesauri. The ALTER-NATIVA thesauri represent the information related to the concepts of four distinct areas, namely Languages, Mathematics, Science, Teaching and Supporting Tools. The definition of each concept, the creation or agreement date and the author(s) of such information compose them. Each of the thesauri is available in a wiki kind of interface enable a public access to its contents. Figure 7.4 illustrates a Teaching ans Supporting tools Tools concept page. To facilitate the insertion of concepts bu the domain experts, a set of extensions were developed. Thus, users can update the thesaurus information as easily as fill a form. The interoperability establishment between the wiki Data Base (DB) and the ALTER-NATIVA KB require a synchronization interface. It was developed using the approach for knowledge management using semantic wiki modules proposed in the previous chapter. In the following subsections are explained in detail three actions that represent thesaurus typical changes: 1) concept deletion; 2) concept creation; and 3) concept edition.

7.3.1 Delete a concept

The elimination of a concept directly affects the structure of the thesaurus tree. If the deleted concept does not have children associated, it is just removed from the thesaurus tree. But if the deleted concept has children associated it is necessary to take some actions to reflect the wiki status in the KB. In this case, the children of the deleted concept are positioned in the thesaurus root. In 7.5, it is shown the initial and final status of the thesaurus tree after the deletion of a concept with children. The example illustrated is related to the deletion of the 'Concept b', which in the initial thesaurus tree had the children 'Concept c' and 'Concept d'. In the final state is possible to see that these concepts are shifted to the root (Thing).

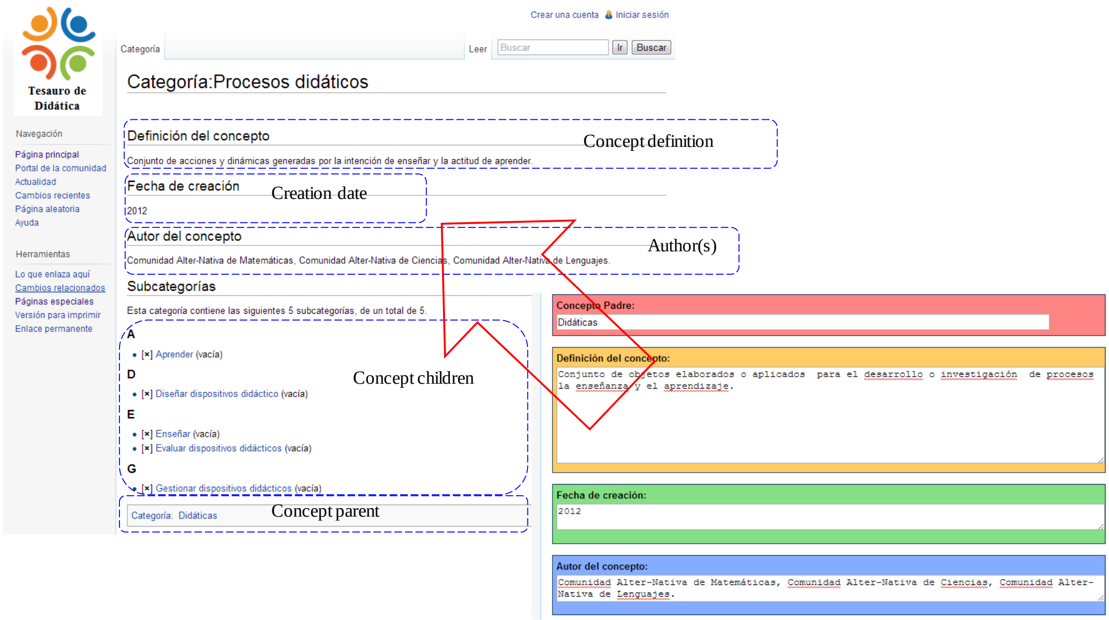


Figure 7.4: Form to build a concept page.

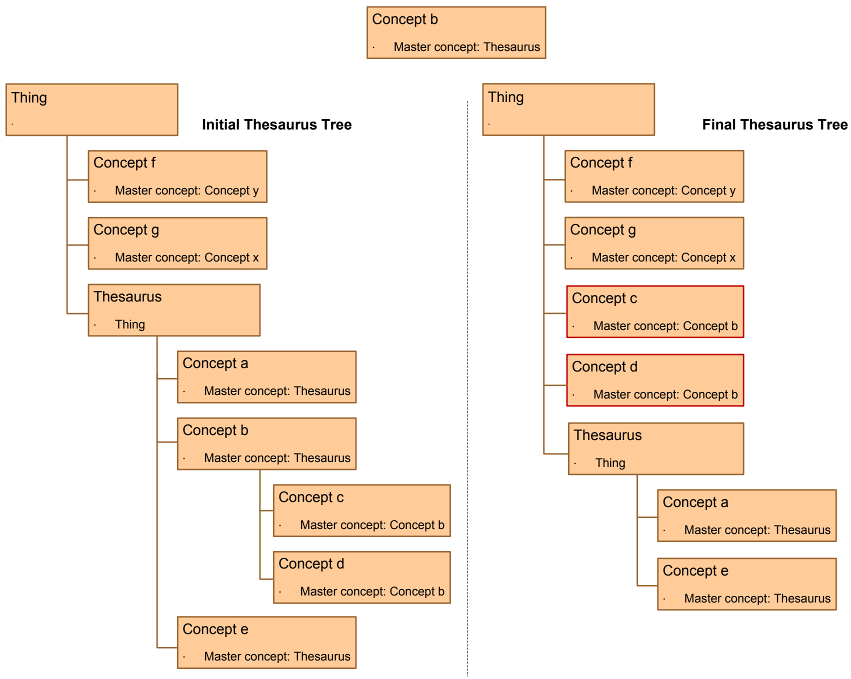


Figure 7.5: Delete a concept with children associated.

7.3.2 Create a concept

As in the concept deletion action, the creation of a concept directly affects the structure of the thesaurus tree. So it is necessary to rearrange the thesauri tree to handle the three possible scenarios explained in table 7.1 that contains a summary explanation of all conjugations of possible atomic situations and which are the steps to execute by the synchronization module.

Table 7.1: Use case of concept creation

Parent exists?	There is any root concept that fits to be this concept child?	Action
Yes	Yes	• Put the new concept under the class referred as master (parent) • Put the other concept root tree under the new introduced concept
	No	• Put the new concept under the class referred as master (parent)
No	Yes	• Put the new concept under the class Thing • Put the other concept root tree under the new introduced concept
	No	• Put the new concept under the class Thing

The Figure 7.6 illustrates the case where the inserted concept ('Concept h') has another existing concept ('Concept b') defined as its master (parent) concept. In this case, the action to be taken is to insert the 'Concept h' has child of 'Concept b'.

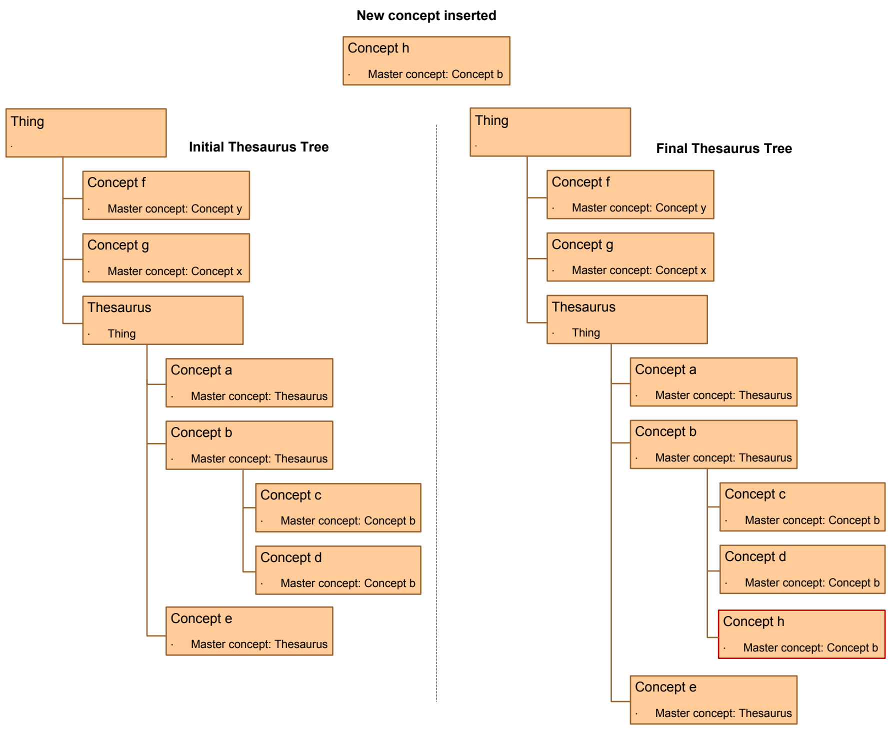


Figure 7.6: Creation of a concept that points to an existing concept master.

The Figure 7.7 illustrates the case where the inserted concept (‘Concept h’) has (by mistake) a non-existing concept (‘Concept z’) defined as its master (parent) concept. In this case, since the ‘Concept z’ does not exist in the initial state of the thesaurus tree, the action to take is to insert the ‘Concept h’ has child of the ‘Thing’ class (to put it in the root).

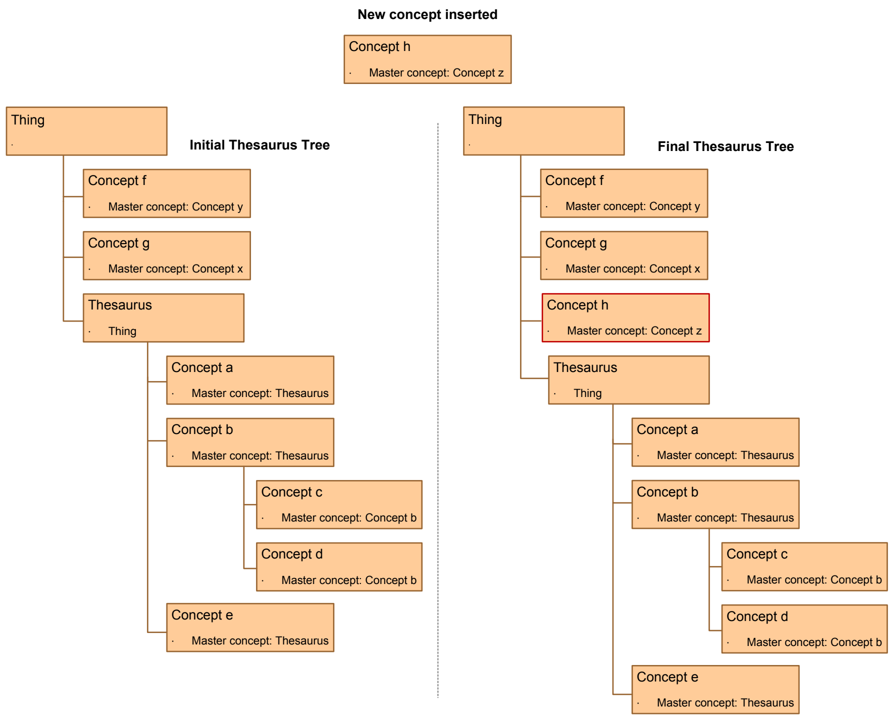


Figure 7.7: Creation of a concept that points to a non-existing concept master.

It is also possible to occur the case where the inserted concept (‘Concept y’) was already defined as a master concept of an already concept existent in the root (‘Concept f’). Figure 7.8 illustrates this mentioned case. When this new concept (‘Concept y’) is inserted in the tree, ‘Concept f’ will be placed bellow it. Figure 7.8 also illustrates a case where two atomic cases occurred simultaneously.

7.3.2.1 Edit a concept

When a concept is edited, if the master concept is not changed, it is only a matter to save (update) the new information in the KB Thesaurus. If the master concept changes, may occur two situations: 1) the new master concept does not exist in the Thesaurus tree, so is necessary to execute the steps illustrated on Figure 7.7; or 2) the new master concept exists in the tree and is necessary to execute the steps illustrated on Figure 7.6.

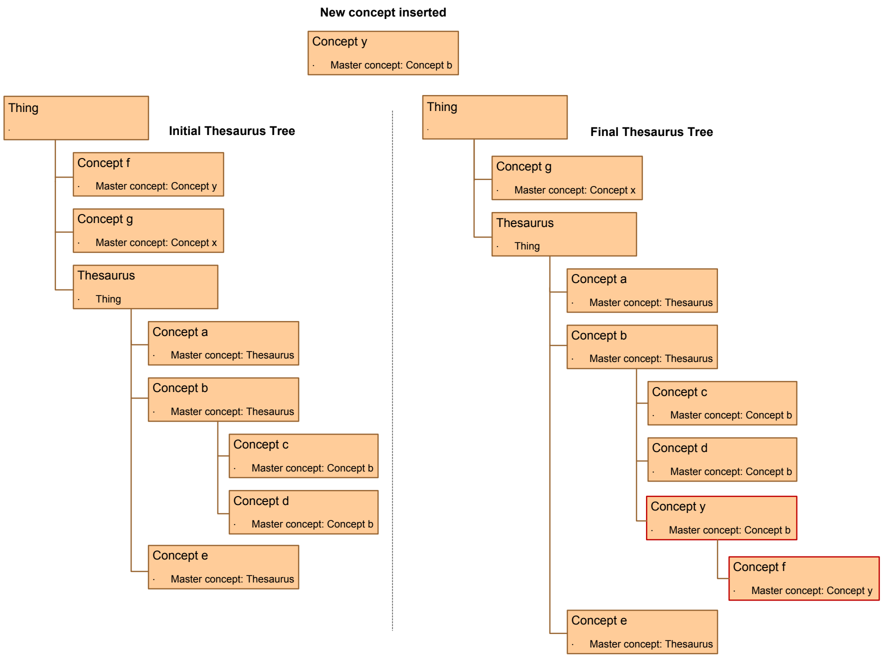


Figure 7.8: The created concept is referred as a master concept already.

7.4 Meta-Knowledge Management

The meta-knowledge is managed by knowledge engineers (ontology managers), which can directly edit the ontology through the protégé editor tool. In this case, two solutions were setup, one offline and another online. The offline way is through the protégé regular tool that connects directly to the ALTER-NATIVA server, which then can edit and upload changes in the OWL file. The online version uses the WebProtégé plug-in, which is a Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web (see Figure 7.9). WebProtégé includes a set of predefined tabs, which contain the most popular functionality in the protégé desktop editor [337].

7.5 ALTER-NATIVA Services

The ALTER-NATIVA KB web services are responsible for giving worldwide users the mechanism of Knowledge Management and reasoning trough the Web. These services use JENA libraries, a java API for managing OWL ontologies. The developed services are classified in two types: the Knowledge Base and Recommendation services.

7.5.1 Knowledge Base Services

KB services are categorized in two groups: 1) Knowledge Management Services - related to the knowledge edition; and 2) Knowledge Query Services - related to the knowledge

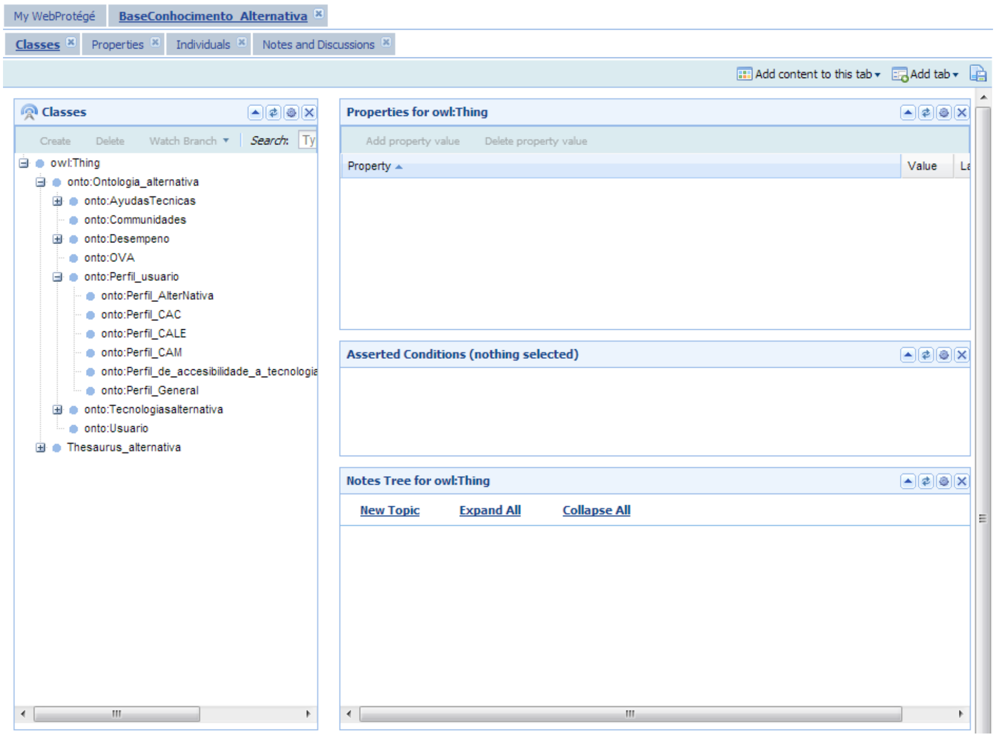


Figure 7.9: ALTER-NATIVA's WebProtégé management interface.

reasoning. These services allow used to update and consult knowledge about consulted VLOs, users profile, performance scenarios, thesaurus, and users' interaction between them and VLOs.

7.5.2 Recommendation Services

The recommendation services have as goal to support users to search for another user or VOLs to interact with. The recommendation of VLOs could be done concerning a user specific interest or diversity characteristic, while the users' recommendation could be focused on a specific profile (e.g. teaching experience or topics of interest). The information/knowledge used to achieve this result is related to: users' historic information (past interactions with ALTER-NATIVA resources), VLOs metadata and keywords, and user profile. Finally, ATutor uses recommendation services to get appropriate VLOs in the development of courses or other VLOs. COLABORA mainly uses these services to facilitate its users on the searching of other users to interact with.

7.5.2.1 Recommendation of Users

As mentioned before, the recommendation of users concerns the suggestion of users to interact with based on:

- Similar teaching experience;
- A specific similar profile characteristics;
- Number of interactions between two determined users;
- Number of common profile characteristics between two users;
- A combination of number of interactions with common user profiles.

When the recommendations deal with more than one characteristic, the process of choosing the most appropriated item is by calculating the one that has more similarities to the set of chosen target characteristics. A function to be used in such process, which calculates the strength of each item with potential to be recommended is used.

7.5.2.2 Recommendation of VLOs

The services of recommendation of VLOs are based on:

- Higher rating;
- Most used;
- Number of similarities with other VLOs consulted by the user (keywords);
- Combination of the previous patters.

To automatically suggest a VLO is necessary to have a log record of the consulted VLOs. Such information is formally represented in the KB, which interrelates the VLOs with the users' interested topics, the number of times each user consulted the it, and his/her associated evaluation rating about such VLO (1- Good or 0 - Bad).

In the top left of Figure 7.10, it is shown how the algorithm for suggestion of similar VLOs based on its characterization keywords. It starts by collecting the identifiers of all the VLO consulted by a user. Then, the keywords that characterize those VLOs (top right) are collected. After this, a new set composed by VLOs other than the ones already consulted by the user and that have one or more common keywords is built. Finally, the recommendation result is represented by the bottom table, which presents the VLO with common keywords, ordered by the high number of keywords in common.

The combined recommendation of VLOs is done based on the previous recommendation services for VLO aggregation. Thus, this service returns the VLO with the higher strength based on the conjunction of the various evaluation patterns characteristics as number of times consulted, rating and number of similarities.

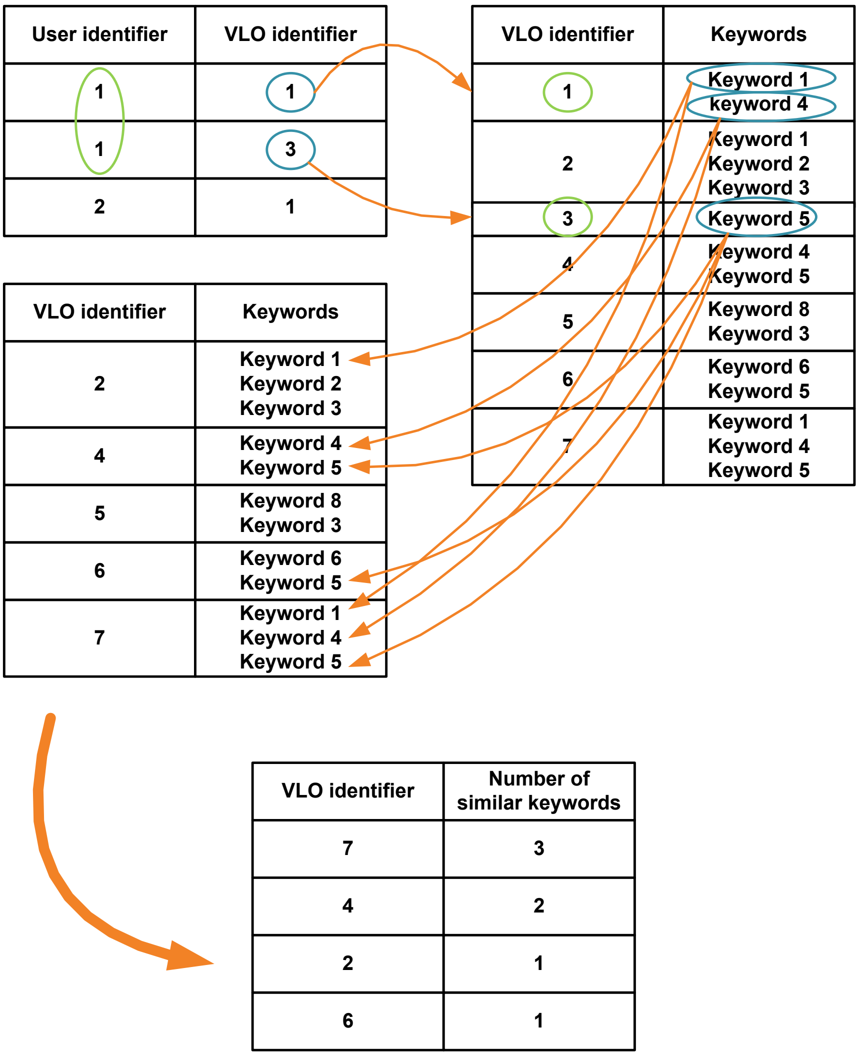


Figure 7.10: ALTER-NATIVA's WebProtégé management interface.

7.6 Discussion

In this research report it is presented a KB developed to represent the ALTER-NATIVA's knowledge. It followed a taxis path to build knowledge and fits as the central element in the proposed Knowledge Management approach. As stated in along this thesis, by defining shared and common domain theories, ontologies help both people and machines to communicate concisely, supporting the exchange of semantics and not only syntax. In align with this, the ontology, to which the developed KB relates to, provide a lexicon and its meanings that describe the ALTER-NATIVA system and domain. It formally represents not only the domain knowledge, but also the system characteristics and objectives in the support of the ALTER-NATIVA's network establishment. Therefore, it integrates technical solutions able to harmonize a community knowledge view providing a semantic interoperable basis to smoothly support the creation of virtual collaborative communities and resources, which has been practiced in the ALTER-NATIVA community.

In this project, none ontology learning methodology was used to build neither the thesaurus or meta-knowledge. However, the author considers that the knowledge base built from scratch could be enriched using ontology learning techniques and increase the knowledge range and improve both users and VOLs recommendation.

Additionally, it was noticed that this kind of knowledge organization systems enables context awareness abilities. This has been reached through available recommendation services. The knowing how users (i.e. professors and trainees) and their profiles are related with the VLOs and what kind of tools can be used to reduce the learning vulnerability (diversity) or to enhance accessibility of trainees, is a contextual information that the system could use to react accordingly. In this case, a better suggestion on the training implementation (e.g. an appropriate VLO suggestion) can be provided.

CHAPTER 8

REQUIREMENTS MANAGEMENT USING SEMANTIC WIKI MODULES

Focusing on the manufacturing domain, the types of requirements that are usually considered are: Business Requirements, Functional Requirements, User-Interface Requirements, Transition Requirements, and Non-Functional Requirements. The creation of an analysis matrix such as the one proposed by Altshuller in [338] and illustrated in Table 8.1 allows to compare the state-of-art elicitation approaches (at the top) depending on its compliance against some specific preconditions (at the left). This matrix made possible to conclude that the brainstorming and brainwriting creative technique are able to cover most of the solution related preconditions and type of requirements. However, one inherent characteristic that can be considered as a disadvantage is the low compliance related to the independency of stakeholder's location. This characteristic is the main catalyzer for the authors work: to find a solution able to take the best of brainstorming and brainwriting techniques, and that reduces the impact of the stakeholder's location.

Brainstorming can be considered as a process for generating creative ideas and solutions through intensive and freewheeling group discussion. Every participant is encouraged to think aloud and suggest as many ideas as possible, no matter seemingly how outlandish or bizarre. Analysis, discussion, or criticism of the aired ideas is allowed only when the session is over and evaluation session begins [60]. This technique is the most widely adopted process for generate creative ideas within organizations [340].

In contrast to the oral sharing of ideas in groups during brainstorming, **brainwriting** involves a group of people *silently writing* and sharing their written ideas. Usually, for brainstorming and brainwriting techniques, is assumed that the group elements are in the

Table 8.1: Applicability for Elicitation Approaches following the work of [338] and [339].

● : high compliance
⦿ : partly compliant
○ : low compliance

	Creativity techniques						Observation Techniques					Interview Techniques				
	Serious Games	Brainstorming	Brainwriting	Six-Hats	TRIZ	Osborn Checklist	Qualitative Observatio	Document Analysis	Ethnography	Netnography	Reverse Engineering	Questionnaire	One-on-one Interview	Focus Group Interview	Requirements Worksho	Prototyping
Stakeholder related factors																
Independent of Stakeholder location	○	○	⦿	○	⦿	⦿	⦿	●	⦿	●	●	●	⦿	○	○	⦿
Independent of Stakeholder temporal availability	○	⦿	⦿	○	⦿	⦿	●	●	⦿	●	●	●	⦿	⦿	○	●
Information Exchange between stakeholders	⦿	●	⦿	●	⦿	⦿	○	○	○	○	○	○	○	⦿	●	⦿
Solution related factors																
Suitable for complex solutions	●	⦿	⦿	⦿	○	○	●	●	●	⦿	⦿	○	⦿	●	●	⦿
Supports radical innovation	●	●	⦿	⦿	⦿	○	○	○	○	○	○	⦿	●	●	●	⦿
Produces fast Results	⦿	●	●	⦿	○	○	○	○	○	⦿	○	⦿	⦿	●	●	○
Types of Requirements																
Business	⦿	●	●	⦿	⦿	●	⦿	○	⦿	⦿	○	⦿	⦿	●	●	⦿
Functional	●	⦿	⦿	⦿	⦿	⦿	⦿	●	○	○	●	●	⦿	⦿	⦿	●
Non-functional	●	●	●	●	⦿	⦿	⦿	○	●	●	○	○	⦿	⦿	⦿	⦿
User-interface	⦿	●	●	⦿	⦿	⦿	⦿	⦿	●	●	⦿	○	⦿	⦿	⦿	●
Transition	○	●	●	⦿	⦿	⦿	○	○	○	○	○	⦿	●	●	●	⦿

same room, sharing ideas. However, due to the dependency of stakeholder location, it may compromise their participation due to lack of availability or even will to meet other group elements. The implementation of a solution that can be accessed remotely and grounds a collaborative environment could reduce the problems associated to the lack of end-users input. Accordingly with the study of Fig. 8.1, other techniques that show good performance in relation to the presented preconditions are **focus group interviews**, **questionnaires** and **requirements workshops**. Those techniques are currently used to incentive end-users and developers in the requirements elicitation process. Thus, the output of these elicitation techniques should be written and used as input of requirements management tools, starting the brainwriting technique from here.

8.1 Going beyond traditional requirements management tools

Requirements management entails being able: 1) to relate many different documents to obtain a synoptic view of these document relations; 2) to retrieve information from within those documents; 3) to create special document view; 4) to handle changes made across the set of documents in a consistent manner; and 5) to accommodate diverse documents structuring requirements and document types [288]. Meeting those demands, several tools have been developed and released almost every day. Indeed, information about several commercial requirements management tools can be accessed from the INCOSE Survey [341]. Some of the most well-known commercial tools are DOORS, RDD-100, RequisitePro, and CaliberRM. Beside those professional tools there are others largely used due its simplicity and lake of commercial costs. Examples are wikis, Microsoft Excel and Microsoft Word. Wikis can provide great value as a requirements management tool, including the incentive for stakeholder participation, the support for more consistent documentation through their simple and consistent layout, the improved search and traceability, the support for base lining and versioning requirements, and the support for collaborative requirements review [342].

Simple tools such as Excel and Word can be used to quickly obtain good results in determining requirements. Experience has shown that when formulating requirements, it is important to choose a practical approach that suits the company's organization and available resources. Good results can then be obtained with surprisingly simple tools. Furthermore, these tools are easy to use and do not generate any additional expenses. However, Excel and Word have its limits. For example, they require other tools for collaborative management, for instance Google Docs.

Table 8.2: Requirement Management tools comparative study (based on the work of INCOSE [341]).

● : high compliance
◉ : partly compliant
○ : low compliance

	DOORS	RDD-100	RequisitePro	Caliber-RM	Wikis	Word and Excel
Preconditions						
Input documents enrichment/analysis	●	●	●	●	●	●
Input documents change/comparsion	●	●	●	●	●	○
Requirements Classification	●	●	●	●	◉	◉
Requirements Derivation (req. to req., req. to analysis)	●	●	●	●	◉	○
History of requirements changes, who, what, when, where, why, how	●	●	●	●	●	○
Visibility into existing links from source to implementation - i.e. Follow the links	●	●	●	●	◉	○
Access control	●	●	◉	●	●	○
Web browser interface?	○	○	●	●	●	○
Single user/ multiple user concurrent users	●	●	●	●	●	○
Commercial?	●	●	●	●	○	○

The summary of Table 2 was built based on the previous tools' descriptions and on the survey made by INCOSE [341]. It represents the compliance of well known requirements management tools, represented on the top, in relation to the characteristics that define them as a good tool. It is possible to conclude that commercial tools have a high compliance in relation to the most of the characteristics considered. However, some companies may not be willing or able to afford them. In that case, solutions like wikis can be a better choice. Like represented in the table, they show similar results in comparison to the commercial tools. However, they are still not fully compliant to some characteristics like: Requirements Classification; Requirements Derivation; and Visibility into existing links from source to implementation. Thus, the work here presented is focused on the adaptation of wiki-based modules so they can meet these characteristics and be able to compete with commercial tools in requirements management.

8.1.1 Semantic Wikis for Requirements Management Support

Wikipedia is an open encyclopedia characterized by being collaboratively edited by its users. This software provides so called categories that are used to classify articles and other pages. The process of assigning categories to wikis page articles is a kind of collaborative tagging like in a folksonomy, but in this case, it is related to the characterization of requirements categories and other classifiers. However, despite this strong usability, wikis are not fully prepared to be integrated with other information systems. Thus, the proposed solution suggests the Semantic Web concept [343]. Semantic Web can be seen as a set of technologies that help knowledge sharing across the Web between different applications. Ontologies based technologies play a prominent role in the Semantic Web [344]. They make possible the widespread publication of machine understandable data, opening opportunities to automatic knowledge reasoning. Web 2.0 focuses in collaboration of users and sharing information between them. Together with the Semantic Web they contribute for software applications with knowledge from web software, e.g. wikis.

Folksonomies aggregates both Web 2.0 and Semantic Web characteristics being great for categorizing documents and resources in a collaborative way [345]. They arise when a large number of people are interested in particular information and are encouraged to describe it [343]. Since ontologies specify a conceptualization of a domain in terms of concepts, attributes, and relations [346], Semantic Web and its applications rely heavily on formal ontologies to structure data for comprehensive and transportable machine understanding. Semantic wikis are a characteristic of the Semantic Web and enrich wiki systems for collaborative content management with semantic technologies [311] allowing domain experts and ontologies to cooperate in one system while wiki pages are presented in a human-readable format in parallel to the formal ontologies.

Several requirements management works supported by wiki modules can be found in the literature. In [347], the authors promoted the collaboration between teams through the implementation of templates for communication establishment. However, these templates can easily be edited by stakeholders breaking the replication of requirements structure. This can be a rollback in what concerns requirements classification. In [348], the Wiki-WinWin approach creates a sequence of steps and instructions to guide stakeholders on requirements management work. During each step, the system displays one or more tools with whose the team can generate, organize, and evaluate concepts and information. An identified problem is related to the need of a *'shaper'* role which function is to integrate, distil, organize & rewrite contributions of others. Hence the proper classification of requirements needs to be always supported a human entity. This could be avoided if the information was already inserted properly in the wiki interface. This is one of the challenges that the authors addressed in the presented work. They also supported the derivation of requirements and their characteristics using proper templates for collaborative tagging. Thus, links can be generated between the several entities involved in requirements management (e.g. user requirements, authors, acceptance criteria). That functionality improved both requirements derivation and links visualization between requirements and implementations.

8.2 Requirements Management Methodology

The proposed methodology for requirements management starts with a **Preparation** step, where the selection of the more appropriated elicitation mothods is a key to reduce the lack of inputs from stakeholders and developers. Several elicitation methods can be selected and used together to achieve better results. In this work, the authors considered brainwriting as a must since it cover the most of the precondition for good requirements elicitation.

After this preparation step, the requirements management methodology represented in Fig. 8.1 starts with the **Elicitation** process. In this process, the requirements begin as ideas or concepts. These can be defined by a single individual, but usually are defined from group's interactions [349]. The novelty of the proposed methodology is that the elicited requirements are written as user stories describing in a comprehensible language what is expected from the system.

Analysis concerns reviewing, analyzing requirements in detail, and negotiating with stakeholders on which requirements should be considered [350]. Thus, mechanisms that allow these interactions should be implemented. These include requirements approval and refinement. The analysis process also encompasses the transition from user to technical requirements and validation criteria definition. It means that, during user requirements analysis, the end-users should be able to provide the acceptance criteria for their requirements. Thus, a requirement is fulfill when it behaves accordingly with its acceptance

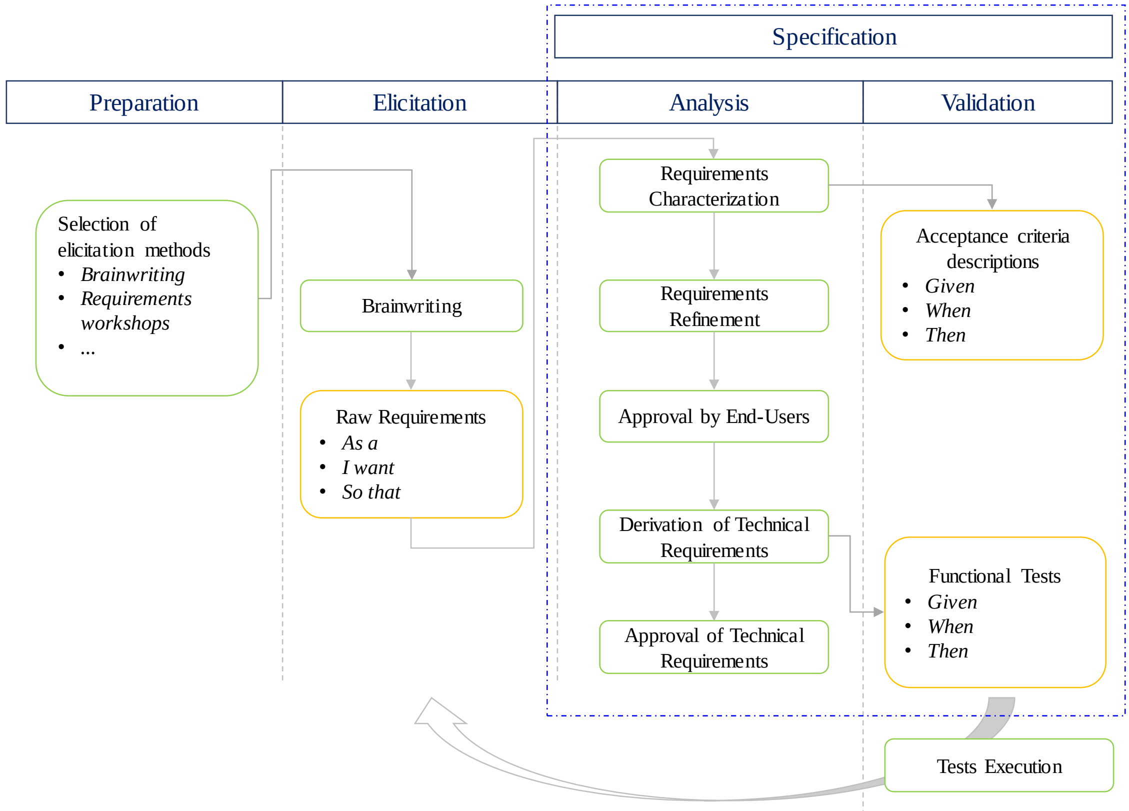


Figure 8.1: Requirements Engineering Methodology.

criteria; and the technical team should also be able to define the functional tests for the derived technical requirements.

Requirements **Specification** describes the phase, where the requirements are brought into a suitable and unambiguous form [351]. The idea of this phase is to make requirements readable and understandable by anyone that was not involved in the elicitation and/or analysis process. The Requirements Specification phase documents the agreed requirements at a certain level [350]. This is a core process in the proposed methodology. The authors want to came up with an approach that makes sure that, since the moment requirements and validation definitions are elicited and analyzed, these are already described in an unambiguous way. This can be done, as an example, using strict forms that only allow requirements to be written in a certain way. Finally, requirements **Validation** is done by checking the compliance of user and technical requirements with their defined acceptance criteria and functional tests correspondingly.

8.3 BDD based Requirements Management Framework

The proposed BDD based requirements management framework was implemented adopting the work presented by Marques et al. in [352]. In that work, the authors argued that the more communication, involvement, and interaction of people, more is the chance for organizations to expose tacit knowledge residing in individuals' heads. To meet this concern, they proposed to use wiki modules for domain experts to expose their knowledge.

The novel proposed framework adapted the previous solution to keep-up with requirements management main steps: Elicitation; Analysis; Specification; and Validation (see Fig. 8.1). Similarly to the work that grounded the proposed framework, it relies on the collaborative aspects of Semantic wikis to allow collaborative contributions and further feedback from interested parties, which might not have the technical skill for complex solutions. Furthermore, due to wikis browser interface, this works allowed stakeholders to collaboratively participate in requirements management independently of their physical location.

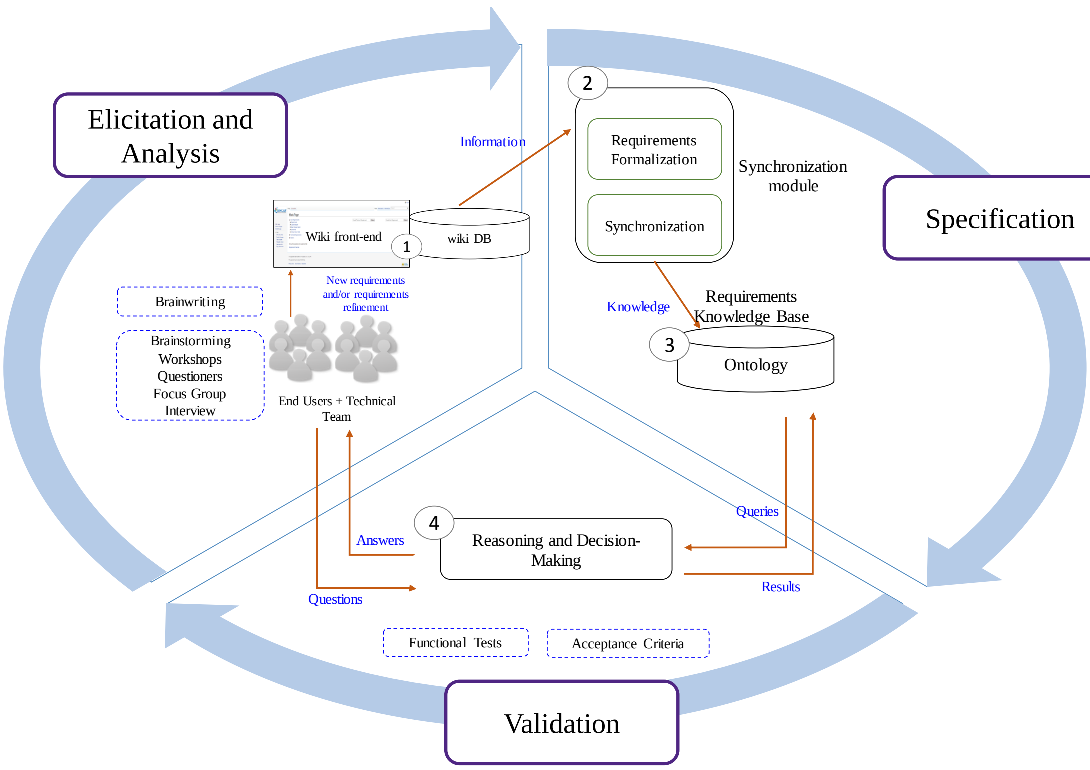


Figure 8.2: Requirements Management Framework.

In the presented work, the classical wiki template was edited to incorporate as many templates as the entities that the system intends to represent (user and technical requirements, authors, etc.). This is the major contribution in terms of wiki front-ends adaptation for requirements management purpose. The mentioned templates were implemented accordingly with the BDD specifications of section 5.2.2.1. The usage of this templates allowed the replicability of requirements structure improving their classification and further analysis.

8.3.1 Instantiation of the Requirements Management Framework

Fig. 8.2 depicts the proposed framework for requirements management which is composed by four modules: 1) wiki front-end; 2) Synchronization module; 3) Ontology; and 4) Reasoning and Decision-Making Module. These are used to implement the sub processes of Fig. 8.1 methodology, namely: Elicitation, Analysis, Specification, and Validation.

8.3.1.1 Requirements Elicitation and Analysis

The requirements management process is established as a cyclic process, since requirements are constantly refined while domain participants increase their contextual knowledge about the system. Consequently, requirements become even more concrete, detailed and complex during the course of a product or service development. It means that, while the contextual awareness of the system increases, the four requirement management processes should be followed cyclically, corresponding to the refinement of the requirements system. Thus, important characteristics that a requirements management tool should address are documents enrichment, modifiability, changes tracking and comparison. Accordingly with the study provided in section 8.1, the **wiki-front end** is compliant with all these characteristics and consequently were selected to be used in requirements elicitation.

8.3.1.2 Requirements Specification

Wikis are known by being collaboratively edited by domain experts based on the knowledge consulted. However, the content of wiki-based front-ends is characterized by being human readable only. This means that its content is not formalized to facilitate computerized use (e.g. reasoning). To overcome this issue, a **synchronization module**, composed by two sub-modules is used:

- **Requirements Formalization** - It is implemented following the requirements formalization methodology of Fig. 8.3. Using this methodology, wikis interface can be adapted to support the structured insertion of requirements from stakeholders despite knowing nothing about its syntax. This methodology is explained in detail in the following subsection.

- **Synchronization** - This module uses Wikipedia Data Base to detect any changes that have occurred in the front-end since it was last run and then updates the Requirements Knowledge Base accordingly. This module runs periodically thanks to a "cron job".

Finally, under the Requirements Specification process, the **Requirements Knowledge Base** is used to represent the knowledge about requirements and related elements like acceptance criteria, functional tests, authors, etc.

Requirements Formalization Methodology

This methodology, illustrated in Fig. 8.3, encompasses the design of an adapted wiki front-end and the ontology which will handle requirements associated knowledge. The first phase of the methodology consists in the knowledge base's classes taxonomy establishment. This phase is composed by the steps 0-2. In step 0 is made the preparation of the platform to handle the domain users' knowledge. In the step 0.a) is made the software installation, and in step 0.b) is created the wiki root class in the ontology in order to represent the knowledge about the requirements.

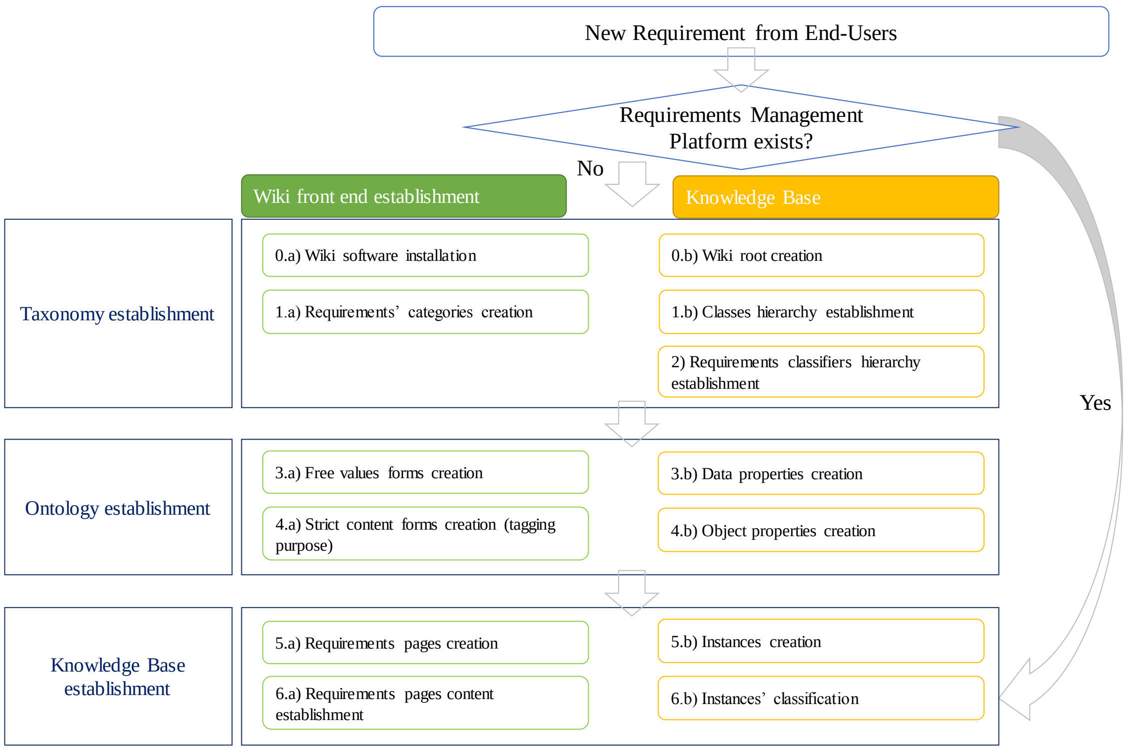


Figure 8.3: Methodology for wiki-based front-end contents formalization.

The process of assigning categories to other categories, in the proposed methodologies (step 1.b)), is used to build the ontology's instances taxonomy, being the tagging between them handled as the ontological relation 'is a'. In the step 1.a) are created the necessary category pages to allow the elicitation and analyses of requirements. The classification of requirements' contents can be facilitated if a classifiers taxonomy of those contents is defined (step 2). This allows to better structure the gathered knowledge and visualize relations between Knowledge Base's instances.

To guarantee the same structure of the wiki articles under a category, the wiki-based front-end was adapted through the development of forms. They encompass the possibility of insert free valued texts like requirements description sentences (step 3.a), and values that need to be inserted in a strict way to allow proper tagging between wiki pages (step 4.a)). With that specification of wiki pages (requirements), is possible to proceed with the steps 3.b) and 4.b) of the methodology. In those steps, for each article section is created a data property or object property to represent the elicited knowledge in the ontology. The object properties allow the connection of the classes under the wiki root class and those under the classifiers taxonomy previously defined (step 4.b)). Data properties will represent knowledge that is not under that taxonomy (step 3.b)).

The process of assigning articles to categories, in the proposed methodology (steps 5 and 6) will be used to instantiate the ontology. This is done by creating an instance under the class with the article's category name (step 5). Then, based on HTML analysis of articles' content, the knowledge of its sections can be represented in the data and object properties of the previously created instance (step 6).

8.3.1.3 Requirements Validation

The requirements validation is done through the compliance analysis in relation to the **acceptance criteria** and **functional tests gathered**. However, the authors also considered that requirements validation could be complemented by the reasoning allowed from requirements specification. Thus, requirements validation is supported by the **Reasoning and Decision Making Module**. This module is able to provide to the community structured and contextual information about requirements. One example is: *Which are the most relevant architectural components to be added to the platform concerning the requirements priority?* The capability of the ontology to answer these questions allows some coordination efforts to the consortium, namely: Prioritize the requirements that are indicated as a priority; Prioritize the implementation of architectural components that embrace a large number of user requirements. The reasoning on the formalized requirements can also be useful after architectural components' implementation. Let's consider that some management activity is need regarding one or more components. Then, the reasoning and decision-making module allows to verify which are the requirements that are compromised and notify the corresponding authors.

8.4 Implementation of the Requirements Management Framework in Industry

The Sensing Liquid Enterprise[1] concept has been introduced by the Future Internet Enterprise Systems (FInES) Research Cluster with the support of the European Commission. The FInES community acknowledge the fact that businesses are facing unprecedented challenges, given the current economic crisis, but also more systemic changes related in particular to the shortness of resources, environmental changes, and ever changing societal needs. Therefore our enterprises are in need of innovative ideas to adapt to these changes and remain competitive, or sometimes, even simply survive in the digital era. The Sensing Enterprise concept is an attempt to reconcile traditional (non ‘pure’ Internet) organisations with the tremendous possibilities offered by the cyber worlds (from the clouds to the dust) [353, 354].

The OSMOsis applications for the Sensing Enterprise (OSMOSE)[2] project has the main objective of developing a reference architecture, a middleware and some prototypal applications for the Sensing-Liquid Enterprise, by interconnecting Real, Digital, and Virtual Worlds in the same way as a semi-permeable membrane permits the flow of liquid particles through itself [355]. The worlds represent a way of organizing the structure of an entire manufacturing enterprise, and the business applications in three types of data management environments: **Real World** - related to data that comes directly from devices that is handled by physical components; **Digital World** - related to data management available in data and knowledge bases or Internet (big data); and **Virtual World** - related to specific management of data with the support of artificial intelligence related programs for specific simulations.

The approach presented in this paper follows the necessity of implementing a requirements management tool able to handle the sensing liquid enterprise transition with the Requirements Engineering Methodology presented in Fig. 8.1.

8.4.1 Wiki front-end

Requirements management tools can be considered as generic. That means they need to be configured to support specific requirements engineering and system development processes. That configuration can be supported by the creation of document templates, schemes of attribute and relation types, and document views. That kind of solution, if applied to wikis, could improve requirements characterization and requirements derivation. Thus, to facilitate the insertion of requirements in the wiki front-end, several extensions

[1] http://finespedia.epu.ntua.gr/Sensing_Enterprise.html
[2] http://cordis.europa.eu/project/rcn/189013_en.html

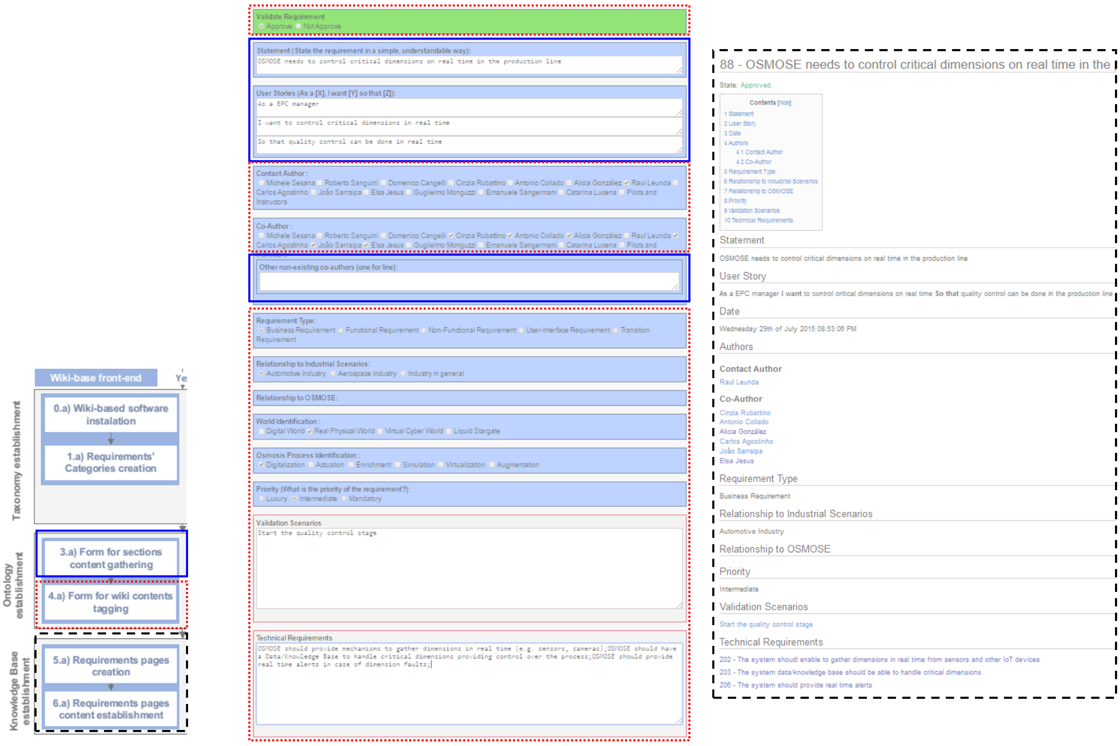


Figure 8.4: Creation and edition of User Requirements form.

were developed following the left part of the methodology of Fig. 8.3 to originate a form able to facilitate requirements creation and edition (see Fig. 8.4). The resulting form (requirements article template) is composed by checkboxes and text areas able to suggest the most suitable values in order to facilitate the tagging between pages and sections (step 4.a)). To handle description sentences, simple text areas are used (step 3.a)).

The representation of the developed form can be observed in the left of Fig. 8.4, where is possible to verify that the usage of the form resulted in a well-structured article page to represent the requirement, which enables the synchronization with the ontology (steps 5.a) and 6.a)). The adapted wiki front-end, besides the elicitation of requirements through brainwriting, allows its analysis (Fig. 8.1).

First, requirements can be <u>characterized</u> by its: 1) Osmosis world: Real World, Digital World, Virtual World; 2) Type: Business; Functional; Non-Functional; User-Interface; and Transition Requirements; 3) Priority; and 4) Industrial Scenario. The priority is related with the relevance of the requirement in the scope of the project. In the OSMOSE project, two industrial scenarios are considered: aviation and automotive industry.

Since requirements are handled in the wiki-front end, community users can <u>refine</u> the requirements any time, and follow its life cycle while they become even more clarified, focused, consistent, unambiguous and complete. At some point, the requirements are refined enough to be <u>approved</u> by the end-users. The approval of the requirements is made using the checkbox on the top of Fig. 8.4. Then, the technical team can start the derivation of technical requirements. Those are presented in the wiki front-end in a different category (with a specific creation and edition template). The tagging between user requirements and technical requirements is supported by an extension able to suggest technical requirements while typing. Finally, the technical requirements can be approved by the technical team in the same way the user requirements were approved by the industrial team.

The developed wiki front-end can be consulted through the OSMOSE project page[3] or using the direct link[4].

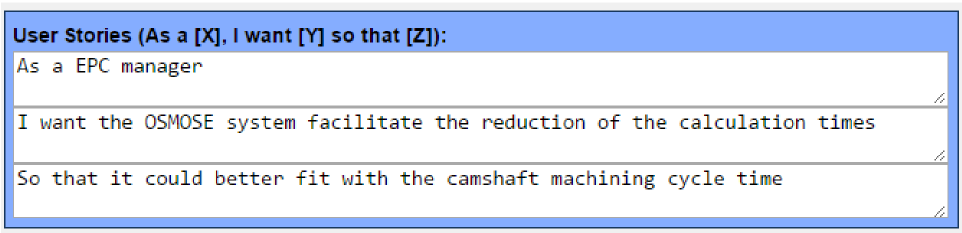


Figure 8.5: User requirements form developed for user stories (behavior) gathering.

8.4.1.1 BDD implementation in requirements management

BDD features were incorporated in the proposed tool to allow users to insert requirements: 1) user stories; 2) acceptance criterias; 3) and functional tests. in a structured way and following the "rules" described in section 5.2.2.1). As an example, for users stories gathering from industrial partners, the form of Fig. 8.5 was developed. It made sure that all user requirements were inserted using an unambiguous language enabling the technical team to understand their purpose and define technical requirements from them.

8.4.1.2 Requirements folksonomy

The proposed solution goes towards the folksonomy concept where the OSMOSE community collaboratively contributes for the requirements categorization process. During that process, the developed tool allows the establishment of tags between the wiki contents, namely between requirements and architectural components (developed solutions). The synchronization module is, then, able to translate that tagging into ontological relation allowing further management (e.g. traceability and reasoning).

The wiki main page enables to navigate in the created folksonomy (see Fig. 8.6), starting by consulting the several user requirements. Then, from each user requirement,

[3] http://www.osmose-project.eu/
[4] http://gris-dev.uninova.pt/osmose

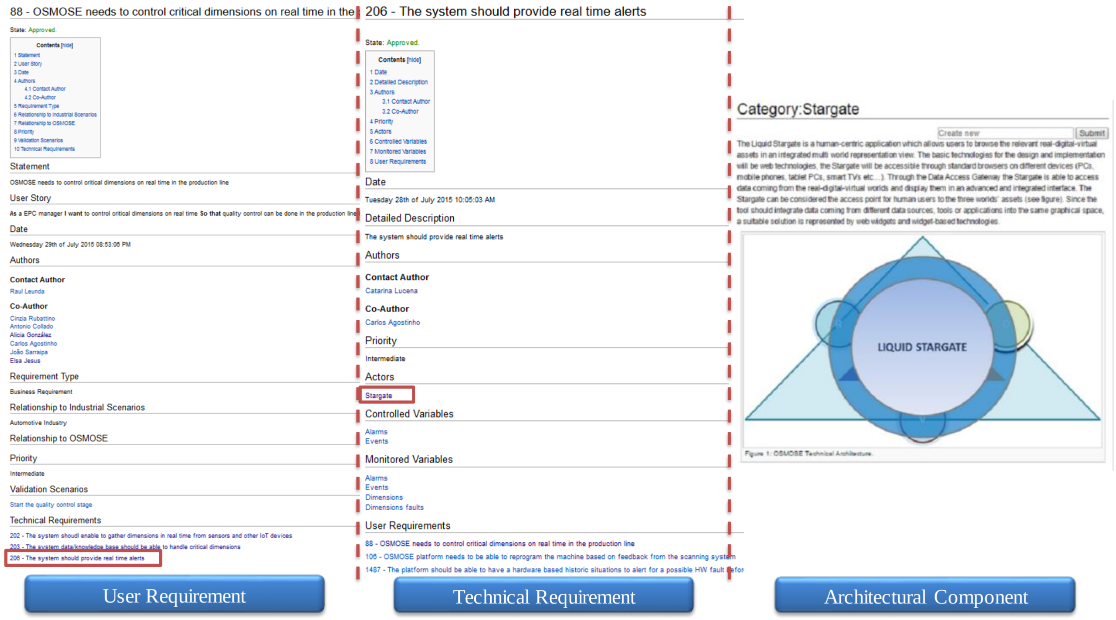


Figure 8.6: Requirements pages tagging (folksonomy).

several links (tags) can be used to access other information like technical requirements, authors, and architectural components.

8.4.2 Ontology establishment for requirements specification

The OSMOSE requirements knowledge base is a component which purpose is to capture OSMOSE requirements and its relation with other project elements (e.g. authors, architectural components). As explained along the paper, it also serves as a facilitator for requirements management, allowing different views of the information gathered from the wiki. Having this kind of knowledge specified would facilitate the search of specific information.

The Fig 8.7 represents the application of the wiki contents formalization methodology (see Fig. 8.3) to establish the knowledge base's taxonomy. As can be observed, the step 0.b) of the methodology consists in the creation of the class *Knowledge Representation* where all the instances are recorded (step 1):

- **Authors** - These are characterized in two groups: end-users and technical team;
- **User Requirements** - Characterization of what the system should do;
- **Technical requirements** - Characterization of How the system should be implemented;

Figure 8.7: Knowledge Base's taxonomy establishment.

- **BDD Features** - Composed by acceptance criteria and functional tests. Their characterization is made accordingly with section 5.2.2.1);

- **Variables** - That will be controlled or monitored. These are elicited in Technical Requirements;

- **Actors** - Participants/actors of technical requirements (e.g. simulator and simulator programmer);

The OSMOSE knowledge base structure is organized such a way that enables to represent conceptually the requirements classification features and the instances of the Wiki requirements, relating them both while keeping them physically separated. To achieve this result a classifiers taxonomy was build following the step 2 of the methodology for wiki-based front-end contents formalization (Fig. 8.3). In this step, the necessary information for requirements management is used to create the classifiers taxonomy. Based on it, the three main classes of the classifiers taxonomy are:

- **OSMOSE Technical Architecture** - enables to define direct links between technical requirements and architectural components;

- **OSMOSE Features** - used to classify requirements using concepts from the OSMOSE project (e.g. Industrial Domain, world; Osmosis Process);

- **Requirements Features** - handles generic requirements features (e.g. priority, Status, Requirement Type).

In the steps 3.b and 4.b of the methodology the data properties to handle free valued properties, and object properties to enable the representation of relations with the classifiers taxonomy are created (left part of Fig. 8.4). Two of the created data properties that are worth to highlight are the *Wiki Page*, which consists in an unique value that identifies an article in the wiki; and the *Requirement Version* that can range from the version 0 (requirement creation) to the version N (last version of the requirement). Thanks to the adoption of the Wiki page id, a unique value is associated to each requirement. Thus, future developments are clearly indexed to the original requirement contributing to its traceability.

Like is represented in Fig. 8.8, a direct correspondence between the front-end and the ontology is achieved, and consequently, all the requirements content is successfully migrated.

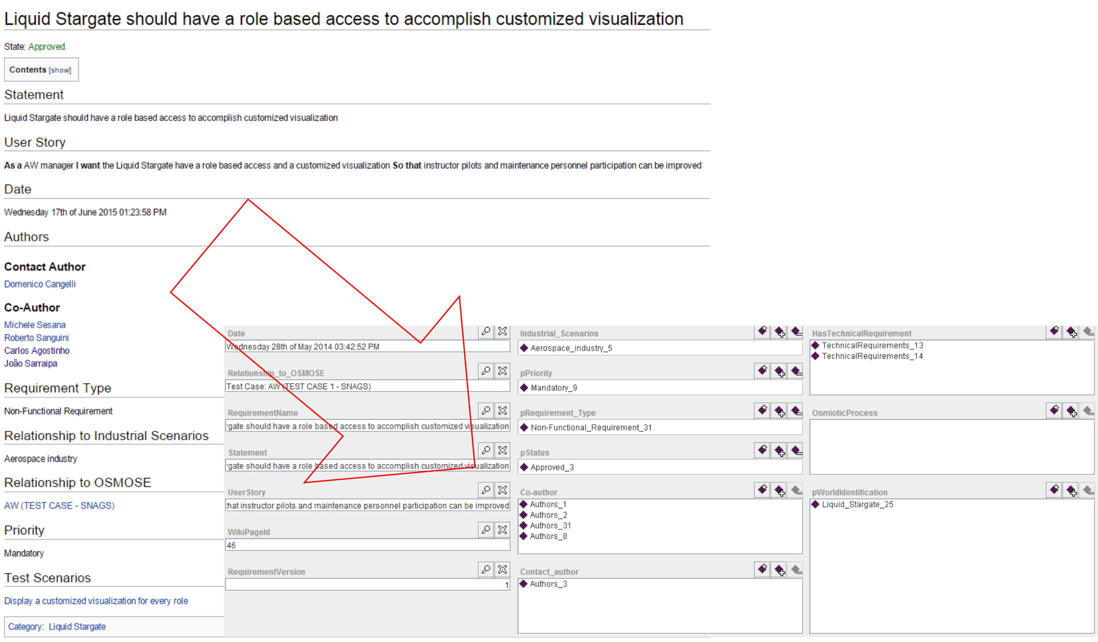


Figure 8.8: Direct correspondence between front-end and back-end establishment.

8.4.3 Synchronization Module

The synchronization module runs periodically and starts by connecting to the wiki frontend database in order to verify if any changes occurred since its last run. JDBC (Java Database Connectivity) is used to querying the front-end database. By querying the wikimedia table 'recentchanges', the authors have access to the set of changed pages, and its type: edition, creation, or removal. If the change is an edition or a creation, through the link to the table text (links to new & old page text) it is possible to have access to the current content of the front-end page. After the collection of the recent changes the HTML of each article or category's page is processed in order to create/ populate the necessary instances, data properties and object properties in the knowledge base.

8.4.4 Reasoning and Decision-Making

The formalization of the requirements in a knowledge was developed accordingly with Fig. 8.2 and supports management features like reasoning over requirements and their historical evolution. These management activities can be sub-categorized, such as:

- Change Management and Traceability;
- Prioritize of implementations;
- Common requirements analysis.

8.4.4.1 Change Management and Traceability

Requirements Traceability is considered as one of the most important characteristics for requirements management. Some important information can be retrieved from requirements evolution, or in some cases may be necessary to roll back to a past model representation (requirement representation). Since the ontology handles the different versions of a specific requirement taking (directly) into account its properties, it is possible to trace not only between specific dates, but also verify where a specific change occurs (e.g. what where the changes that a requirement suffers to become "Approved"?). If one desires to perform any query on past versions of the requirements without having to do a full "rollback" to the Wiki to find a specific change the OSMOSE requirements it becomes extremely important.

8.4.4.2 Prioritize implementations

The defined ontology handles the knowledge about categorized user requirements (e.g. priority and status), their derived technical requriements and consequent architectural components. Therefore, some reasoning to prioritize implementations can be made:

- Which are the approved user requirements that are considered a priority?;
  Selection of all user requirements whose priority has the value "mandatory" and status has the value "approved"

- Which are the most relevant architectural components to be added to the platform concerning the user requirement priority?;
  For these requirements follow the path: User Requirements → Technical Requirements → Architectural components.

- Which are the user requirements that will take advantage of adding one or more architectural components to the architecture?.
  It is done following the path: Architectural component → Technical Requirements → User Requirements

The capability of the ontology to answer these questions allows some coordination efforts of the partners, maximizing efficiency.

8.4.4.3 Common requirements analysis

The major part of the process for common requirements identification is automatic. However, it starts with the manual step of selecting, among the ontological properties available, the characteristics considered as relevant for common requirements election ($\{C_{C_n}\}$). This process uses the full ramification: user requirement → technical requirements →architectural components. Hence, several properties where elected both from the user and technical requirements:

- **User Requirements**
  - Requirement Type;
  - Technical Requirements;
  - OSMOSE World;
  - Osmosis process;
- **Technical Requirements**
  - Actors;
  - Controlled variables;
  - Monitored Variables;
  - Architectural Components.

As part of the manual initial step, it is also defined the minimum shared characteristics (threshold - $min_{C_{XN}}$) from which the requirements are considered common or not.

```
*** Group 1 ***:
159 - OSMOSE shall be able to modify the planning in concordance with maintenance
activities
53 - OSMOSE needs to enable actuation/reaction changes in RW based on the
maintenance and usage data


*** Group 2 ***:
84 - OSMOSE platform needs to send information to the customer about the camshaft
104 - OSMOSE needs to have different user interfaces
93 - OMOSE has to be able to share the 3D model of the camshaft among different
EPC departments
103 - OSMOSE needs to have different screens of results
81 - OSMOSE needs to display different levels of information depending who is
accessing to the data
92 - OSMOSE has to provide a customized (3D) viewer


*** Group 3 ***:
100 - OSMOSE has to facilitate the reduction of manufacturing errors
98 - OSMOSE has to contribute to reduce the maintenance and/or the replacement
costs
```

Figure 8.9: Groups of common requirements identified.

At this stage the automatic process can be executed. Therefore, for each requirement, the values of the elected properties are extracted and compared with the values of the full list of available requirements (Y, Z, ...). If a property match, a score is associated to the requirement compared (1 for each characteristic C_{xn}). This is repeated for the set of characteristics selected at beginning, increasing the score each time there is a match ($\sum C_{xn}$). Afterwards, the group of requirements considered common with the one under analysis (X) is the set that share a score above the threshold ($minC_{XN} > C_{xn}$). The process is repeated for all requirements and all the distinct groups are identified. In Figure 8.9 is possible to find the 3 groups of common requirements identified by the implemented algorithm. The identification of the groups can change accordingly with the threshold value specification. In the implemented case, the threshold value considered was 26 (very restrictive value).

Recommendation of resources based on common requirements

Thanks to the OSMOSE folksonomy established, User Requirements, Technical Requirements, and Architectural Components are connected between them. These relations, together with requirements characterization, can be used for the recommendation of solutions (architectural components) when a new requirement enters in the requirements knowledge base.

For the example provided in Fig. 8.10, the requirements '*The platform should be able to have a hardware based historic situations to alert for a possible HW fault before it happens*'. Based on the characterization of the input requirement, the requirements 106, 66, 65 were identified like having a higher level of similarity. Since the requirements identified as similar needed the architectural components: Stargate, Event Manager, and

```
target : 1487 - The platform should be able to have a hardware based
historic situations to alert for a possible HW fault before it happens


Most Common Requirements:

106 - OSMOSE platform needs to be able to reprogram the machine based on
feedback from the scanning system
66 - OSMOSE maintenance and scheduling application shall be linked to a
relational database management system
65 - OSMOSE shall store information about every simulator HW
configuration, software components version.


Recommendation of Architectural Components based on Common Requirements:

Stargate
Event_Manager
Osmose_data
```

Figure 8.10: Example of architectural components recommendation.

Osmose data for its implementation, these implementations will be suggested to implement the new requirement in the OSMOSE Platform.

8.4.5 Execution and Validation of Results

This section reports the compilation and consolidation of the technical requirements validation results. It is done by comparing the actual outcomes of a program execution with its expected behaviour. The complexity of the comparison depends on the complexity of the data to be observed. At the end of the analysis step, a test verdict is assigned. There are three major kinds of test verdicts:

- Pass - If the program produces the expected outcome and the purpose of the test case is satisfied;
- Partially passed - If the program produces only part of the expected outcome;
- Fail - If the program does not produce the expected outcome.

The Technical requirements were implemented using the OSMOSE Technical Architecture and their validation was done through behaviours comparison with the functional tests defined by the technical team. As represented in Table 8.3 using a technical requirement associated to the context manager module of the architecture, each technical requirement can yield one or more functional test (validation scenario). Each one of them describe one expected behaviour of the system (technical requirement). A technical requirement is validated in it is compliant with all of those behaviours.

During the technical evaluation 41 requirements have been examined to be included in a prototype: 32% were not addressed since they did not fit the required functionalities while the remaining 68% were implemented and consequently validated. Following the

Table 8.3: Technical Requirements Tests and Validation of Results (illustration).

Technical Requirement	**Scenario**	**Given** (preconditions)	**When** (action that happens)	**Then** (Outcomes/Output)
The system should have an events knowledge base	Event persistence in the Context Manager	**Given** an event happening in a world	**When** the event is of relevance for the event history	**Then** a new instance for the event is created in the Context Manager
	Event examination in the Context Manager	**Given** Relevant events in the past	**When** Stargate users open the event history	**Then** The context Manager returns all relevant events.

process just explained for technical requirements validation it has been inferred that 96% of the evaluated requirements were passed while 4% (one requirement) only partially passed. No failures where reported. Fig. 8.11 depicts both the requirements coverage and results of the evaluation of the addressed requirements.

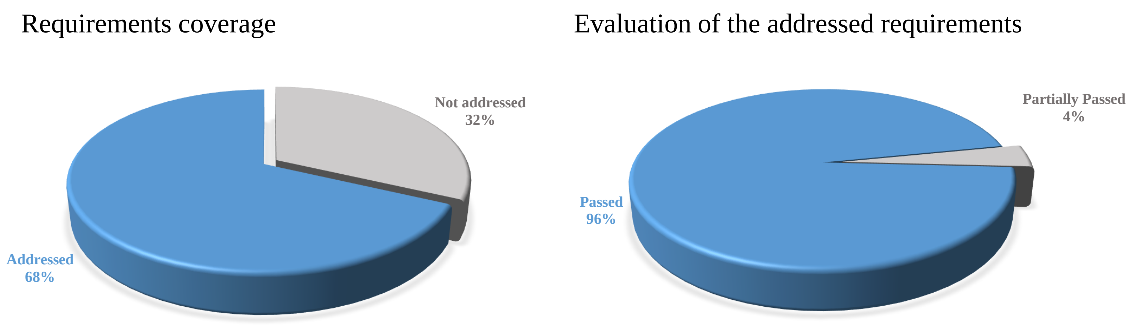


Figure 8.11: Execution and validation results.

8.5 Discussion

On the proposed Requirements Engineering methodology, behaviour driven features where used to facilitate communication between end-users and technical teams. It facilitated the understating of user requirements from technical teams since an unambiguous language was used to describe them.

The approach was concretized in a wiki based tool for requirements management. Using it, users were able to collaboratively participate in the requirements management process through an intuitive front-end. Consequently, the lack of domain experts input in requirements elicitation was reduced, which is considered one of the main causes of project failure. The developed tool consists in a wiki which edition section was transformed into a form to facilitate inputs insertion. The developed forms allowed to keep requirements specification, namely their structure and style during its life cycle, addressing the modifiability of requirements.

The collaborative tagging supported by extensions established links between user requirements and its authors, technical requirements and architectural components. Using this folksonomy, further reasoning could be made to optimize resources (some solutions might fit to similar requirements), prioritize implementations and make a trace of requirements evolution.

The proposed BDD based requirements management methodology was successfully implemented in the OSMOSE project when implementing the new sensing liquid enterprise concept. Using the proposed methodology and consequent framework, 92 User Requirements where elicited. Table 12.1 provides a summary of the number of requirements elicited accordingly with their type, OSMOSE World and Osmosis Process. From these requirements 49 where considered as mandatory for the project development.

Table 8.4: OSMOSE Outcomes in relation to User Requirements Elicitation.

Requirement Type	**OSMOSE World**	**Osmosis Process**
Business: 21 Functional: 56 Non-Functional: 10 User-Interface: 5	Real: 17 Digital: 24 Virtual: 11 Liquid Stargate: 15	Digitalization: 6 Actuation: 5 Virtualization: 4 Augmentation: 4 Enrichment: 3 Simulation: 1

After User Requirements Elicitation and Refinement, 72 Technical Requirements where derived and associated to OSMOSE Architectural Components [356]. The implemented methodology also supported the project coordination and validation through acceptance criteria and functional tests definition. Section 8.4.5 describes in detail the coverage of tests execution in relation to the defined requirements.

CHAPTER 9

PROCESSES MODELING APPROACH

The OSMOSE Project[1] aims at developing a reference architecture for modelling and managing sensing liquid enterprises [357], this is achieved by interconnecting three worlds: the Real World (RW), the Digital World (DW), and the Virtual World (VW) (see more in 8).

The osmosis concept is a process of passing the molecules from a less concentrated solution (individual perception of each world) to a more concentrated one [358]. The project follows this concept, where each enterprise has a special type of business processes used to moderate the information exchanged between the different worlds, named Osmosis Processes. Although having a different meaning inside the project they can me modelled using the same strategies of regular ones. The six Osmosis processes considered are [339]:

- **Digitalization** (RW-DW) - Model and representation of real world data in a computer-tractable form;
- **Actuation** (DW-RW) - Plan and implement highly distributed decision-making;
- **Enrichment** (VW-DW) - Extends the computational and experiential capabilities of the Digital World annotations and projections coming from simulations and what-if hypothetical scenarios;
- **Simulation** (DW-VW) - Instantiate and run hypothetical future scenarios fed by Digital World data;
- **Virtualization** (RW-VW) - Provides data for simulation of hypothetical simulations from the real world and runs the simulation;

[1] www.osmose-project.eu

- **Augmentation** (VW-RW) - Annotates Real World objects with Virtual World information.

To facilitate the collaborative processed modelling from the project (and domain) participants, The OSMOSE Process Manager Modelling Toolbox was developed. It supports users in the definition of the osmosis processes starting from a high-level conceptual view, provided directly by business actors, until the definition of the osmosis process model, which needs to be conducted by more technical actors such as system architects. These different processes are used to interconnect the different worlds by extending the enterprise processes running in each world, to link them with each other, by crossing the osmosis membrane between them. From the OSMOSE perspective the 3 worlds are represented also as a set of enterprise processes and sub-processes, with their internally generated tasks and events. The toolbox supports the composition of a process by means of specifying activities and events, which trigger depends on specific performer indicators (e.g. raise of the temperature, emergency button pressed, etc.).

9.1 General Approach Model to Instantiate the OSMOSIS Processes

The Osmosis Liquid Sensing Enterprise (LSE) processes are a special type of enterprise process, embedded in the organization's business process. They extend the business processes running in the different worlds linking them with each other by crossing the (osmosis) membrane between the different worlds. They can be associated to certain services, which in turn can trigger events as well as act upon organizational resources.

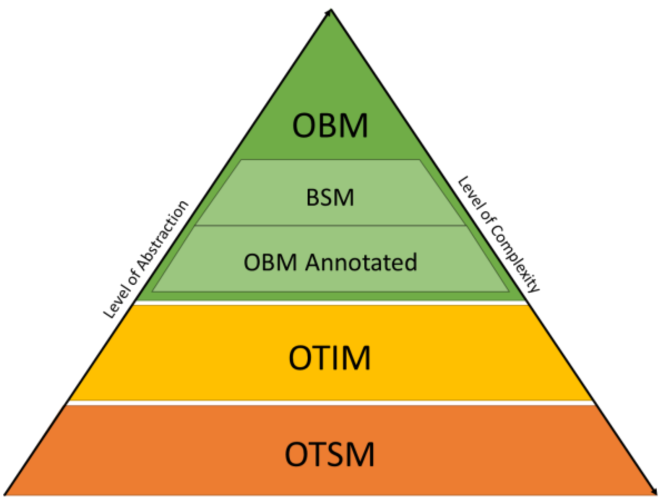


Figure 9.1: OSMOSE Process Manager Modelling Architecture.

Following the work of Ducq et al. (in [286], [359]) on model driven service engineering, and as explored by Agostinho et al. (in [298]) one can consider that the design of services is inherent to the manufacturing Liquid-Sensing Enterprise. Hence, the LSE design could benefit from the methodology behind MDSEA in order to accelerate the transition of the traditional enterprise to the 'internet-friendly' and context-aware organization envisaged in OSMOSE. Thus, even though the initial objectives where different (processes modelling instead of services) the MDSEA can be adapted to support the osmosis process development, envisaging modelling at three abstraction levels that are integrated among themselves following semi-automatic top-down transformation mechanisms (see Figure 9.1):

- **OSMOSE Business Model (OBM)** - specifying the business service model, and identifying the innovation requirements and behaviour for the novel service (using business language). This level is extended with specific annotation possibilities (OBM Annotated) that provide the possibility to separate the concerns among the IoT-worlds. This enables to identify osmosis processes. For these reasons OBM was divided in two parts, the BSM (Business Service Modelling) to design the Business model, and the OBM Annotated;

- **OSMOSE Technology Independent Model (OTIM)** - complementing the upper level model with detailed technology independent functionally. This modelling level is indented for product engineers or software architects to model the osmosis processes, their full behaviour and constraints. In transformation from OBM are oriented and prepared for the osmosis process instantiation;

- **OSMOSE Technical Model (OTSM)** - which is the last level and consists in the instantiation of the identified behaviours and constrains with architectural components and in the process execution.

This three-layer paradigm facilitates the innovation potential involving multi-disciplinary teams, and bringing together skills that go from business and marketing, to design and implementation, progressively deriving and reusing knowledge down-to the implementation level.

A guidance on how the OSMOSE paradigm is adopted in the processes modelling and execution task is described step by step following the waterfall approach illustrated in Figure 9.2. The first two steps are part of the OBM layer. The first one consists in the definition of the application goal, and the second one is focused on the identification of basic (business) participants involved (e.g. organizations and resources). The categories of participants that can be identified at this level are represented in the top of Figure 9.2.

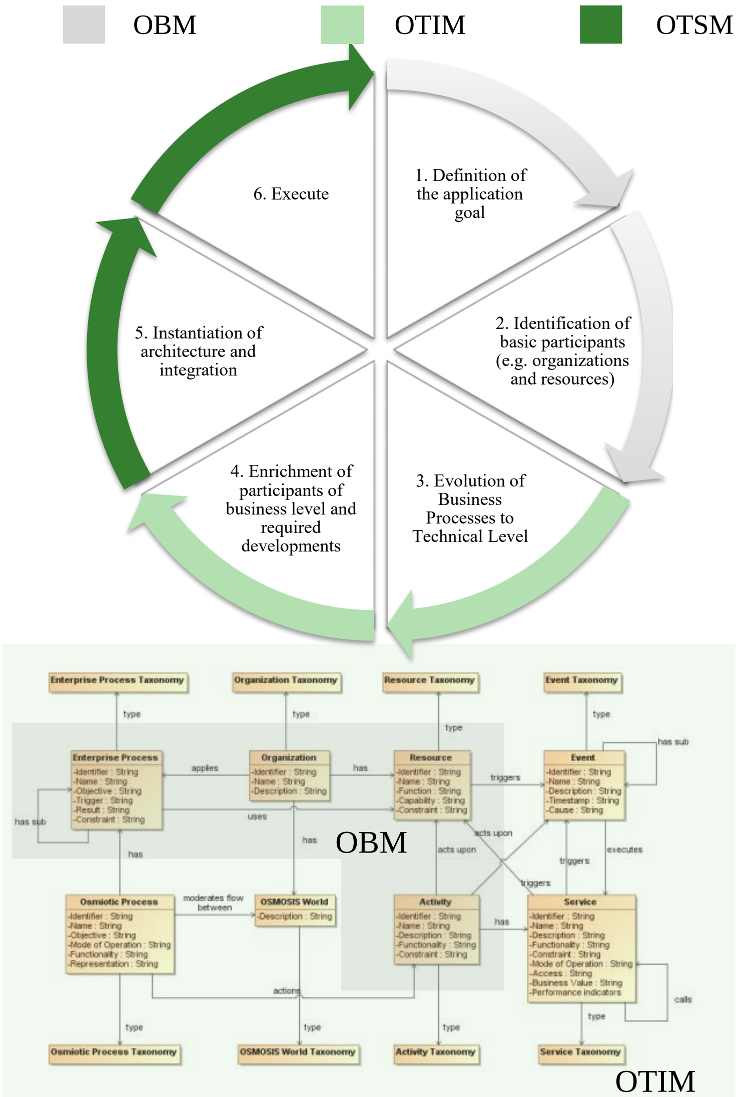


Figure 9.2: Waterfall Approach to instantiate OSMOSE processes modelling and execution.

The next two steps are related to the OTIM layer. The step 3 (Evolution of Business Process to Technical level) starts with the identification of the involved worlds and separation of concerns to identify the osmosis processes. Then, the definition of the right events is essential (starting point of step 4). Some of them can be already in the business model and should be recognized. Others have to be created from scratch exposing new items or features to the system. The participants identified in the OBM layer can also be enriched with technical information (e.g. relations between services and activities, since services where only identified at this point). In the OTSM steps, the required elements of the architectural components are instantiated to accomplish the OTIM level specifications and the processes are executed.

9.2 Liquid Sensing Enterprise Modelling Environment

As the general approach instantiates a model driven framework, extending MDSEA. It is a logical choice that the modelling framework is a continuation of MSEE's for service modelling. Hence, the LSE toolbox for process modelling and development is taking up the MSEE service lifecycle toolbox [360]. Several modelling languages have been chosen at each level to enable the osmosis processes modelling. At the OBM level, following the MSEE's MDSEA, the main language used is Extended Actigram [361]. The advantage to use this language at this level is its simplicity and easiness in the use for the companies' business members. At the Technology Independent Model (TIM) level, the modelling languages must allow a more detailed representation of each concept and process. Hence, also following MDSEA, BPMN 2.0 is the appropriate choice. At the TSM level, the jBPM environment for processes execution, was elected, accompanied by the Esper complex event processor for the events management, and RabbitMQ for messages profiles (E-mail and SMS) creation and management.

9.2.1 OBM-OTIM Process Transformation

In order to reach the OTIM layer, an ATL[2] engine is used to automatically execute the predefined transformation rules between the OBM and OTIM models. ATL is one of the most used transformation languages, having a large user base and being well documented [362]. It is composed by a set of rules that define how the source model elements are linked and how the target model elements are instantiated. Subsequently, BPMN 2.0 is used (via the BPMN2.0 Visual Editor for Eclipse) to model technical details of osmosis processes. Thus, every time a user finalizes the design of their Business model, the tools offers to the user the possibility of transforming the business model into the technical model and enrich it with more technical information.

To enable this feature, it was necessary to define the mappings between both models used in the transformation. These mappings include a mapping from EA* language to the BPMN language and subsequent extension to handle the modelling of osmosis behaviour. An initial transformation from EA* to BPMN can be found on the work of Bazoun et all [361]. Taking this work as a basis, some changes were made to handle the modelling of osmosis behaviours that are summarized in Table 9.1.

As can be observed in Table 9.1, the EA* process transformation to BPMN 2.0 triggers a rule that automatically creates all the necessary elements to describe the worlds interactions (pools, processes and lanes). The resources transformation was also changed. In the current result, both the EA* resource participants and responsibles are transformed

[2]https://eclipse.org/atl/

to elements belonging to the osmosis process staring world. The way that the current transformation affects the LSE modelling environment is explained in more detail in the following subsections.

9.2.1.1 Osmosis Process Modelling as Collaboration Diagrams

One change in the transformation is related to the fact that the transformation of an OBM EA* will always result in a OTIM BPMN 2.0 Collaborative Diagram. The advantage against using a regular BPMN, (where the elements are represented within a pool or participant [363]) is that a collaborative process allows to not only represent the activities of each osmosis word or process, but also the interactions between them. Thus, it can be considered, in a global point of view, that not only a perspective of a particular world is represented, but also the osmosis events that might cause the transition between worlds and also the messages exchanged between participants.

Table 9.1: Summary of changes made in EA* to BPMN 2.0 transformation.

EA* (OBM)			BPMN 2.0 (OTIM)	
Process			Pool	Starting World Pool
				Ending World Pool
				Osmose Platform Pool
			Process	Middleware Process
				Starting World Process
				Ending World Process
			Lane	Membrane Lane
				Osmosis Process Lane (e.g. Digitalization)
				Ending World Lane
Resources	Resource	Responsible for	Lane	Lane in Staring world
		Participant in	Resource	Resource (added to the list of resources of a task)
	Human	Responsible for	Lane	Lane in Staring World
		Participant in	Resource	Resource (added to the list of resources of a task)
	IT	Responsible for	Lane	Lane in Staring World
		Participant in	Resource	Resource (added to the list of resources of a task)

As represented in Table 9.1, when the transformation between EA* and BPMN is executed, independently of the osmosis process selected, the output BPMN diagram will have 3 pools (or participants): 1) Osmosis Platform - composed by the Membrane (with is logic included) and the Osmosis Process; 2) The starting world; and 3) the destination world. This is also represented in Figure 9.3.

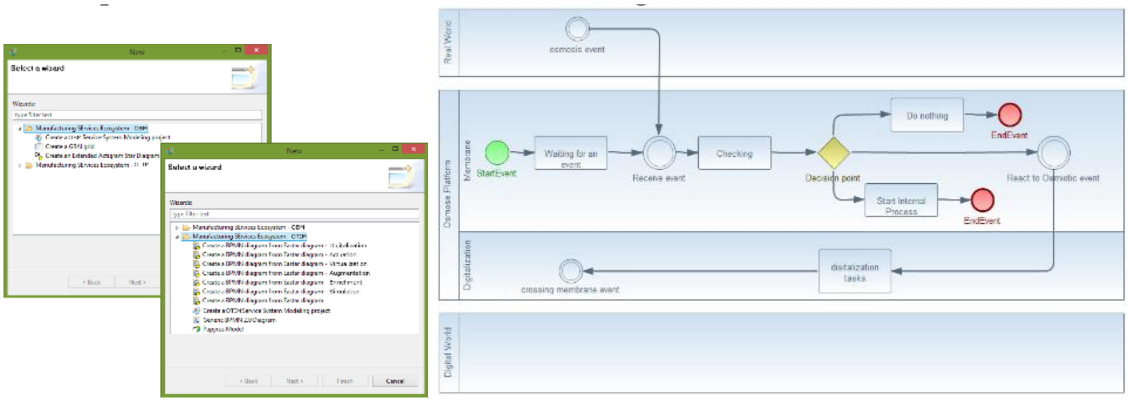


Figure 9.3: Osmosis process selection in tool interface.

9.2.1.2 Osmosis Processes and Worlds

Another change in relation to MSEE's developments is related to the tool interface. As illustrated on the left of Figure 9.3, at the moment a user decides to evolve the business model to its technical representation, he/she needs to know a prior which is the Osmosis Process to represent. Thus, accordingly with the osmosis process selection, the conditions provided in Table 9.2 are triggered (e.g. the selection of a digitalization process will result in the creation of the RW and DW pools, as represented on Figure's 9.3 right).

Table 9.2: Osmosis Processes changes in the EA* to BPMN 2.0 transformation.

EA* (OBM)	Condition	BPMN 2.0 (OTIM)
Process	Digitalization	Creates Real and Digital World Pools
	Actuation	Creates Digital and Real World Pools
	Virtualization	Creates Real and Virtual World Pools
	Augmentation	Creates Virtual and Real World Pools
	Enrichment	Creates Virtual and Digital World Pools
	Simulation	Creates Digital and Virtual World Pools

9.3 OSMOSE Process Manager Modelling Toolbox

The tool corresponds to the design time in the OSMOSE project and it is divided in three parts. The first part is the Business Design which aim to give the capacity for any user to design a process in a high level perspective, by defining actors, goals and activities. The second part is the BPMN process. After the user finish the Business Model, the tool is able to transform the model into the BPMN model. The last part is able the process to be executed since the process follows the BPMN standard. By providing these features to the project, the tool will improve the performance by managing and optimizing the processes.

The OSMOSE Process Manager offers the following features:

- **Modelling and Analysis** - It supports the strategy (To-Be) of the business model, enabling an analysis of the goals and clarifying how the business process will work, starting from a business perspective (without any notions of OSMOSE). After that, it is time to design the Technical Model, with the support of the Toolbox where is described the process in BPMN, identifying tasks, events, inter-dependencies between the different worlds etc.

- **Execution (External)** - The execution of the business processes is not included in the Toolbox. However, it is provided an export option, allowing the user to export the BPMN model into a number of executable processes from the perspectives of the Real, Digital and Virtual worlds. For example, in the case of a digitalization, it is created a Real process, a Digital process and an Osmosis Event process ready for the execution phase.

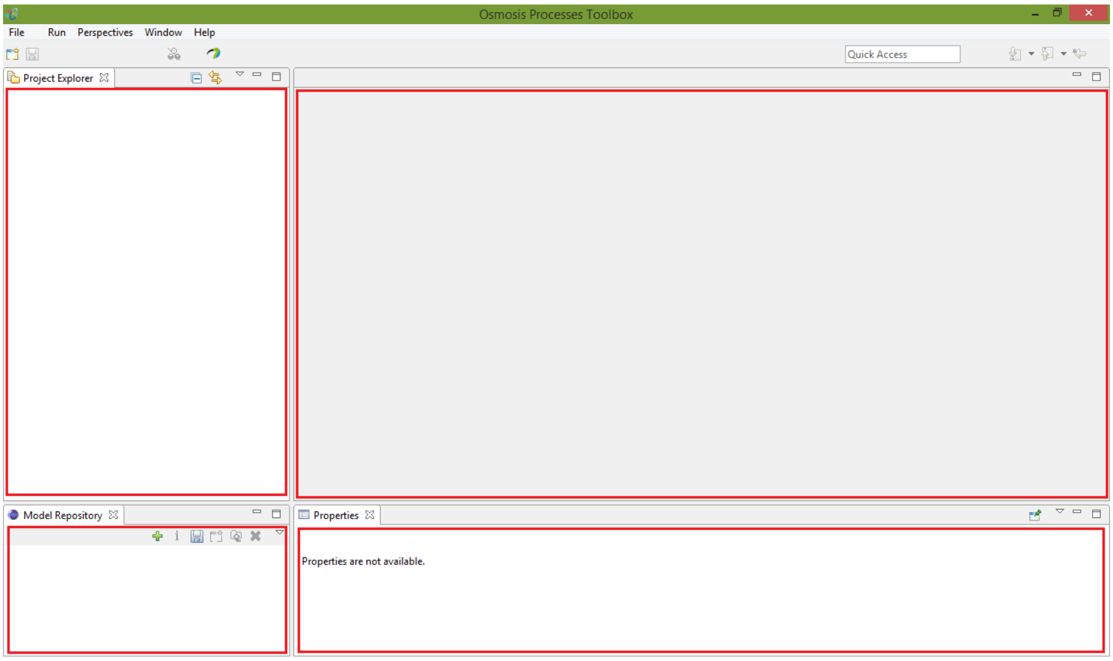


Figure 9.4: Osmosis Process Manager Workspace.

9.3.1 General Overview

The main graphical interface of the application is composed of four basic areas or views, which are depicted in Figure 9.4 and described below:

- **Project Explorer** - Located on the top left side, allows to browse the existing projects in the workspace;
- **Model Repository** - Located on the bottom left side, allows to connect and manage model repositories;
- **Properties** - Located on the bottom right side, allows to view and edit the properties of a given entity;
- **Workspace** - Located on the centre, it displays the editor view currently being used by the user.

To start modelling a process, a new project can be created by using the *File –> New* option from the main menu. A dialog box should then appear enabling the user to choose a wizard according to his needs, as depicted in Figure 9.2. At this point one may choose what type of modelling level it wishes to work on, whether it is at OBM, or OTIM level. Each level may contain more than one specific wizard that guides through a different set of options in order to generate a project according to the user's own preferences.

9.3.2 Modelling a Process

For the purposes of this instructional section, the following illustrates the development of a simple "Hello World" process, detailing all the fundamental information that is required in order to create a process from the higher level of abstraction, OBM, until the middle level of abstraction, OTIM. The last level of abstraction (OTIM) is only addressed after BPMNs export and consequently it is not directly addressed by the toolbox.

9.3.2.1 Modelling at OBM level

To start, the user should initialize the OBM project wizard in order to create an OBM project. Once all the basic details regarding the project names have been inserted and the project has been created, it is then possible to start detailing the process by opening the ".bsm" file and complementing all the relevant fields that concern the process specification, as illustrated in Figure 9.5.

In this specific case, the process is comprised of a service, a single enterprise that has one associated resource, and a process. Once the process has been fully specified, the user can then create the corresponding EA* diagram and start modelling it. This is done by right-clicking the project name in the Project Explorer and selecting *New -> Other ->Create an Extended Actigram Star Diagram*. By opening the newly created EA* diagram, the Workspace view will open up a graphical editor with all the features required to model it, as displayed in Figure 9.6.

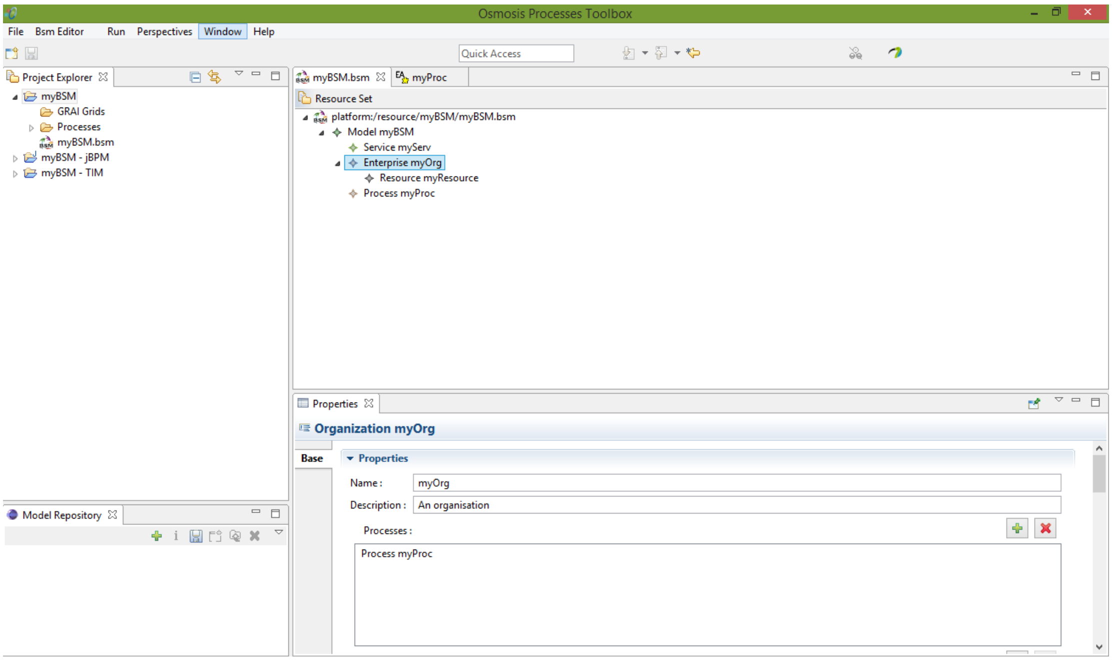


Figure 9.5: Creating an OBM project.

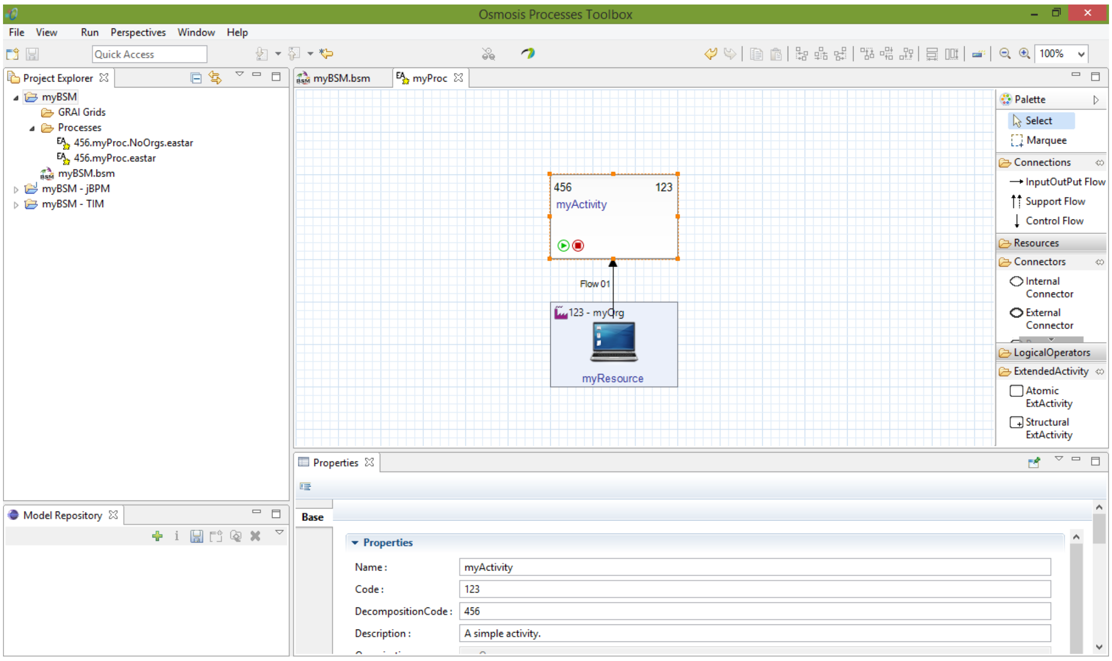


Figure 9.6: Modelling an OBM process with EA*.

In this case, a simple Atomic *ExtActivity* was created, name *myActivity*, in order to illustrate a given task that is to be performed during process execution. The IT resource *myResource*, which belongs to the organisation *myOrg*, is responsible for performing the aforementioned activity. In summary, the diagram describes a process in which a single activity will be executed by a given resource.

9.3.2.2 Modelling at OTIM level

Once the diagram at OBM level has been modelled, it is then possible to transform it to OTIM level in order to further detail it and complement it. This is achieved by creating a new OTIM project, by choosing the *Create an OTIM Service System Modelling project* wizard, and then, once the project is created, executing the *Create a BPMN diagram from Eastar diagram* wizard in the *New->Other* menu. This wizard allows the user to choose an EA* diagram already existent and transform it into a BPMN process. It should be noted that if the user knows at start which osmosis processes it desires to approach, then it is possible to choose the specific osmosis process type from the wizard dialog, allowing simplify the modelling process by automatically complementing the resulting BPMN with tasks related to that specific osmosis process. As mentioned before, once the new BPMN model is opened, a BPMN graphical editor can be used to further detail and complement it. The result in consequence of our scenario is depicted in Figure 9.7.

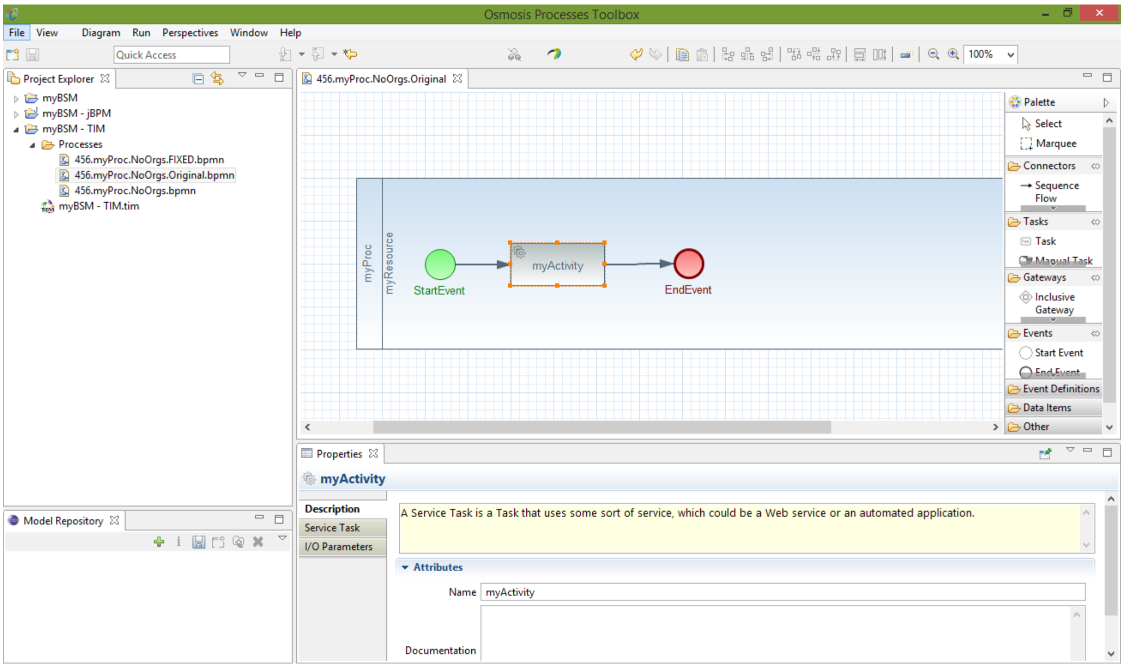


Figure 9.7: BPMN derived from an EA* process.

As it can be observed, the EA* diagram modelled previously was effectively converted into a BPMN process comprised of a single task that runs within the specified resource belonging to the specified organization. As illustrated in Figure 9.7, the activity created in the EA* diagram was transformed into a BPMN Service Task. However, because this tutorial wishes to create a "Hello World" like process, a Script Task will be required instead in order to run a script and display a message in the console. Thus, the process was improved according to this specific need, and the Service Task replaced with a Script

Task, as illustrated in Figure 9.7.

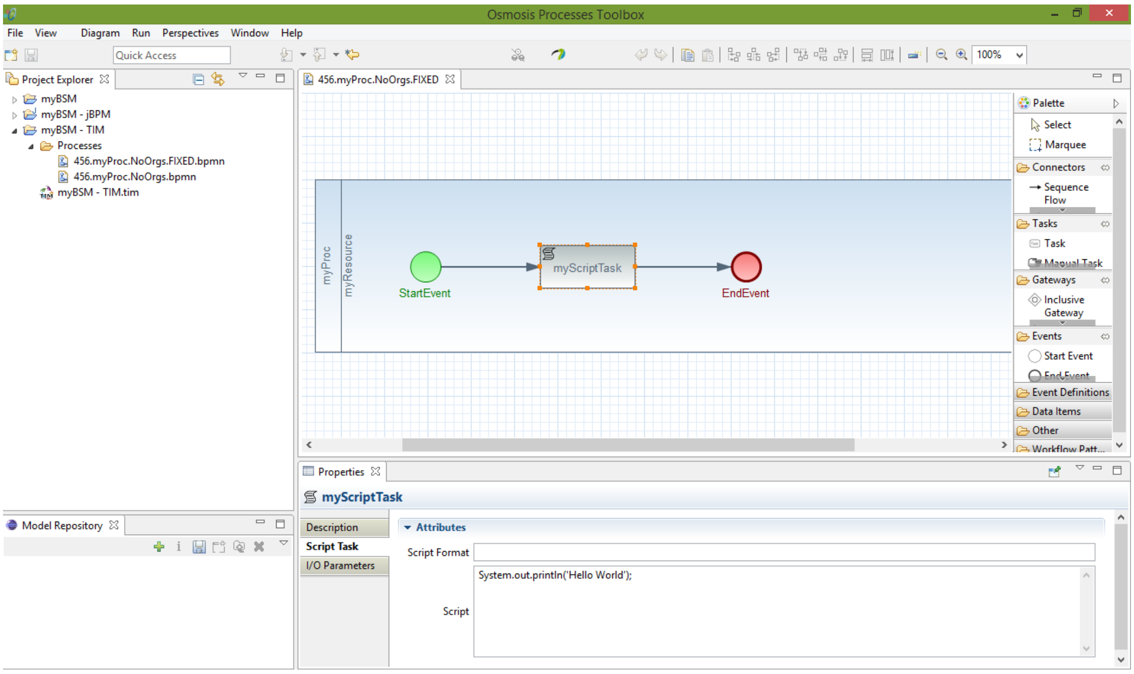


Figure 9.8: Complementing an existing BPMN process.

As illustrated, the taks will be responsible of running a simple Java function that should display the messate "Hello World", thus concluding the modelling process.

9.4 OSMOSE Use Case Example

The use case provided is related to the monitoring of a camshaft manufacturing tool which as a short life-time left. When it reaches a certain number of years or performance levels, a recalibration/replacement is needed and camshafts production should be interrupted. This monitoring is made by a RW smart toolbox which has the capability to evaluate if the manufacturing tool is near expiration.

9.4.1 OSMOSE Business Model

Following the first two steps of the waterfall approach presented section 9.1, the modelling starts by the identification of the application goal (Enterprise Process) and in the identification of the business participants:

- **Enterprise Process** - Monitoring of the manufacturing tool;
- **Organizations** - Manufacturer Organization;

- **Resources** - Monitoring System; Tool; Maintenance and Production Manager; Scheduler System; Technician;

- **Basic Activities** - Monitoring Lifetime of Manufacturing tool, Block Tool Usage, Notify Responsible, Schedule Maintenance Procedure for Tool, Select Appropriate Technician, Notify Technician, Execute Maintenance Procedure.

To describe the relations between the identified participants and the business strategy, an EA* model should be provided next (see Figure 9.9). This model starts with the monitoring of the tool activity, and then, depending on the identified issues, other activities should be carried on: block the tool usage and notify the responsible entity if it is a severe problem; or just notify the responsible entity if the issue does not justify the tool shutdown. After that, a maintenance procedure should be scheduled and the appropriate maintenance technician selected and notified, so the actual maintenance can be executed and the camshaft produced.

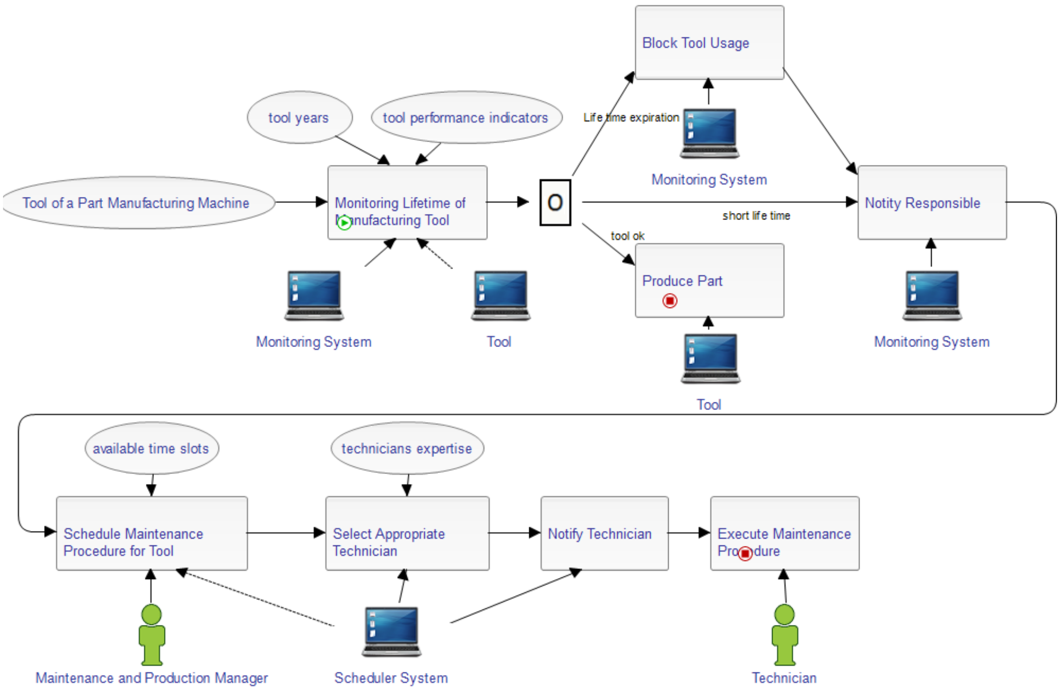


Figure 9.9: EA* (OBM) model of Monitoring of the manufacturing tool.

9.4.2 OSMOSE Technological Independent Model

The steps 3 and 4 of the waterfall approach are executed by the technical team, which has the knowledge about the OSMOSE concepts and the needed technical modelling skills. Thus, after the evolution of the business process to the technical level through the automatic transformation of EA* to BPMN 2.0 (step 3), the technical team is in charge of enrich the model with information about participants not identified at the business level (step 4). Although, for each OBM one or more OTIM models can be derived, for the sake of simplicity, in the example provided it is only considered the flow between the 'Monitoring of Lifetime of Manufacturing Tool' as a RW activity and the need to 'Notify Technician', a DW activity. After the identification of the flow as a digitalization osmosis process, new participants are identified:

- **Osmosis Process** - (Digitalization) Maintenance Schedule update;
- **Worlds** - Real and Digital Worlds;
- **Events** - Info available, Maintenance Required (osmosis event), Maintenance info available;
- **Activities** - Process tool information, Get tool location, Store data.

As illustrated in Figure 9.10, the output of the transformation is composed by the OBM participants (only the ones in the identified flow) plus the elements characteristic of the osmosis behaviour explained in section 9.1. Afterwards, the diagram is completed with further technical detail using the newly (OTIM) participants identified.

9.4.3 OSMOSE Technological Specific Model

After the conceptual modelling phase, it is the implementation model, which is the design of the BPMN to be launched to execution. Between the OTIM and OTSM and because both are using BPMN 2.0, a simple export/import of the BPMN file is done. Hence, it is not really a transformation, however, if needed some predefined rules could be implemented together with the export functionality.

Then, in the model is made the configuration of the BPMN with the respective calls of the services and events, which are going to be used in the run-time mode. Finally, the model is complemented with more detail and ready to be tested and then ready to be launch in execution mode, making the osmosis process runnable.

9.5 Discussion

In this chapter, the author presented the Osmosis Processes concept and its associated modelling challenges for the liquid-sensing enterprise. In order to address them, a three-layer paradigm based on the MSDEA approach is used to facilitate the coordination and cooperation potential between multi-disciplinary teams. This three-layer paradigm is instantiated

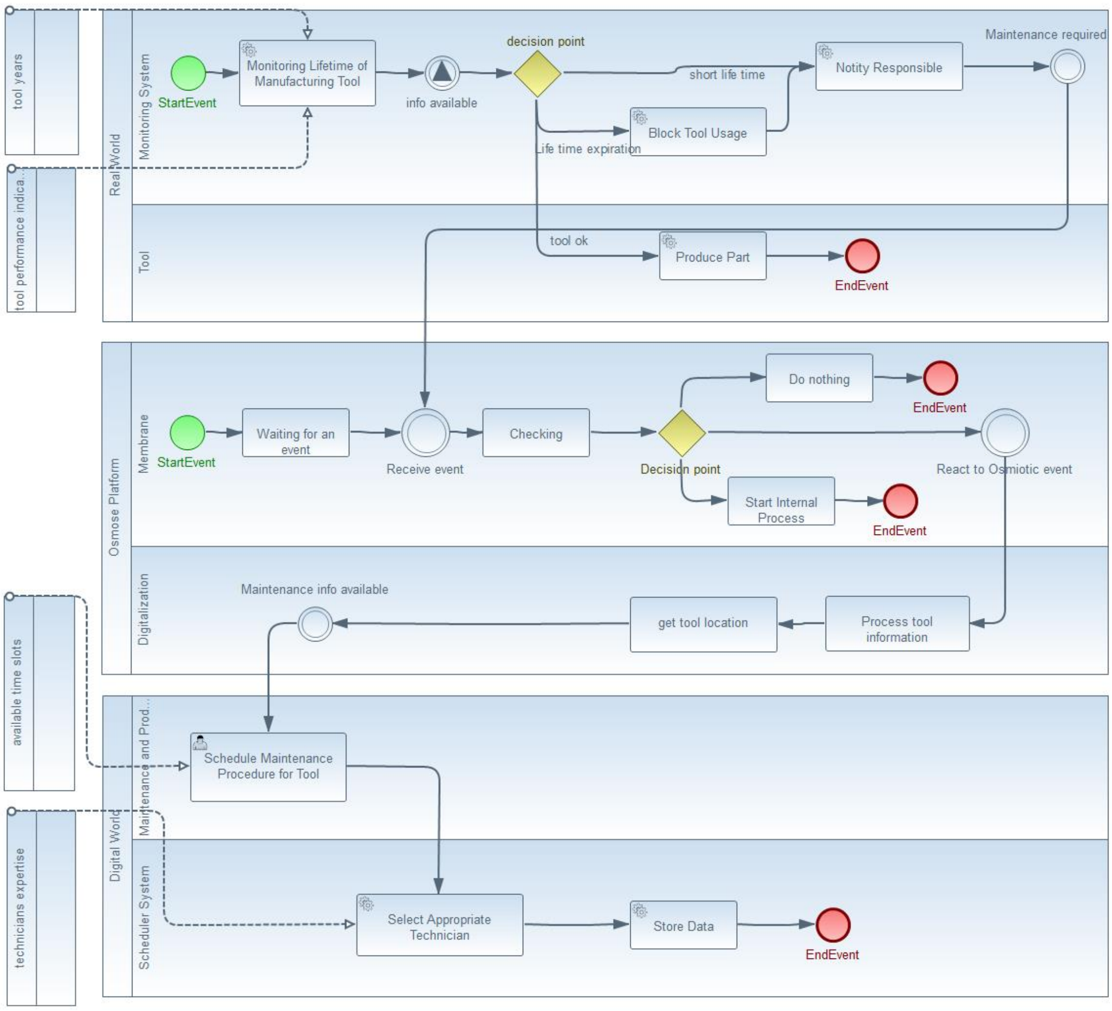


Figure 9.10: OTIM model of the Example.

by a 6 steps waterfall approach that allows the modeller to separate concerns regarding business priorities, technological independent models and final processes execution.

The 6 steps of the waterfall approach are supported by a modelling tool adapted from MSEE's project results to the purpose. The changes made to the tool follow the need to model the osmosis processes concept, namely the interactions between worlds, and the middleware membrane decision logic. As described in section 9.2, the evolution ranges from the EA* to BPMN 2.0 mappings and transformations, to the interface itself (see Figure 9.3). These changes enable enterprise's business members to use a simple and easy to modelling language like EA* to model the business strategies, and then, the technical team, with their knowledge about the osmosis worlds concept and technical modelling skills, to enrich the business model with the osmosis behaviours and constrains that need to be instantiated with technical components in the third layer paradigm level.

CHAPTER 10

SERVICES MODELING AND INFORMATION EXCHANGE

One of the biggest issue in the Service Oriented Architecture (SOA) vision is related to the data heterogeneity between inter-operating services. This is because, typically, enterprise systems are developed over several periods of time, by diverse organization and not necessarily with the same structures or vocabulary [364]. As a consequence, a substantial heterogeneity in Web Services elements' syntax, structure and semantics is verified.

Given the pivotal importance of service discovery for service-oriented computing, several attempts to improve the quality service discovery tasks are currently being pursued. One of the major efforts in this direction is promoted by the World Wide Web Consortium (W3C) which strongly advocates the introduction of semantic information in the description of Web services[1] [365]. Semantics acquires a particular importance to share services in a Peer to Peer (P2P) system where the lack of common understanding of the world generated the need of explicit guidance in the process of discovering available resources [366–368].

Currently, services description is expressed by means of the Web Services Description Language (WSDL)[2] by declaring a set of message formats and their direction (incoming/outgoing) [365]. WSDL, in essence, allows the specification of the syntax of the input and output messages of a basic service, as well as other details needed for the invocation of the service [369]. However, it can be considered that WSDL do not offer sufficient semantic richness for services to be machine-processable [370]. Human intervention is

[1] http://www.w3.org/2003/10/swsig-charter
[2] http://www.w3.org/TR/wsdl

often needed to interpret the meanings in order to discover, compose, and invoke Web Services. Here, the notion of ontology is seen as an effective way to provide the required specifications to allow the automation of the mentioned tasks [370]. Five projects to support the idea of ontologies incorporation in Web Services description are: OntoGov [371], TERREGOV [372], SemanticGov [373], Access-eGov [374] and FUSION [375].

10.1 Web Services Interoperability

There are several tools that allow mappings establishment across several data formats. One example is the *Advanced Mapping and Transformation* module of talend [376]. This module uses a Eclipse-based tooling environment to enable users to build, tests and maintain mappings across several data formats, including transformations between java and complex XML data, JSON, databases, flat files and EDI applications. It can be used to integrate applications, services, and APIs without coding. The *SAP NetWeaver Composition Environment Library*[377], allows users to define how the data coming from one service operation (as input) in the service flow has to be mapped (definition of what will be the output) in the service flow. Thus, allows to directly define how the data from one service data structure is mapped to another data structure. Other application tool to consider is IBM's *WebSphere Application Server V8.5.5* [378]. It supports services mappings establishment by providing a simple way of performing content-based routing and message transformation that can be applied to web service messages being sent from service clients without needing to make changes to either the service client or service provider. All this tools are valuable solutions contributing to the deployment of interoperable Web Services across platform, applications, and programming languages. However, the author didn't found any indication that these tools are capable to import OWL, used to define semantic relationships between concepts. This goes against the need of semantic descriptions usage to facilitate services discovery.

To face this issue, the author proposed a solution able to represent connections or mappings between information systems, or more precisely, between web services elements and semantic concepts described by ontologies. The semantic representation of the domain is accomplished by a domain reference lexicon establishment in collaboration with the domain experts. Thus, the (external) cooperation and awareness building regarding the domain is supported by the usage of a common reference lexicon (domain semantics) to be the intermediary in the communications between enterprises Web Services and to the outside. Consequently, the establishment of mappings between enterprises Web Services and the reference domain lexicon will allow each of the enterprises to keep its own knowledge and semantics unchanged, and still able to smoothly interact with its

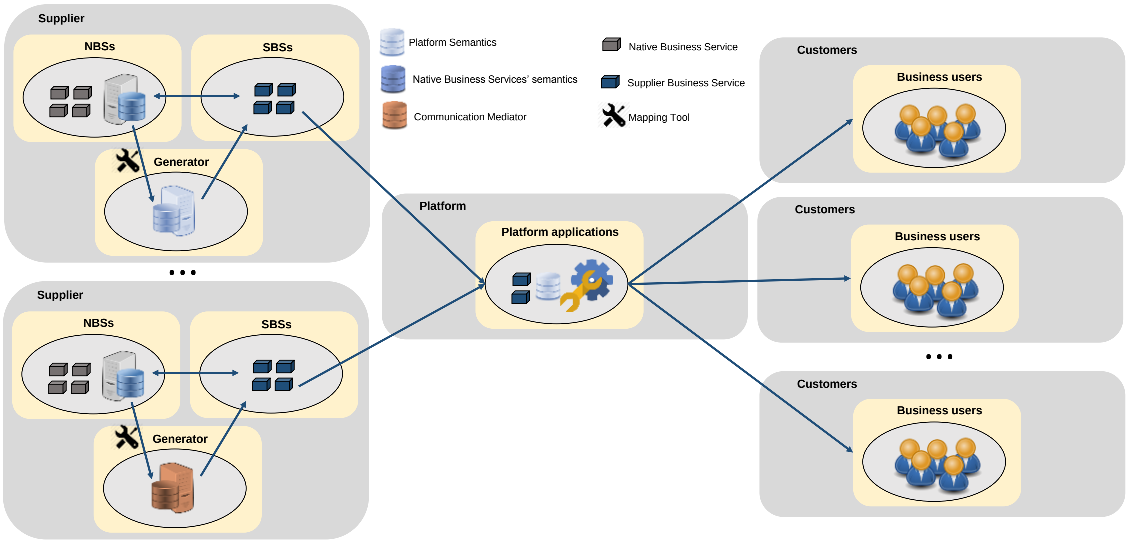


Figure 10.1: Approach for the interoperability of the information exchange.

domain.

10.2 Approach for the interoperability of the information exchange

The author's approach aims to provide a solution (see Figure 10.1) able to represent connections or mappings between enterprises Native Business Services (NBS) (message elements of services) and semantic concepts described by ontologies (reference domain lexicon - knowledge). NBS are already existing software installed and exploited within the context of each supplier. The mappings with a reference lexicon can, then, be used to generate Supplier Business Services (SBS) accordingly to the elements of the domain ontology. SBSs are a set of features that are exposed from each supplier infrastructure and consist in the externalization of one or more NBSs. The generation of these SBSs promotes interoperability of enterprises applications in the domain allowing integration across multiple stakeholders (suppliers) and domains of interest (customers).

The first step to materialize this software-based system is to acquire a **platform reference lexicon** together with the domain experts. These are the main actors of their domain, thus their participation in the domain glossary (or reference lexicon) would increase the awareness of the domain terms. Glossaries are lists of specialized terms, mostly in alphabetic order, that sometimes are unique to a specific subject. Each term is composed by its corresponding description. It includes descriptive comments and explanatory notes, such as definitions, synonyms, references, etc. [72]. Then, a **mapping tool** that allows to represent both the domain glossary and the supplier services elements is required. This tool is intended to be used for administrative purpose (supplier's manager) and supports the process of mappings establishment and storage between the platform and suppliers' data models. These mappings will be stored in a specific ontology - the Communication Mediator (CM) (detailed in section 10.4). Finally, the **generator** is responsible for the generation of SBSs based on the mappings (represented in the communication mediator) created between the NBS and the platform ontology.

10.3 Modeling Morphisms to Enable Sustainable Interoperability

Models are used to capture the essential features of real systems by breaking them down into more manageable parts that are easy to understand and to manipulate. They are used in systems development activities to draw blueprints of the systems and to facilitate communication between different people in the team at different levels of abstraction. People have different views of the system and models can help them to understand these views in a unified manner [379].

In mathematics, Morphism (MoMo) is an abstraction of a structure-preserving map between two mathematical structures. It can be seen as a function in set theory, or the connection between domain and co-domain in category theory [380]. Recently, this concept has been gaining momentum applied to computer science, namely to systems interoperability. This new usage of *morphisms* specifies the relations (e.g. mapping, merging, transformation, etc.) between two or more information model specifications (M as the set of models).

In this context, the research community identifies two core classes of MoMo: non-altering and model altering morphisms [380, 381]. In the non-altering morphisms, given two models (source A and target B), a mapping is created relating each element of the source with a correspondent element in the target, leaving both models intact. In model altering morphisms, the source model is transformed using a function that applies a mapping to the source model and outputs the target model [382]. Other relations, such as the merge operation, can also be classified as model altering morphisms, however they are not detailed in this work.

10.3.1 Knowledge Enriched Tuple for Mapping Representations

The research community has developed many proposals to morphisms representations [383]. As analysed in [303], graph theory has been used in some, although other theories can be considered to achieve the envisaged goals, e.g., set theory [384], model management [385], or semantic matching [336]. However, there is not a single perfect solution that can be used to achieve all the morphisms goals at once. Some are ideal for structural issues, others for semantics providing good human traceability, and others are more formal and mathematical based. For that reason, in this work, is used a 5-tuple mapping expression (Mapping Tuple - MapT), presented in [303], with the goal of consolidate existent approaches to morphisms:

$$< ID, MElements, KMType, MatchClass, Exp > \tag{10.1}$$

being:

- ID - unique identifier of the $MaptT$;
- $MElements$ - pair indicating the mapped elements;
- $KMType$ - stands for Knowledge Mapping Type and determines the type of mappings represented in a specific instance of $MapT$;
- $MatchClass$ - stands for Match/Mismatch classification and depends on the $KMType$:
  - if $a = b$, the mapping is absolute and $MatchClass$ = Equal;
  - if $KMType$ = conceptual, the mapping is relating terms/concepts: $\{Equal, Naming, Coverage, MoreGeneral, LessGeneral, Disjoint\}$
  - Otherwise the mapping is either non-existing or more concrete addressing structural issues
- EXP - stands for mapping expression that relates and further specifies the previous tuple components. Normally, this expression can be translated to an executable transformation language (e.g. ATL).

10.4 Communication Mediator

The Communication Mediator Ontology has been built up as an extension of the Model Traceability Ontology defined by [386], and it is able to represent ontology semantic operations, like: (1) Semantic mismatches; (2) Semantic transformations; (3) Ontologies mappings; and other ontologies operations. Thus, the Mediator Ontology (MO) is able to log ontology and entity operations in a way that is possible to trace changes in all the ontology life cycle. It addresses traceability as the ability to chronologically interrelate the uniquely identifiable objects in a way that can be processed by a human or a system.

The mapping relations can be related to a traceability element, in such sense that a specific term defined in the reference ontology has a related one in the organization member ontology, making possible a way to trace ontology elements. This way, the morphisms are modelled with traceability properties in a sense that they enable to store different versions of information model elements, as well as mappings between specific objects defined in a model or ontology [336].

The MO structure has two main classes: 'Object' and 'Morphism'. The 'Object' represents any 'InformationModel' (IM), which is the model itself, and 'ModelElements' ($MElements$) (also belonging to the IM) that can either be classes, properties or instances. The 'Morphism' associates a pair of 'Objects' (related and relating), and classifies their relationship with a 'MorphismType' ($MType$), 'KnowledgeMappingType' ($KMType$) (if the morphism is a mapping), and 'Match/Mismatch' ($MatchClass$) class [336].

10.5 Model Transformations

The models mapping specifications can be performed either at a high level of formalization using graphs, sets, tuples, etc. or at lower levels, i.e. specifying the mappings by text. However, in both cases is necessary to implement them using a transformation language.

Atlas Transformation Language (ATL) is one of the most used transformation languages, having a large user base and being very well documented[387]. An ATL transformation is composed by a set of rules (matched rules) that define how the source model elements are linked to the target model elements. These elements can then be filled with information from the source model by so called rules (similar to functions in usual object languages like JAVA) and action blocks (blocks of imperative code which can be used by matched rules and called rules).

Model transformations are a current practice to enable interoperability among two organizations or ecosystems of organizations, each with their own service system. With them, one can specify P2P mappings to translate any data provided from one side's format specification into the other, thus allowing a seamless exchange of information.

When performing a horizontal model transformation (e.g. converting instances of a model to instances of another model) an explicit or an implicit mapping of the 'meta-model' has to be performed. Thus, as depicted in Figure 10.2 the idea is that when performing a transformation morphism at a certain level n, this transformation has (implicitly or explicitly) to be designed by taking into account mappings at level $n+1$. Once the $n+1$ level mapping is complete, executable languages can be used to implement the transformation at n level, e.g. using ATL and the Query/View/Transformation (QVT) [388].

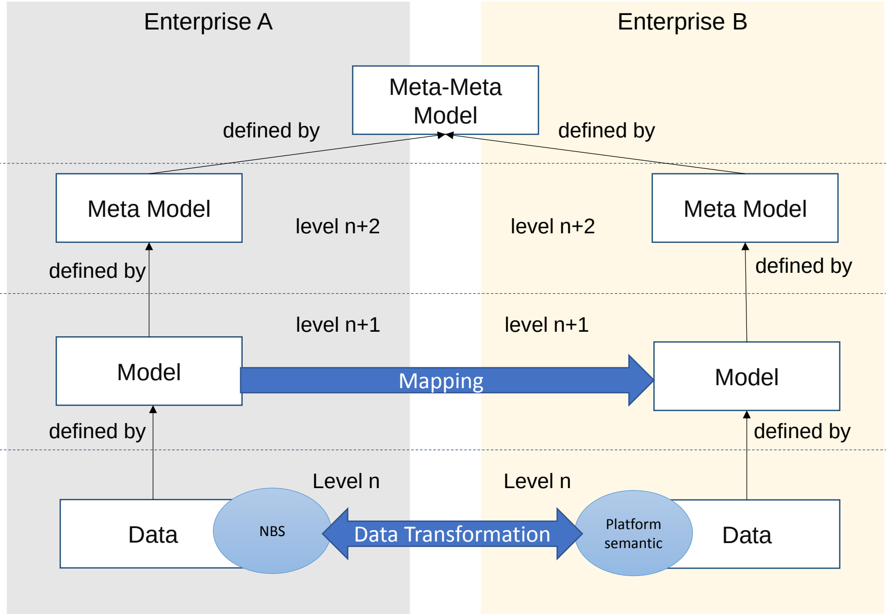


Figure 10.2: Method for Horizontal Transformation.

This type of transformations are normally static processes that once defined can be repeated any number of times achieving the same results. However due to the constant knowledge change caused by the dynamics of the global market (external environment), services and models that regulate enterprise systems, are not. In fact, some researchers have attempted to extrapolate results from a 'general systems theory' or 'complexity theory' that could explain the importance of evolution of systems in all fields of science [389]. These theories view all systems as dynamic, 'living' entities that are goal-oriented and evolve over time, thus, information systems, services and the mappings that connect them should be prepared to respond to the environment dynamics which is in a constant update. To support this dynamicity, horizontal transformations should be provided with traceability features, and mappings stored in a parseable and structured knowledge-base [303].

Since it is used a knowledge enriched tuple for mappings representation that are stored on a CM is possible to keep traceability of model mapping changes so that readjustments can be easier to manage, and data exchange re-established automatically using the model-driven development paradigm.

10.6 ISOFIN cloud Application Scenario

The ISOFIN cloud project envisage the establishment of technical solutions, guidelines and standards to support the interoperability resolution, facilitating a seamless sharing of artefacts at the knowledge and software level that would enable a fast and efficient creation of new Financial (e.g. bench and Insurance) products or services. Thus, ISOFIN's overall aims are:

- Analyse the problems resulting from the internalization of financial applications and identify potential solutions;
- Create ways of reducing the time that software developers take to achieve interoperability between domain financial stakeholders (e.g. bench and insurance);
- Standardize the domain semantic level, where no harmonization exists;
- To implement mechanisms to facilitate the generation of new services in a ubiquitous context that could easily be discovered and accessed by domain customers;

Based on the project aims, the result was the development of a platform that speeds up the development and deployment of services to be sold to the clients and users of the financial domain. For that, the steps identified in the approach presented in section 10.2 for the interoperability of the information exchange section must be followed: (1) acquisition of a platform reference lexicon together with the domain experts; (2) definition of mappings between enterprise NBS and the reference lexicon using a mapping tool; and (3) Generation of SBSs based on the established mappings.

10.6.1 Domain Reference Lexicon establishment

It is a fact that when an information system intends to represent a domain's knowledge it needs to be aligned to the community that it represents. Consequently, it is required to have a solution where community members could present their knowledge about the domain and discuss it with their peers. Additionally, such knowledge must be available and dynamically maintained by all the involved actors [56].

The solution adopted by the author for the domain reference lexicon establishment is based on MENTOR - methodology for enterprise reference ontology development [336], supported by an interface accordingly with the works presented in [56, 390] to implement the following MENTOR's steps (1) terminology gathering; (2) glossary building; (3) thesaurus building. The result is a domain lexicon, whose semantics are constantly refined through a specific front-end in order to handle the knowledge provided by the domain experts.

10.6.2 Mapping Tool

The used Mapping tool consists in an adaptation of a previously work with the aim of provide a graphical means to define different kinds of mapping between models [391].

This tool was developed with the intention of open Language Independent Meta-Model (LIMM) files, an abstract interface on top of information systems.

JGraph has been elected and modified to read the input information model files and store morphisms at the MO ontology. It is a widely used open source project for graph visualization and manipulation, similar to Microsoft Visio, with good documentation and several examples. Features include a complete selection of layouts to automatically position the graph, many styles of shapes and edges, validation of connections, as well as an undo and redo manager.

Some adjustments were made to enable the interaction (mapping definition) between two different information models' graphs, and to become integrated to the Communication Mediator Ontology. To the integration with the CM ontology, JENA was used - a Java API for OWL providing services for model representation, parsing, database persistence and querying ontologies.

In the ISOFIN's application scenario. the mapping tool supports the mappings definition between NBSs elements (e.g. services and message concepts names) described in WSDL 1.1[3] and Schema Definition Language (XSD)[4], and the ISOFIN's reference lexicon described using OWL 2.0[5] For the purpose, the enabled features are the ones represented in Figure 10.3 use case: 1) Open wsdl file; 2) save mappings; 3) import mappings; 4) generate WSDL file. The import of the platform reference lexicon is made when the open WSDL file option is triggered. That means, when the user decided to open a WSDL file for mappings establishment, both the WSDL and reference lexicon (downloaded from a domain server by calling a provided service) graphical representations are presented side by side (see Figure 10.8).

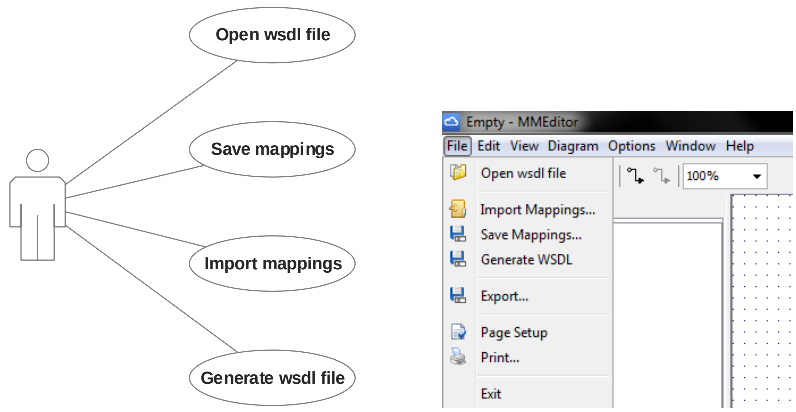


Figure 10.3: Mapping Tool Use Case.

In addition to the menu bar, three other areas compose the tool (see Figure 10.4):

[3]https://www.w3.org/TR/wsdl
[4]https://www.w3.org/TR/xmlschema11-1/
[5]https://www.w3.org/TR/owl2-overview/

- Mapping text descriptions - where are described the selected mappings and elements;
- Minimized view of the full mapping screen - allows to navigate in the full map and finally;
- Mapping screen - where is done the representation of the models and the mappings

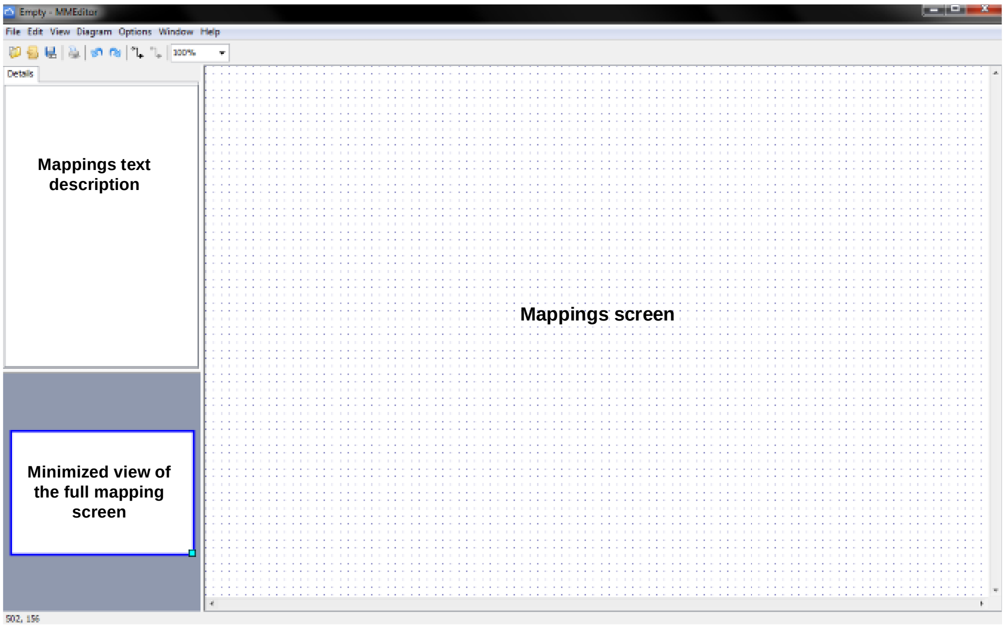


Figure 10.4: Mapping Tool working area.

Starting from the work presented in [391], some changes were made to the mapping tool to allow the interpretation of WSDL 1.0 (suppliers web services standard language) and OWL 2.0 (platform semantics description language). The most relevant changes are related to the transformation of both the input files into LIMM, consequently both the reference lexicon and NBS WSDL can be opened for mappings establishment

10.6.2.1 Language Independent Meta Model

This language enables the abstraction in relation to technologies and implementation details associated with the different modelling languages, and thus, enlarge the scope of users involved in a traditional mapping definition activity. Thus, having domain experts (suppliers) involved in the mapping process increases the quality of the mappings enabling the interoperability between suppliers and the rest of the platform participants.

10.6.2.2 OWL to LIMM transformation

The transformation of OWL 2.0 to LIMM already existed in the context of the European Project CRESCENDO [392] and explained in [391]. To enable the mapping among the OWL meta-model [393] and the LIMM, one needs to firstly put the OWL data in XMI serialization following the OWL meta-model specifications. Nevertheless, the procedure to do so is not straightforward as desirable since, in spite of the inputting OWL model is already XML serialized, it cannot be directly processed by the ATL toolkit which needs XMI as an input. The complete process for accomplishing the language mapping is illustrated in Figure 10.5, where the first step consists in doing an injection of the original model to an XML MOF meta-model specification. Following that, the second preparatory step consists in mapping that XML format to the reference OWL meta-model which will be the starting point for the actual $\theta(OWL, LIMM)$ language mapping (step 3). The, the step 4 takes place to execute the transformation between models itself using ATL.

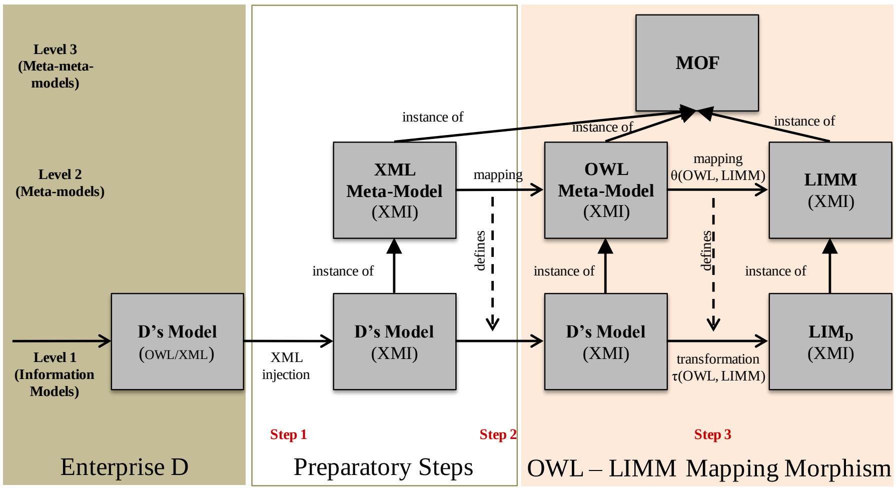


Figure 10.5: OWL 2.0 to LIMM Transformation [391].

10.6.2.3 WSDL to LIMM transformation

The transformation of WSDL 1.0 to LIMM follows the same approach of the transformation of OWL 2.0 to LIMM described before. Since not all the WSDL model elements are relevant for the described use case, only the elements depicted using yellow (dotted line) on Figure 10.6 where transformed. These are the elements containing the description of the services name, input and output messages, and the categorization of the messages. To allow the representation of services input and output messages, the container types were also considered (continuous line in red). The WSDL 'types' element is a container for XML Schema type definitions. The type definitions are referenced from higher-level message definitions in order to define the structural details of the message. To get more structured representation of the NBS content, the representation of the services was split in two parts, namely the types and the services (Figure 10.7).

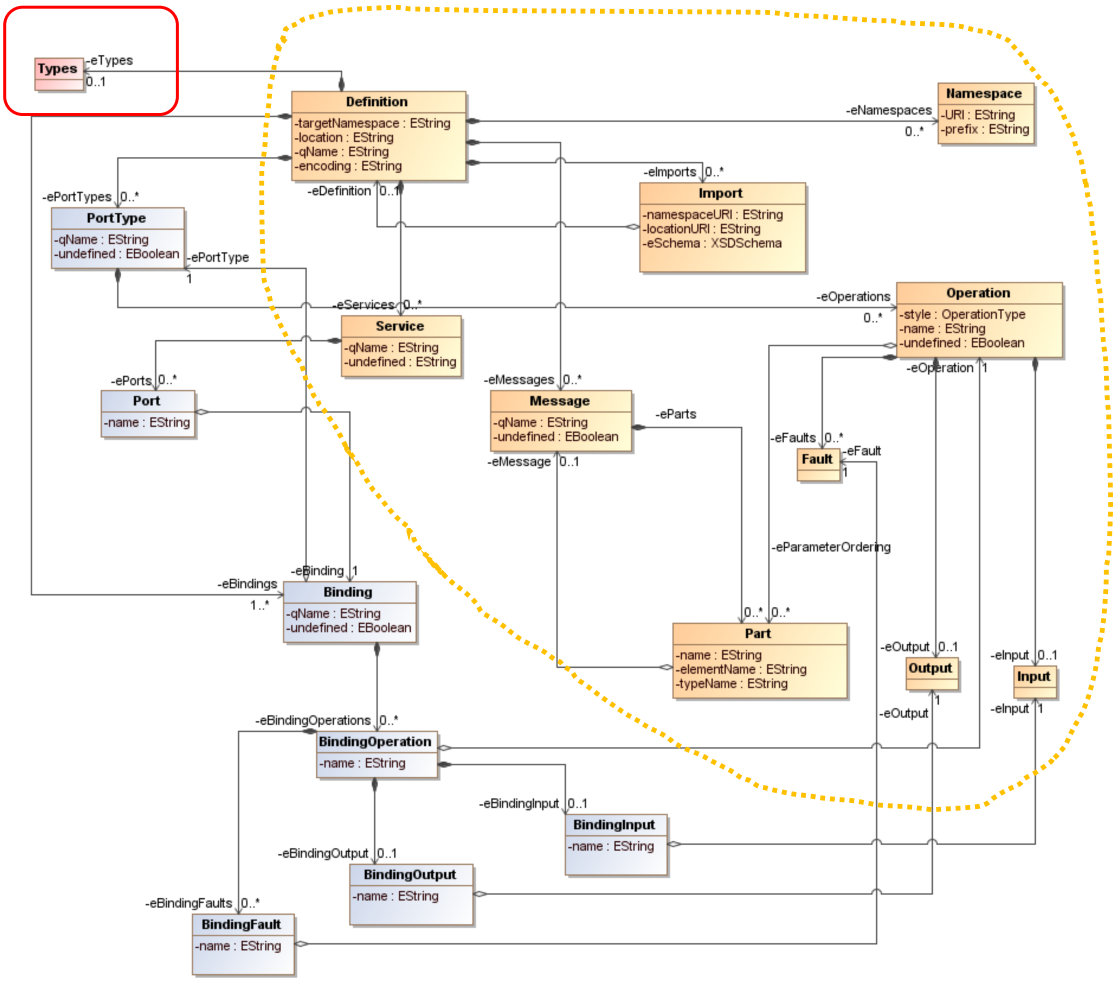


Figure 10.6: WSDL 1.0 Meta-model.

10.6.3 Mappings Storage

As mentioned before, the mapping tool provides a graphical means to define different kind of mappings while storing them in the Communication Mediator Ontology. Using the mapping tool developed, it is possible to open and show mappings between information models, namely, the NBS services and the ISOFIN reference lexicon, represented respectively in WSDL and OWL (Figure 10.8).

Since a mapping definition is a complex and time-consuming task, the Mapping tool is capable of halting the process at any time without losing the progress made so far by the user. This functionality still allows keeping mappings traceability.

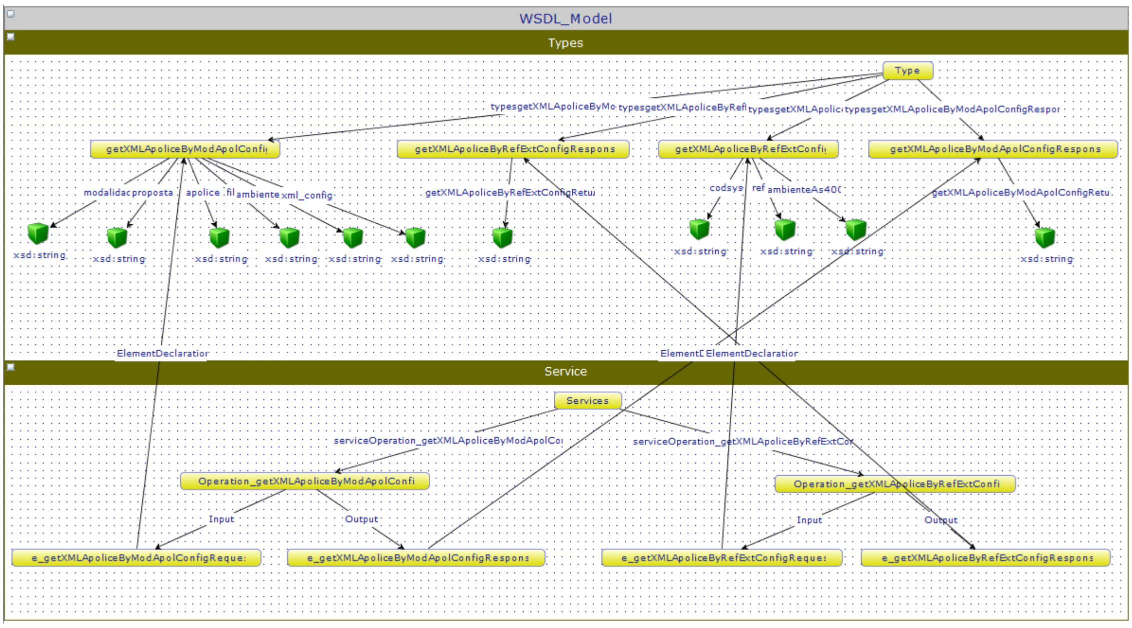


Figure 10.7: WSDL 1.0 LIMM Graphical representation.

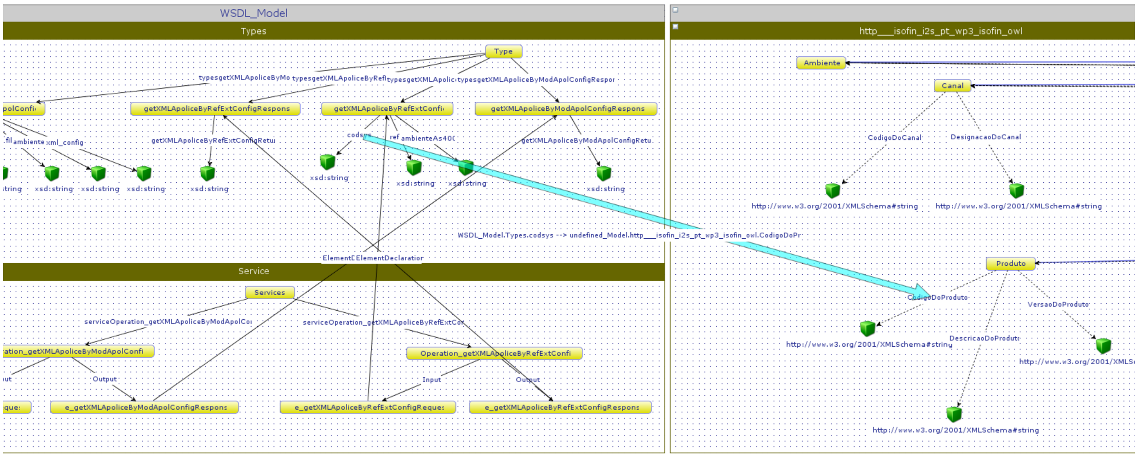


Figure 10.8: Mapping Tool demonstration.

10.6.4 Generate SBS WSDL file

To mediate information between the ISOFIN suppliers (NBS) and the ISOFIN platform it is needed to align the semantics of such information. To accomplish that, a generator to enable the representation of NBS services according to the ISOFIN nomenclatures was developed. This generator consists in a component of the mapping tool and it generates a WSDL of the supplier services accordingly to the ISOFIN semantics from the established mappings.

10.7 Discussion

The aim of this work is to provide interoperability integration on the information exchanged between domain suppliers and a specific platform, thus providing support to organizations' external processes. To achieve that goal, an approach capable to represent connections or mappings between information systems is proposed by the author. The outcome is that systems information subsystems are able to interact, but keeping their internal nomenclature thanks to the generation of Supplier Business Services.

In the application scenario, a mapping tool was used to support the establishment and definition of mappings between the ISOFIN ontology (knowledge) and supplier data models. These mappings are logged accordingly to a determined specification able to facilitate further traceability mechanisms implementation to facilitate maintenance of the established interoperability between the information systems. Such specifications are represented by a tuple that formalizes the mappings, which then can be stored as knowledge elements in the Communication Mediator Ontology. The stored knowledge elements enable the generation of SBS (WSDL) web service already compliant with ISOFIN's reference lexicon. Such new web services will facilitate the process of sending and receiving interoperable data between suppliers and the ISOFIN system.

However, due to the infinite possible model formats (meta-models) to represent any kind of information in web services contents composition (e.g. NBS), it is difficult to create a full interoperability solution. Thus, despite the fact of the complexity of reaching a complete automatic or dynamic interoperability it is concluded that partial solutions could be reached when it is acknowledged in advance the nature of the meta-models of the information exchanged and when there is a specification for traceability representation able to represent all the kinds of mismatches.

Chapter 11

Customers' Feedback Analysis

In contrast to earlier times, when finding sources of information was the key problem to companies and individuals, today's issues are that digital information is so easily shared and replicated that the amount of data and information available for organizations to analyse is exploding [308]. However, the analysis and knowledge gathering from this data is essential to help organizations to take suitable decisions [16]. As a response, the big data initiative seeks to glean intelligence from data and translate that into business advantage. According with the result of a survey conducted in 2012 by NewVantage Partners [13] there are seven groups of tangible benefits to achievable by big data initiatives. On the top of them (22% rate) are: better fact-based decision-making and improvement of customers' experience. The remaining ones are: increased sales, new product innovations, reduced risk, more efficient operations and higher quality of products and services.

The author contribution toward this issue is to enable the extraction and analysis of customer's comments focusing on their manifested sentiments. Sentiment analysis is associated to the automatic analysis of evaluative data and tracking of the predictive judgments [394]. This big data usage would allow organizations managers to take decisions based on evidences rather than intuition [308], being the evidence related to the customer satisfaction level of a specific service or product.

11.1 Framework for customers' sentiment analysis

Companies' web software layouts usually integrate a customers' review/comments area, where customers can express their sentiments related to the used products or services. These comments are used as inputs of the proposed framework for customers' sentiment analysis (see Figure 11.1). This framework encompasses **data collection**, **pre-processing** and sentiment **analysis** to be returned as useful knowledge and finally to support companies **decision-making** (accordingly with section 2.3).

The first step of the framework consists in the HTML pages' collection. Then, in the next step, each page HTML is processed to return customers' comments in a textual form.

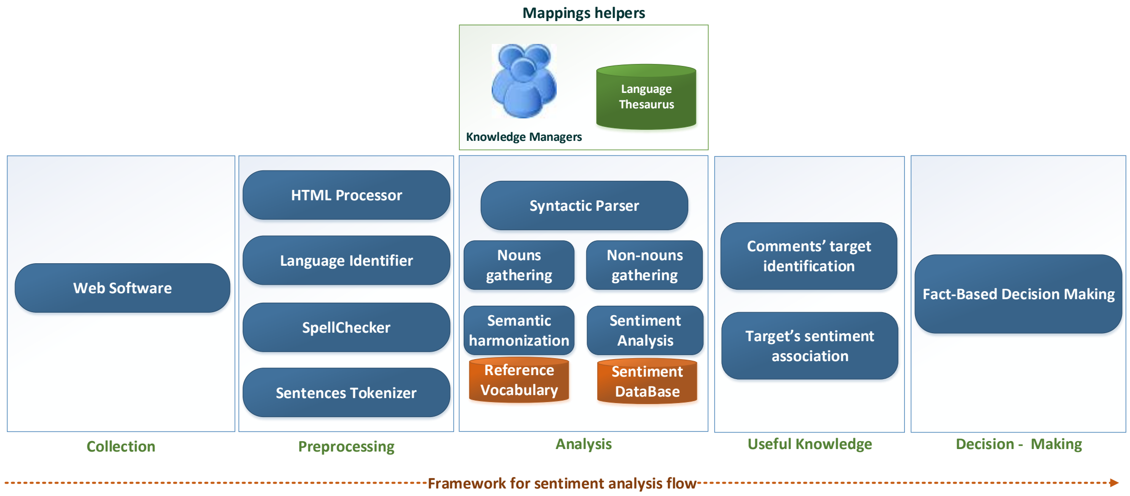


Figure 11.1: Framework for customers' sentiment analysis.

The essential issue in sentiment analysis is the identification of how sentiments are expressed in texts and whether the expressions indicate favourable or unfavourable opinions toward the subject [305]. Consequently, a typical approach to sentiment analysis is to start with a lexicon of positive and negative words and phrases [395]. However, in order to obtain an accurate sentiment analysis is required a sentence-level or even phrase-level sentiment analysis. Even though most of the sentiment extraction algorithms perform poorly in this aspect, without phrase-level sentiment analysis, the association of the extracted sentiment to a specific topic is difficult [396].

Since, in this framework we are interested in customers' reviews, which can be written both in the services (or products) providers' native language or in a foreign language, after HTML processing for comments extraction, a language identification is made. Automatic language identification is the process of using a computer system to identify the language of a spoken utterance [397]. An accurate language identification can facilitate the use of background information about the language and use of more specialized approaches in many Natural Language Processing (NLP) tasks dealing with a collection or a stream of texts, each of which can be written in a different language [398].

Since many web sites contain no spell checker, comments written by customers can have spelling errors. The existence of these errors complicates the natural language processing task. For that reason, in the author's proposed framework a spell checker to be used prior to comments analysis was integrated.

Most processing tasks in NLP, presupposes a preliminary phase of tokenization, after which the input sequence is provided with its individual tokens explicitly identified and isolated from each other [399]. Tokenization is the process of breaking up the given text into units called tokens. The tokens may be words, numbers or punctuation marks. Tokenization does this task by locating word boundaries (ending point of a word and beginning of the next word) [400].

There is a great interest of using syntactic information as part of an information retrieval strategy. Humans gather information not only based on the meaning of the words, but also based on the structure in which words are put together [401]. The knowledge of customers' comments syntax can be used to increase the performance of sentiment analysis task. Accordingly with [401], parsing intends the recognition and explanation of concatenation patterns that conform to a system of rules (grammar) and are presented to the analyser as a string (sentence) of consecutive units (words) which correspond to instances of a pre-established set (vocabulary). One of the possible output types of parsing process are parsing trees. These consists in an ordered, rooted tree that represents the syntactic structure of a string (sentence) according to some formal grammar [402].

At this point, it is possible to extract the nouns, adverbs, verbs and adjectives of customers' comments. The nouns will be assumed as customers' sentiment targets and the rest of the grammatical terms will be used to extract the polarity of the comment. For the sentiment polarity extraction, it is used a dictionary of the comments' language with polarities associated to each word. After the semantic harmonization process, the output is a common syntax and semantics to describe the customers' comments targets. The semantic harmonization process is implemented by merging or mapping the knowledge of the various sets of terms with the existing version of a semantic reference lexicon. This lexicon is built using the methodology for domain's glossary building and enrichment based on synonyms explained in the next subsection. The use of such kind of glossaries increases the probability of matches between nouns and classifiers from customers' comments with a specific reference lexicon, which will result in a greater number of concepts that the framework can handle.

Finally, the knowledge output of the framework is a common vocabulary to identify the targets of the comments, the adjectives adverbs and verbs, and the polarity associated. Thus, companies' managers will know: where to act (comments' target with negative polarity associated) and how to act (comments' adjectives, adverbs and verbs). This will result in a fact-based decision making that enables organizations to improve their services and products quality (meeting customers' expectations).

11.1.1 Methodology for Domain's Lexicon Building

In this work, the author is mostly interested in a lexicon building that considers the taxonomic representation of concepts and its enrichment by adding synonyms to them. The organization of domain's terms in a taxonomic form allows the classification results from deeper nodes of the taxonomy to be propagated up to help classification of their ancestor.

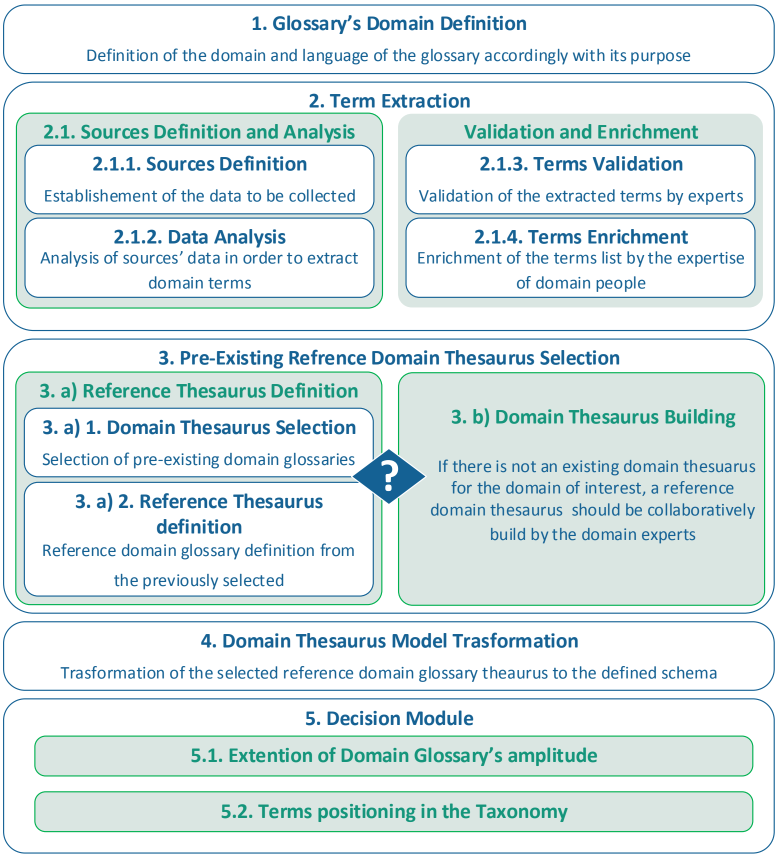


Figure 11.2: Methodology for Glossary Building and Enrichment.

As can be observer in Figure 11.2 the methodology is composed by five main steps: 1) Glossary's Domain Definition; 2) Term Extraction; 3) Pre-Existing Domain Thesaurus Selection; 4) Domain Thesaurus Model Transformation; and 5) Decision Module.

The first step, Glossary's Domain Definition, encompasses the specification of the language and domain in which the glossary intends to be applied.

The second step consists in domain's terms extraction from several sources. To acquire useful terms, at first, it needs to define the data sources to be collected. Since domain experts are most likely aware of their domain's sources of information, they should be the ones to provide input sources accordingly with the chosen language and domain. Then, these are analysed to extract relevant terms that will be used to build the domain lexicon. Domain experts should also validate the extracted terms and add other pertinent ones.

The third step of the methodology can be divided in two distinct branches accordingly with the existence (or not) of pre-existing domain thesaurus for the specified language. If some are already available for the domain of interested, they should be considered. Then, a reference thesaurus from this set needs to be selected or built. Regarding the non-elected thesaurus its terms need to be extracted and added to the collected terms of step 2 of the methodology (Terms Extraction). If there is not any pre-existing thesauri it should be collaboratively build by domain experts (step 3.b), using, as an example, the methodology explained in [390] and chapter 7 where the authors created a user friendly wiki interface to allow thesaurus building by domain experts.

However, to build a good domain thesaurus structure, is not enough to just collect the domain terms or pre-existing domain glossaries and relate them. It is also necessary to define a model to formalize the domain's lexicon. The adopted thesaurus model is represented both in Figures 11.4 and 11.5, where each term of the reference nomenclature lexicon corresponds to a class in the taxonomy. For each reference term a synonyms list is then created and formalized as a set of annotation properties named as 'synonym'. The step 4 of the methodology consists in the transformation of the selected reference domain thesaurus to the defined model schema. To avoid information losses, both knowledge engineers and domain experts should support the transformation between models. One possible example is related to the adoption of a domain thesaurus with no indication of the reference nomenclature in a bulk of terms. Like in OntoPT [403], the representation of synonyms could be made using object properties (synsets). If such kind of thesauri is adopted, the domain experts should define the reference terms inside the represented lexicon, so knowledge engineers could proceed to the transformation between models, handling as many annotation synonyms as the existing object properties (synsets) for the selected reference term.

The last step of the methodology consists in the Decision Module, which purpose is to extend the selected reference domain lexicon amplitude and the correct placing of terms in the taxonomy. The amplitude increasing is done through dictionaries usage and supported by domain experts. As can be observed in Figure 11.3, the synonyms gathering module runs cyclically, retrieving several intermediate synonyms lists in order to obtain the final synonyms list.

In the synonyms gathering module (part of step 5.1), illustrated in Figure 11.3, the total of dictionaries are queried to gather all the available synonyms for the input term. More synonyms are also collected from the domain experts, which can use their expertise in the domain to complement the available information. The terms that were already input of the synonyms gathering module are marked to not be treated again. This process is repeated cyclically, being the input of the synonyms gathering module each of the terms that were not processed yet. When all the terms are processed, the output is the final synonyms list.

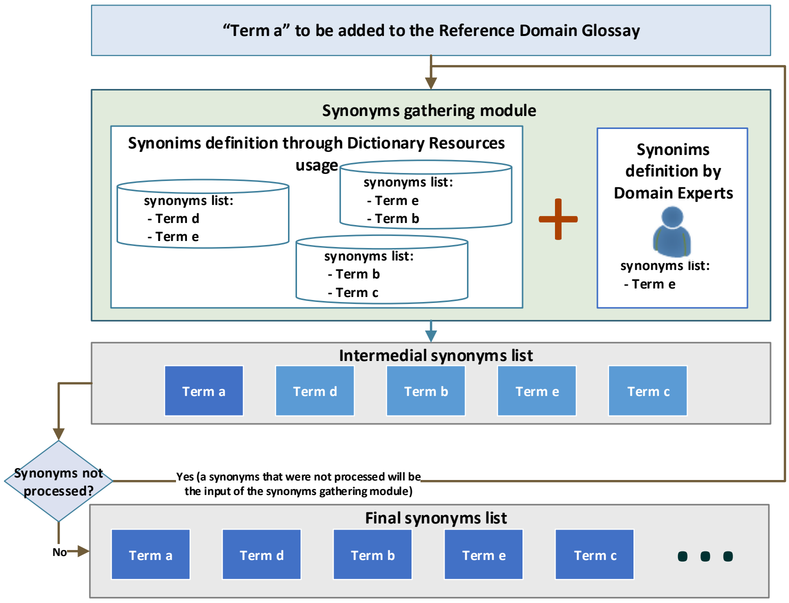


Figure 11.3: Synonyms gathering cycle.

The second step of the Decision Module (step 5.2) encompasses the right placing of the synonyms list in the domain thesaurus. In this step two situations may occur: 1) one of the terms of the synonyms list is already in the domain thesaurus; and 2) none of the terms of the synonyms list is yet in the domain thesaurus.

The first occurrence scenario is illustrated in Figure 11.4. In this case is possible to verify that the Term b, part of the gathered synonyms list, is already a term of the domain thesaurus.

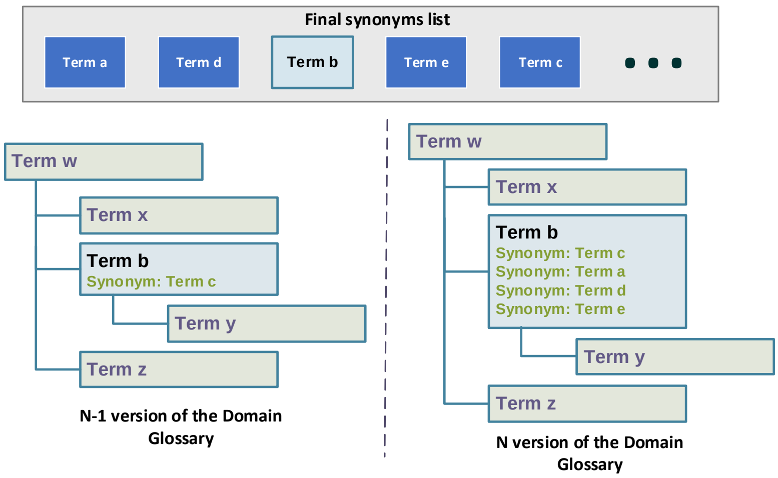


Figure 11.4: Term placing in the domain thesaurus: one of the terms of the synonyms list is already in the domain thesaurus.

The other scenario consists in terms placing in a thesaurus that does not contain any synonym of the input term, both in the reference domain and its corresponding synonyms list. In this specific scenario, domain experts and knowledge engineers should collaborate to place the synonyms list in the right place of the taxonomy. As can be observed in Figure 11.5, the first step consists in the selection of the reference term from the list of synonym terms by the domain experts and the specification of the ancestor of the selected reference term from the set of the thesauri already existing terms. In the figure is possible to see that the Term b was selected as the reference term and its ancestor is the Term w. Then, after these specifications from the domain experts, knowledge engineers should update the domain's taxonomy accordingly. The output of the methodology is a formal domain lexicon able to provide a better understanding of the domain to which is related.

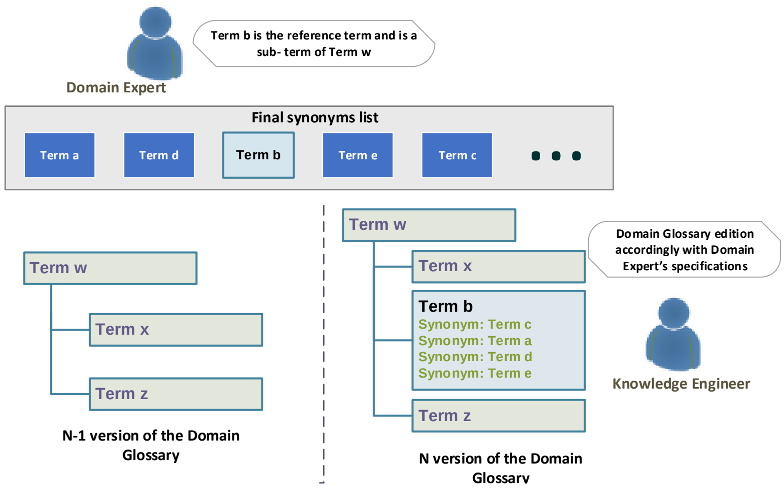


Figure 11.5: Term placing in the domain thesaurus: none of the terms of the synonyms list is yet in the domain thesaurus.

11.2 Portuguese hospitality industry customers' comments analysis

Semantik project aimed at developing a platform of services, in a Software as a Service logic, to allow the usage of structured, semi-structure and unstructured data from the web to serve enterprises management necessities. It proposed the use of semantic web (Web 3.0) methodologies and technologies to help enterprises to identify, analyse and classify information in a more correct and effective way. Thus, Semantik allowed enterprises to take advantage of the enormous amount of available information (and knowledge) on web using automated data collection and presentation. It was done by organizing information from various sources using a decision support framework which included intelligence classification according to each organization business model, signal strength and management level.

The chosen application scenario was the hospitality industry were web technologies have an enormous impact on marketing and sales strategies. As part of the project developed work, the previously presented framework was used to develop a methodology for customers' comments analysis. The output of the methodology were both the targets and their associated polarity (customers' sentiments about specific services and/or products). Based on that information, the organization managers can improve the decision-making by directly address customers' opinions.

For this scenario, Booking[1] was the web software used to analyse hospitality industry customers' satisfaction. Consequently, the input of this methodology was Booking's web software whose layout integrates a customers' review/comments area (see step 1 of Figure 11.6), where customers post their opinion about the services provided.

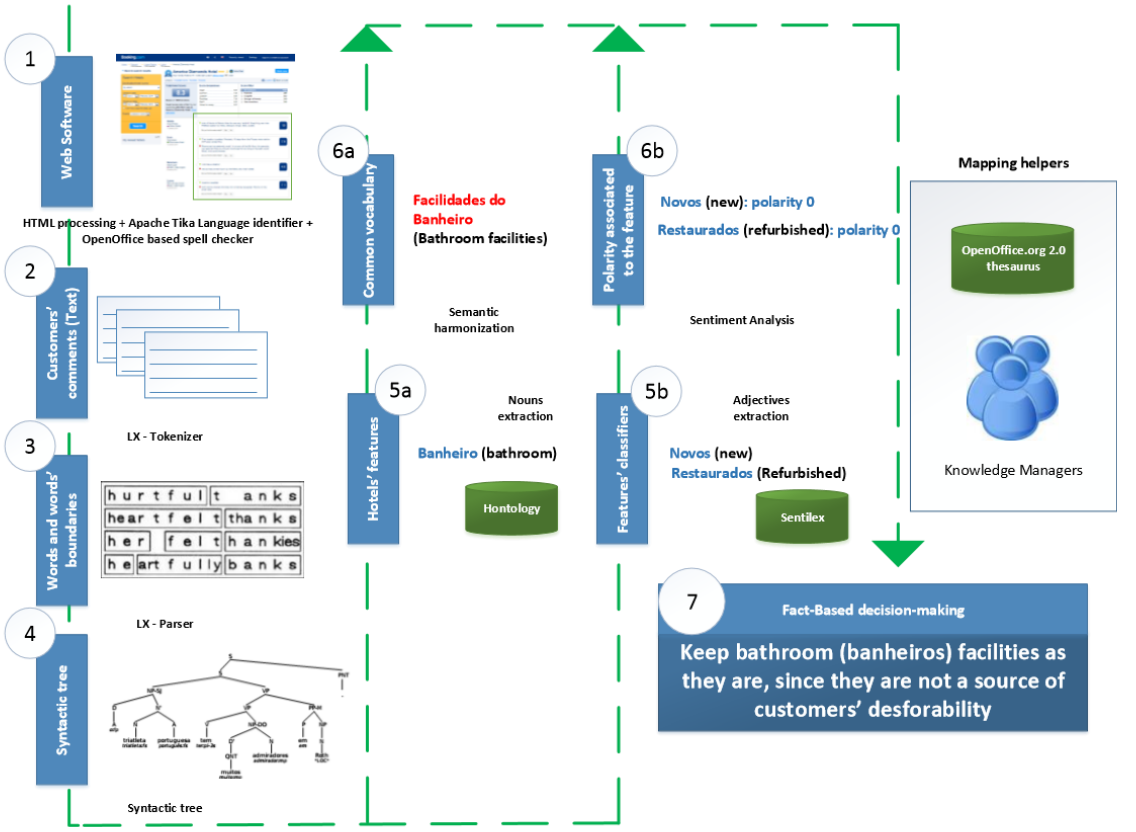


Figure 11.6: Methodology for Portuguese hospitality industry customers' comments analysis.

11.2.1 Data collection and pre-processing

The first step of the proposed method encompasses the data collection. For that, some HTML processing is required to identify in bookings layout the comments area. The step

[1]http://www.booking.com/

two begins the processing of the comments with the language processing and spell checker application. The comments used in the next steps are the ones written in Portuguese. The use of a language identifier was no need, since the comments' language were already presented in comments' metadata. Due to in this scenario we were only interested in Portuguese comments, a spell checker based on Open Office Portuguese dictionary was used.

In the third step was made the tokenization of the extracted sentences to allow the application of the syntactic parser. This is made with LX-Tokenizer[404], which is a tokenizer that takes into account Portuguese non-trivial cases that involve ambiguous strings. The outputs of LX-Tokenizer application to the positive comment are the correspondent words and words' boundaries.

11.2.2 Data analysis

In the fourth step, it was made the syntactic parsing of the comments through the LX-Parser. LX-parser is a freely available online service for constituency parsing of Portuguese sentences. The parser was trained and evaluated over CINTIL-Treebank, a treebank produced from the output of a deep processing grammar by manually selecting the correct parse for a sentence from among all the possible parses produced by the grammar [405, 406]. The parser produces several outputs that consists in the several phrases that compose the full customer comment. This functionality allows the isolation of the several targets of interest, since they usually are in separated phrases or tree branches, and individually analysis.

Taking as an example the comment *'Banheiros novos. Reformados'*, which translation is 'New bathrooms. Refurbished', the application of the syntactic parser is presented in Figure 11.7.

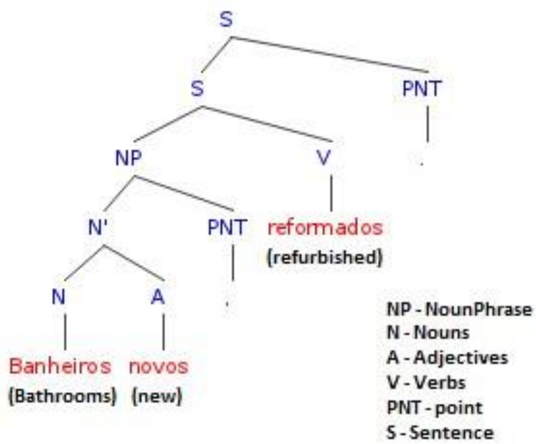


Figure 11.7: Example of phrase syntactic tree.

By Analysis of the phrase using a syntactic tree is possible to verify that the phrase is composed by a sentence (S) and a point (PNT). The sentence is then composed by a noun phrase (NP) and a verb (V). The noun phrase contains the noun (N) 'banheiros' and its adjective (A) 'novos'. By sentence analysis, the verb of the sentence, 'reformados', can also be associated to the noun. Thus, based on the parsing of Figure 11.7 phrase is possible to high-light the target of the comment (banheiro - bathroom) and the corresponding sentiment indicators (novos - new, and reformados - refurbished).

The fifth step is dived in two distinct parts, one for the semantic harmonization of the targets (5a) and other for the sentiment indicators analysis (5b). It is done through mappings establishment between comments terms and a domain's lexicon obtained using the methodology presented in section 11.1.1.

11.2.2.1 Portuguese hospitality industry lexicon establishment

In this application scenario, domain experts also suggested Booking as a relevant source for terminology acquisition (methodology's step 2.1.1). In Booking's pages content, namely hotels and rooms descriptions, the gathering of terms can be easily done by HTML processing, because terms are placed following a very specific layout (see Figure 11.8 representing methodology's step 2.1.2).

Comodidades dos quartos: Cofre, Ar Condicionado, Duche, Secador de Cabelo, WC, Casa de Banho, Telefone, Canais por Cabo, Mini-bar

O acesso Wi-Fi está disponível gratuitamente em todos os quartos.
Área do quarto: 25 m²
Tipo(s) de Cama(s): 1 cama de casal grande ou 2 camas individuais

Figure 11.8: Excerpt of Booking room description.

Hontology was selected as the reference domain thesaurus source (methodology's step 3.a). Hontology is a multilingual ontology for the hospitality industry in the tourism industry [407]. The Portuguese nomenclatures were originally placed as an annotation property of each class of the Hontology's. Thus, the first step of the transformation process (step 4) encompasses the exchange of class's name to their respective Portuguese annotation (see Figure 11.9). After this step, Portuguese reference terms appeared in a taxonomic form (thesauri). Then, the next step encompasses the enrichment of Hontology's terms with the ones extracted from booking website and selected dictionaries, (step 2.1.2 of Figure 11.2 methodology and represented in steps 2 and 3 of Figure 11.9) together with domain experts suggestions. In this application scenario the chosen dictionary was OpenThesaurusPT[2], which is an open source project to the development of a synonyms dictionary for the Portuguese language.

[2] http://openthesaurus.caixamagica.pt/

Figure 11.9: Portuguese hospitality industry glossary building demonstration scenario.

In Figure 11.9 is also possible to observe the attribution of the several annotation synonyms to the term "Hostel" gathered from booking (methodology's step 5.1). For this term, domain experts suggested "Hospedaria" as a synonym. Regarding the term "Hostel", the used dictionary (OpenThesaurusPT) didn't retrieve any synonyms, however, for the synonym defined by the domain experts ("Hospedaria"), the collected synonyms were: "Albergaria", "Albergue", "Estalagem", "Hospedagem" and "Pousada".

Since the term "Hostel" was not already in the chosen reference domain thesaurus, the placing of the concept was made both by the domain experts and knowledge engineers (Figure 11.2 step 5.2). In the third square of Figure 11.9 is possible to verify that the term was placed under the class "Acomodação" (Acommodation) and each of its synonyms were added as annotation properties.

Hostel synonyms can be useful when analysing domain competitors. Considering that a domain actor has an accommodation classified as hostel and is placed in the region of Lisbon. The competitors' recognition can be made by looking for all the accommodations in the same region, with the same star classification, and classified as Hostel or any of its synonyms. Another advantage of methodology's output is the hierarchical representation of hospitality industry lexicon which enables the propagation of the sentiment assigned to an object to its children.

Then, for each noun detected in a customer comment the sentiment analysis methodology tries to do the mapping with the domain lexicon concepts. If the mapping does not exist, the domain experts and knowledge engineers should act to properly increase the lexicon range with it.

As an example of the domain lexicon application, Figure 11.6 shows the mapping between the comment concept 'Banheiros' (Bathroom) and the reference term 'Facilidades do Banheiro' (Bathroom Facilities).

11.2.3 Knowledge Acquisition and Decision Making

For comments polarity identification, the author opts for using a list of adjectives, verbs and adverbs and their polarities, namely Sentilex. Sentilex was built with a methodology to automatically enlarge a Portuguese sentiment for mining social judgments. Sentilex is a Portuguese lexicon with 7.014 lemmas and 82.347 inflected forms [408]. The polarity attributed by sentilex is: positive (1), negative (-1), or neutral (0). The mapping of the comments classifiers with sentilex terms is made in a similar way of nouns harmonization with the domain lexicon, meaning that the range of known terms by sentilex is extended using a dictionary (OpenThesaurus).

Giving as inputs the classifiers 'novos' and 'reformados', the attributed polarity from sentilex is 0 for each one. Based on Sentilex sentiment retrieval, and since both sentiment indicators have neutral polarity (0), the decision to make (step seven) is to keep bathroom facilities as they are, since, in this case, they are not a source of customers' unfavourability.

11.3 Discussion

The proposed Framework establishes a set of components that aim to guide companies to customer's sentiment analysis. Its main characteristic is the ability of gather customer's data from web software and analyse the associated sentiments. This allow companies to make fact-based decision making to improve the services and products provided.

In the example of Semantik project, focused in the Portuguese hospitality industry, a methodology was implemented to retrieve both the targets and target's associated polarities from booking customer's comments. The methodology output can, then, feed the decision-making process, contributing to context awareness decisions. It aims to help hospitality companies to analyse and identify useful knowledge in a more effective and establish fact-based decision-making able to improve the quality of services and products. The proposed framework can be applied to several domains being only a matter of choose another web source and reference ontology for the vocabulary harmonization. The framework can also be applied to another language by changing the tokenizer, syntactic parser and polarity dictionary.

CHAPTER 12

CONCLUSIONS

Although the presented research is very ambitious, the need of organizations and individuals to create and employ mechanisms to search and retrieve data from the huge amount of available information and mine it into knowledge makes it an important research challenge. It is important for organization to be capable of identifying its sources of knowledge, whether tacit, or present in domain documents, and use them for decision-making support and new strategies definition. However, the lack of motivation of individuals to share knowledge and the difficult processing of available documents may lead to a difficult knowledge acquisition.

Other challenge identified is the need for organizations to collaborate in order to survive in the emerging society and competitive markets. Organizations are willing to improve their competitiveness through partnerships, thus interoperability between systems and applications is required. One of the main problems regarding the interoperability between systems and applications is related to the high number of semantic representations of the same segment of reality - it results in a difficult semantic interoperability achievement. Thus, a mean to enable the knowledge transference between organizations without losses due misunderstandings is needed. To face this issue, the author consider that a possible solution is to use a common reference lexicon to be the intermediary in the communications between enterprises and to the outside. Then, the establishment of mappings between enterprises and the reference domain lexicon will allow each of the enterprises to keep its own knowledge and semantics unchanged, and still able to smoothly interact with its domain.

Several topics where identified as part of the solution to be formulated during the PhD work, namely: 1) Knowledge Creation Process; 2) Knowledge Management and

Collaborative Systems; and 3) Business Intelligence. These supported the establishment of a knowledge management framework (proposed in chapter 5).

The validation of the overall work was accomplished by the integration of its results in research projects (bridge to both industrial and scientific domains) depicted in 12.2, and scientific acceptance by peers as detailed in section 12.3.

Table 12.1: Outcomes related to the Research Questions and Sub-Questions.

Research Question	**Development and Outcomes**
How to facilitate tacit to explicit knowledge transference?	1. A new approach was proposed for organizations knowledge management (chapter 6). This approach supported the implementation of several tools like: • Wiki-based knowledge management framework - Supported FInES community in domain experts knowledge acquisition and integration with other research activities (chapter 6) • Wiki-based glossary building tool - Supported domain experts in their domain glossary building creation (see chapter 7) • Requirements Management tool - Supported both business and technical requirements elicitation and linkage with organization's processes (see chapter 8)
How to ease knowledge sharing between organizations multi-disciplinary teams?	1. A new approach was proposed for organizations knowledge management (see outcomes of RQ 1). 2. A organizations processes modelling tool that allows the collaboration and knowledge transference between teams was proposed and developed (see chapter 9). It allowed knowledge sharing between business and technical teams during the processes design.
How can knowledge be effectively transferred, within and among organizations, facing the high number of semantic representation of the same segment of reality?	1. A new approach was proposed for organizations knowledge management (see outcomes of RQ 1). 2. A methodology for domains lexicon building was proposed (see chapter 11). It allows both: • Organization of domain's terms in a taxonomic - the classification results from deeper nodes of the taxonomy to be propagated up to help classification of their ancestor. • Synonyms integration - The range of concepts that domain participants (e.g. humans and pieces of software) is capable to interpret is increased. 3. A services modelling and information exchange approach and consequent tool (see chapter 10). Its main outcomes are: • Supporting of mappings establishment between organizations' and suppliers' services. Thus facilitate the process of sending and receiving interoperable data between suppliers and the organization. • Traceability of the established mappings to further maintenance.

12.1 From a research question to validation

That question was then decomposed in three sub-questions whose answers were considered by the author as the key to solve the main research question. Those are related to: 1) the creation of mechanisms to facilitate tacit to explicit knowledge transformation; 2) knowledge transference between multi-disciplinary teams while collaborating; and 3) knowledge transference between individuals with different background while facing a high number of semantics to represent the same segment of reality. The main outcomes of this research work organized in relation to its research questions is presented in Table 12.1.

12.2 Integration with other research activities

The research work was successful integrated in several projects. The following subsections describe the main achievements of the research integration as well as some resulting publications.

12.2.1 ENSEMBLE

The research performed under ENSEMBLE aimed in the establishment of an approach able of gathering tacit knowledge from the FInES community and transforming it into explicit knowledge, so that it could be available to the full community.

The proposed approach (see Chapter 6) proposes the usage of simple web-based interfaces, in the form of wiki modules, which allow domain experts to contribute with their tacit knowledge through an intuitive front-end. That knowledge is then transformed into explicit knowledge, in the form of ontologies, following a semi-automatic methodology. With this process, knowledge becomes available for querying and intelligent reasoning. Other knowledge bases can be integrated, providing users extended awareness of the domain and enriched feedback information that can motivate the refinement of the front-end and more suitable decisions.

Main achievements

- Increase of the number of participants in the collaborative knowledge gathering (Motivation for knowledge sharing);
- Increase of Quantity and Quality of the knowledge shared by the community;
- Integration of the domain community with other initiatives.

Publications

- Knowledge Management Framework using wiki-based front-end modules [352]
- Collaborative Knowledge Management Using Wiki Front-End Modules [409]

12.2.2 ALTER-NATIVA

ALTER-NATIVA has an e-Learning repository of accessible learning objects, which aims to organize, store and retrieve educational resources produced by the members of the ALTER-NATIVA network. The virtual learning objects in the repository are organized into areas such as science, mathematics, and language. Then, it is necessary to facilitate the categorization of virtual learning objects and professors' expertise from community members (e.g. professors) using a common and well known lexicon. The main problem is related with the constant insertion of virtual learning objects and professors in the platform that can be categorized with keywords, which are not available in the reference nomenclature (lexicon), resulting in the 'crash' of the established semantic interoperability. In this situation, the need of facilitating community members to actively participate in the constant domain knowledge updating (manage) process is observed. Consequently, the contributions to ALTERNATIVA (see Chapter 7) were to:

- Provide to domain experts the necessary tools and mechanisms to constantly update the reference domain lexicon.
- Provide to domain experts the ability to categorize both the virtual learning objects and professors using the established domain lexicon. It can be used for:
  - Search for professors and/or virtual learning objects directly based on the keywords that characterize them.
  - A professor of Mathematics normally has interest in differential calculus. Thus, next time a professor of this profile enters the system, the system should automatically recommend to him a set of resources in accordance (e.g. VLO or users related to differential calculus or with one of its ancestor in the reference classification taxonomy).
  - Recommendation of users to interact with, based on common interests and levels of expertise.

Main achievements

- Semantic interoperability in the domain;
- Accuracy of the recommendation system;
- Quality of the established domain reference lexicon;
- Solution capable to represent and refine the established domain lexicon.

Publications

- A Knowledge Management Framework to Support Online Communities Creation [410]

- The ALTER-NATIVA knowledge management approach [72]

12.2.3 ISOFIN

In ISOFIN's scenario was necessary to develop a mechanism capable to represent the connections/mappings between ISOFIN's suppliers' services (NBS) and a ISOFIN's reference nomenclature. NBS are the already existing software installed and exploited within the context of each ISOFIN supplier. The mappings can, then, be used to generate SBS accordingly to the elements of ISOFIN's ontology. SBS's are a set of features that are exposed from the ISOFIN Supplier infrastructure. A SBS is the result of externalizing of one or more NBSs. The creation of SBSs is dependent upon the ISOFIN Supplier strategy (business needs) and/or legal requirements that it is required to fulfill.

The generation of these SBSs promotes interoperability of applications in the financial domain allowing integration across multiple stakeholders (ISOFIN suppliers) and domains of interest (ISOFIN Customers). It can be regarded as a software-based system that allows information subsystems to interact together keeping ISOFIN suppliers nomenclature, but able to interact to the ISOFIN Platform thanks to the generation of SBS services.

The author research work can contribute to this scenario by providing a solution capable to represent connections or mappings between systems, or more precisely, between services elements and semantic concepts described by ontologies (reference domain lexicon). The implemented solution had in consideration the ability of performing a trace of the mapped elements. It will allow to domain users to adapt their mappings accordingly with observed changes in the environment (e.g. changes in the reference nomenclature, new services' parameters).

Main achievements

- Services mappings traceability;
- Quality of the established domain reference lexicon;
- Solution capable to represent and update the established service mappings.

Publications

- Model-driven Approach for the Interoperability of Enterprises' Services Information Exchange [411]

12.2.4 OSMOSE

The OSMOsis applications for the Sensing Enterprise (OSMOSE) project has the objective of developing a reference architecture a middleware and some prototypal applications for the Sensing-Liquid Enterprise, by interconnecting Real, Digital, and Virtual Worlds in the same way as a semi-permeable membrane permits the flow of liquid particles through itself [355]. The worlds represent three types of data management environments: Real World – related to data that comes directly from devices that is handled by physical components; Digital World – related to data management available in data and knowledge bases or Internet (big data); and Virtual World – related to specific management of data with the support of artificial intelligence related programs for specific simulations.

Under this projects, two scenarios were considered: Requirements Management and Processes Modelling.

12.2.4.1 Requirements Management

This scenario follows the necessity to implement a Requirements Management approach and tool capable to handle elicitation and/or characterization and further refinement of requirements. Then, technical requirements derivation from the set of approved requirements should be conducted. Those technical requirements can also be associated to OSMOSE architectural components. Finally, the requirements are formalized and handled in an ontology allowing simple queries or advanced reasoning like:

- Prioritize implementations;
- Automatic identification of common requirements based on their characterization;
- Recommendations of implementation solutions based on similar requirements;
- Requirements change management and traceability.

Main achievements

- Multi-disciplinary teams were able to collaborate in requirements elicitation using the developed solution;
- Increased context awareness of the elicited requirements;
- Visibility into existing links from requirements to implementations;
- Aggregation of requirements (e.g. according with priority and commonalities) to coordinate its implementations.

Publications

- Collaborative Management of Requirements Using Semantic Wiki Modules [339]
- A Semantic Wiki Approach to Enable Behaviour Driven Requirements Management (To be published).

12.2.4.2 Processes Modeling

This scenario aims at facilitate teams with different backgrounds to collaborate in organizations' processes modeling. For that the OSMOSE Process Manager Modelling Toolbox was developed. It supports users in the definition of the osmosis processes starting from a high-level conceptual view, provided directly by business actors, until the definition of the osmosis process model, which needs to be conducted by more technical actors such as system architects.

Main achievements

- Enabled collaboration between teams with different backgrounds in organizations' processes management;
- Enabled knowledge transference between teams while collaborating;

Publications

- Process Modelling Approach for the Liquid-Sensing Enterprise [412]
- Osmosis Process Development for Innovative Product Design and Validation [413]

12.2.5 SEMANTIK

Under SEMANTIK, the chosen scenario and domain to validate the project was the hospitality industry. In this domain, web technologies had an enormous impact on marketing and sales strategies. Following this direction, the aim of SEMANTIK is to help enterprises to become more competitive in the global market by making an efficient use of all generated data from external agents (e.g. customers and competitors). Thus, the Overall objectives are to establish a natural language processing upon hospitality industry web sources in order to allow:

- Natural language query to the system:
  - Which are features that my competitors have available to their clients?
- Opinion Mining:
  - Which are the characteristic that the clients appreciate the most;
  - Which are the aspects to improve to increase competitiveness.

Main achievements

- Increased context awareness of the domain;
- Speed up the time to answer;
- Decreased effort to coordinate tasks - customer oriented solutions.

Publications

- Framework for Customers' Sentiment Analysis [77]

12.3 Publications Summary

The research work published in conference papers or journals addresses several areas that can benefit from the proposed framework. In Table 12.2 those areas are presented and the papers related to those areas are listed.

Table 12.2: Publications by Research Area.

Research Area	Publications
Human and Enterprises Interoperability	• Semantic Adaptation of Knowledge Representation Systems • A Knowledge Management Framework to Support Online Communities Creation • Model-Driven Approach for the Interoperability of Enterprises' Services Information Exchange
Knowledge Transmission and Learning	• E-training development approach for enterprise knowledge evolution [327] • The ALTER-NATIVA knowledge management approach [72]
Knowledge Sharing and Management	• Knowledge Management support in Sensing Enterprises Establishment [414] • Knowledge Management Framework using wiki-based front-end modules [415] • Collaborative Knowledge Management using wiki front-end modules [409] • A Semantic Wiki Approach to Enable Behaviour Driven Requirements Management (To be published)
Organizations' Processes Management	• Collaborative Management of Requirements using Semantic Wiki Modules [339] • Osmotic Event Detection and Processing for the Sensing-Liquid Enterprise [416] • Osmosis Process Development for Innovative Product Design and Validation [413] • Process Modelling Approach for the Liquid-Sensing Enterprise [412]
Emotions and Sentiments assessment	• Framework for Customers' Sentiment Analysiss [77] • Framework for Management of Internet Objects in their Relation with Human Sensations and Emotions [417]

Table 12.3: Publications made in the Control and Decision Area.

Research Area	Publications
Control and Decision	• Adaptive quasi-optimal gains tuning of PI-fuzzy controllers [418] • Optimal gains tuning of PI-fuzzy controllers [419] • Gains Tuning of Fuzzy Controllers Based on a Cost Function Optimization [420] • Outliers Accommodation in Fuzzy Control Systems over WSAN [421] • Fuzzy controllers gains tuning: a constrained nonlinear optimization approach [422] • Gains Tuning of Fuzzy PID Controllers for MIMO Systems: A Performance-Driven Approach [423] • Optimal Tuning of Scaling Factors and Membership functions for Mamdani Type PID Fuzzy Controllers [424] • Fuzzy-PID Gains Tuning: Differential Evolution versus Second Order Analytical Algorithms [425]

As stated in chapter 4 Fuzzy Logic is a multivalued logic able to absorb vague information, usually described in natural language, and convert it into a numerical format for easy computational manipulation, searching for shaping or emulate the human reasoning. Thus, it made sense to explore this topic in parallel with the main PhD research. With that, the author got the opportunity to explore how human knowledge (described by natural language rules) could be directly applied to systems control and decision. The publications made awhile of this research are the detailed in Table 12.3.

12.4 Future Work

As future work, there is the further incorporation of ontology learning techniques in the proposed approach for domain glossary building. It would reduce the effort of a domain lexicon achievement contributing for systems interoperability - The author considers that a possible solution for systems interoperability is to use a common reference lexicon to be the intermediary in the communications between enterprises and to the outside. This future work could take advantage of the state of the art study presented in section 4.4.

Also as future work, a methodology that supports organizations in their knowledge management typology selection should the established. It would allow knowledge management approaches to better meet each organization profile. The author already identified some of the criteria that could be follow in section 3.1. These could be used to help future works to came up with a formal methodology.

BIBLIOGRAPHY

[1] M. Osterloh and B. S. Frey. *Motivation, Knowledge Transfer, and Organizational Form*. Tech. rep. Institute for Empirical Research in Economics - University of Zurich, IEW - Working Papers, 2000.

[2] M Polanyi. *The Tacit Dimension*. Terry lectures. Anchor Books, 1967.

[3] I. Nonaka and H. Takeuchi. *The Knowledge-Creating Company. How Japanese Companies Create the Dynamics of Innovation*. New York: Oxford University Press, 1995.

[4] S. M. Jasimuddin and Z. J. Zhang. "Knowledge management strategy and organizational culture". In: *Journal of the Operational research society* (2013), pp. –.

[5] R. S. de Alwis and E. Hartmann. "The use of tacit knowledge within innovative companies: knowledge management in innovative enterprises". In: *Journal of Knowledge Management* 12.1 (2008), pp. 133–147.

[6] J. Reason. *Human Error*. Cambridge University Press, 1990.

[7] C. Gutwin, S. Greenberg, and M. Roseman. "Workspace Awareness in Real-Time Distributed Groupware: Framework, Widgets, and Evaluation". In: *Proceedings of HCI on People and Computers XI*. HCI '96. London, UK, UK: Springer-Verlag, 1996, pp. 281–298.

[8] B. A. Pasternack and A. J. Viscio. *The centerless corporation: A new model for transforming your organization for growth and prosperity*. New York: Simon & Schuster, 1998, p. 312.

[9] J. Pfeffer and R. Sutton. *The Knowing-doing Gap: How Smart Companies Turn Knowledge Into Action*. Harvard Business School Press, 2000.

[10] M. Alavi, R. H. Smith, D. E. Leidner, M. Alavi, and D. E. Leidner. "Knowledge management systems: issues, challenges and benefits". In: *Article 7*. 1999.

[11] J. Liebowitz. *Knowledge management handbook*. CRC press, 1999.

[12] E. Craciun. *Benefits of a Knowledge Management System Within an Organization*. 2014. URL: `https://www.linkedin.com/pulse/20141029121125-2108052-benefits-of-a-knowledge-management-system-within-an-organization` (visited on 08/14/2016).

[13] A. Montoyo, P. Martínez-Barco, and A. Balahur. "Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments". In: *Decision Support Systems* 53.4 (Nov. 2012), pp. 675–679.

[14] A. Liew. "Understanding data, information, knowledge and their inter-relationships". In: *Journal of Knowledge Management Practice* 8.2 (2007).

[15] C. Lynch. "How do your data grow ?" In: 455.September (2008).

[16] A. McAfee and E. Brynjolfsson. "Big Data: The Management Revolution. (cover story)." In: *Harvard Business Review* 90.10 (2012), pp. 60–68.

[17] NewVantage Partners. *Big Data Executive Survey: Themes & Trends*. http://newvantage.com/wp-content/uploads/2012/12/NVP-Big-Data-Survey-Themes-Trends.pdf. 2012.

[18] Capgemini. *The Deciding Factor: Big Data & Decision Making*. Tech. rep. Economist Intelligence Unit commissioned by Capgemini, 2012.

[19] NewVantage Partners. *Big Data Executive Survey: Creating a Big Data Environment to Accelerate Business Value*. http://www.ciosummits.com/media/solution_spotlight/NVP _Big_Data_Survey_-_Business_Value.pdf. 2012.

[20] K. E. Martin. "Ethical issues in the Big Data industry". In: *MIS Quarterly Executive* 14 (2015), p. 2.

[21] Aite. *Reference data Acquisition Challenges: Getting it right from the start*. 2011. URL: https://www.informatica.com/downloads/1645_AITE_Informatica_Reference_Data_Acquisition.pdf (visited on 08/14/2016).

[22] J. R. Olson and H. H. Rueter. "Extracting expertise from experts: Methods for knowledge acquisition". In: *Expert systems* 4.3 (1987), pp. 152–168.

[23] L. B. Holder, Z. Markov, and I. Russell. "Advances in knowledge acquisition and representation". In: *International Journal on Artificial Intelligence Tools* 15.06 (2006), pp. 867–874.

[24] H. A. Smith and J. D. McKeen. "Instilling a knowledge-sharing culture". In: *Queen's Centre for Knowledge-Based Enterprises* 20.1 (2003), pp. 1–17.

[25] N. M. Noor and J. Salim. "Factors influencing employee knowledge sharing capabilities in electronic government agencies in Malaysia". In: *International Journal of Computer Science Issues* 8.4 (2011).

[26] M. B. Ismail and Z. M. Yusof. "Factors affecting knowledge sharing in public organizations in Malaysia". In: *Proceedings of the Knowledge Management International Conference*. Vol. 1. Citeseer. 2008, pp. 165–171.

[27] L. Yap, R. Tasmin, M. S. C. Rusuli, and N. Hashim. "Factors influencing knowledge management practices among Multimedia Super Corridor (MSC) organizations". In: *Communications of the IBIMA* (2010).

[28] H. Suresh and C. O. Egbu. "Knowledge mapping: concepts and benefits for a sustainable urban environment". In: *20th Annual Conference Association of Researchers in Construction Management (ARCOM)*. 2004, pp. 1–3.

[29] D. Asoh, S. Belardo, and R. Neilson. "Knowledge management: issues, challenges and opportunities for governments in the new economy". In: *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on*. 2002, pp. 1745–1754.

[30] K. Mathi. "Key success factors for knowledge management". In: *MBA: International Business Management& Consulting. Master thesis* (2004).

[31] I. Green-Demers, L. G. Pelletier, and S. Menard. "The impact of behavioural difficulty on the saliency of the association between self-determined motivation and environmental behaviours." In: *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* 29.3 (1997), p. 157.

[32] C.-P. Wu. "The Moderation Effect of Motivation on the Relationship between Internet Search and Critical Thinking". In: (2008).

[33] L. Camarinha-Matos. "Scientific Research Methodologies and Techniques - Unit 2: Scientific Method". In: *Phd Program in Electrical and Computer Engineering* (2010).

[34] K. Ichijo and I. Nonaka. *Knowledge Creation and Management: New Challenges for Managers*. Oxford University Press, 2006.

[35] I. Nonaka. "A Dynamic Theory of Organizational Knowledge Creation". In: *Organization Science* 5.1 (1994), pp. 14–37.

[36] H. Takeuchi and T. Shibata. "The New Dynamism of the Knowledge-Creating Company." In: *Japan Moving Toward a More Advanced Knowledge Economy: Advanced Knowledge—Creating Companies*. Washington, D.C.: World Bank Institute (WBI), 2006.

[37] S. Ceptureanu, E. Ceptureanu, et al. "Knowledge creation/conversion process". In: *Review of International Comparative Management* 1.1 (2010).

[38] M. Alavi and D. E. Leidner. "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues". In: *MIS Q.* 25.1 (Mar. 2001), pp. 107–136.

[39] F. Alipour, K. Idris, and R. Karimi. "Knowledge creation and transfer: Role of learning organization". In: *International Journal of Business Administration* 2.3 (2011), p. 61.

[40] H. Takeuchi and I. Nonaka. *Hitotsubashi on Knowledge Management*. Wiley, 2004.

[41] H. K. Eliufoo. "Knowledge creation and transfer in construction organizations in Tanzania: Building and Real State of Economics". In: (2005).

[42] G. Yukl. *Leadership in Organizations*. Prentice Hall, 2010.

[43] T. M. Amabile. "The social psychology of creativity: A componential conceptualization." In: *Journal of personality and social psychology* 45.2 (1983), p. 357.

[44] F. Barron and D. M. Harrington. "Creativity, intelligence, and personality". In: *Annual review of psychology* 32.1 (1981), pp. 439–476.

[45] T. M. Amabile. "A model of creativity and innovation in organizations". In: *Research in organizational behavior* 10.1 (1988), pp. 123–167.

[46] J. P. De Jong and D. N. Den Hartog. "How leaders influence employees' innovative behaviour". In: *European Journal of innovation management* 10.1 (2007), pp. 41–64.

[47] D. O. McDaniel, D. H. Brown, and C. M. Woh. *Section 1- Manual for Media Trainers: A Learner-centred Approach*. Asia-Pacific Institute for Broadcasting Development (AIBD); United Nations Educational, Scientific and Cultural Organisation (UNESCO), 2001. Chap. section 1.

[48] F Bouthillier and K Shearer. "Understanding knowledge management and information management: the need for an empirical perspective". In: *Information Research: an international electronic journal* 8.1 (2002).

[49] T. M. Cao and P. Compton. "A Simulation Framework for Knowledge Acquisition Evaluation". In: *Proceedings of the Twenty-eighth Australasian Conference on Computer Science - Volume 38*. ACSC '05. Newcastle, Australia: Australian Computer Society, Inc., 2005, pp. 353–360.

[50] A. H. Mohammad and N. A. M. AL SAIYD. "GUIDELINES FOR TACIT KNOWLEDGE ACQUISITION". In: *Journal of Theoretical and Applied Information Technology* 38.1 (2005), pp. 110–118.

[51] N. A. Stanton and R. Stammers. “Learning styles in a non-linear training environment”. In: *Hypertext: State of the art* (1990), pp. 114–120.

[52] M. Stollberg, A. V. Zhdanova, and D. Fensel. “h-TechSight — A Next Generation Knowledge Management Platform”. In: *Journal of Information & Knowledge Management* 03.01 (2004), pp. 45–66.

[53] “Information Acquisition and Mental Models: An Investigation into the Relationship Between Behaviour and Learning”. In: *Info. Sys. Research* 7.2 (June 1996), pp. 198–214.

[54] M. Khalifa and K. N. Shen. “Effects of knowledge representation on knowledge acquisition and problem solving”. In: *Proceedings of the 2nd International Conference on Intellectual Capital, Knowledge Management and Organisational Learning (ICICKM 2005)*. Academic Conferences Limited. 2005, p. 277.

[55] K. Kotiadis and S. Robinson. “Conceptual modelling: Knowledge acquisition and model abstraction”. In: *Simulation Conference, 2008. WSC 2008. Winter*. 2008, pp. 951–958.

[56] C. Marques-Lucena, C. Agostinho, S. Koussouris, and J. Sarraipa. “Knowledge Management Framework using Wiki-based Front-end Modules”. In: *ICEIS 2015 - Proceedings of the 17th International Conference on Enterprise Information Systems, Volume 1, Barcelona, Spain, 27-30 April, 2015*. 2015, pp. 79–86.

[57] K. Dalkir. *Knowledge Management in Theory and Practice*. The MIT Press, 2011.

[58] G. Von Krogh, J. Roos, and D. Kleine. *Knowing in firms: understanding, managing and measuring knowledge*. Sage, 1998.

[59] B Faust. “Implementation of tacit knowledge preservation and transfer methods”. In: *International Conference on Knowledge Management in Nuclear Facilities*. 2007.

[60] BusinessDictionary. *Brainstorming @ONLINE*. Aug. 2015. URL: http://www.businessdictionary.com/definition/brainstorming.html.

[61] B. Brešić. “Knowledge Acquisition in Databases”. In: *Management Information Systems* 7.1 (2012), pp. 032–041.

[62] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. “From data mining to knowledge discovery in databases”. In: *AI magazine* 17.3 (1996), p. 37.

[63] U. M. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. “Advances in Knowledge Discovery and Data Mining”. In: ed. by U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. Menlo Park, CA, USA: American Association for Artificial Intelligence, 1996. Chap. From Data Mining to Knowledge Discovery: An Overview, pp. 1–34.

[64] M. Dawande, J. Kalagnanam, H. S. Lee, C. Reddy, S. Siegel, and M. Trumbo. “The Slab-Design Problem in the Steel Industry”. In: *Interfaces* 34.3 (2004), pp. 215–225.

[65] Z. Pawlak. “Rough sets and intelligent data analysis”. In: *Information Sciences* 147.1-4 (Nov. 2002), pp. 1–12.

[66] S. Auer, S. Dietzold, and T. Riechert. “OntoWiki – A Tool for Social, Semantic Collaboration”. In: *The Semantic Web - ISWC 2006*. Ed. by I. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. Aroyo. Vol. 4273. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2006, pp. 736–749.

[67] R. Davis, H. E. Shrobe, and P. Szolovits. “What Is a Knowledge Representation?” In: *AI Magazine* 14.1 (1993), pp. 17–33.

[68] P. Pantel and M. Pennacchiotti. “Automatically Harvesting and Ontologizing Semantic Relations”. In: *Proceedings of the 2008 Conference on Ontology Learning and Population: Bridging the Gap Between Text and Knowledge*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2008, pp. 171–195.

[69] S. Lervad. *The European Association for Terminology (EAFT): Background, Objectives and Perspectives*. Tech. rep. Snow Landsc. Rec. 74, 2: 263-267, 1999.

[70] E. Riloff and J. Shepherd. “A Corpus-Based Approach for Building Semantic Lexicons”. In: *In Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*. 1997, pp. 117–124.

[71] D. Maynard, Y. Li, and W. Peters. “NLP Techniques for Term Extraction and Ontology Population”. In: *Proceedings of the 2008 Conference on Ontology Learning and Population: Bridging the Gap Between Text and Knowledge*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2008, pp. 107–127.

[72] J. Sarraipa, C. Marques-Lucena, S. Baldiris, R. Fabregat, and S. Aciar. “The ALTER-NATIVA knowledge management approach”. English. In: *Journal of Intelligent Manufacturing* (2014), pp. 1–17.

[73] A Muresan and J. Klavans. “A method for automatically building and evaluating dictionary resources”. In: *Proceedings of the Language Resources and Evaluation Conference (LREC*. 2002.

[74] E. Westerhout and P. Monachesi. “Extraction of Dutch definitory contexts for eLearning purposes”. In: *Proceedings of CLIN* (2007), pp. 1–15.

[75] I. Fahmi and G. Bouma. “Learning to identify definitions using syntactic features”. In: *Proceedings of the EACL workshop on Learning Structured Information in Natural Language Applications*. 2006.

[76] P. Velardi, R. Poler, and J. V. Tomás. “Methodology for the Definition of a Glossary in a Collaborative Research Project and its Application to a European Network of Excellence”. In: *Interoperability of Enterprise Software and Applications*. Ed. by D. Konstantas, J.-P. Bourrières, M. Léonard, and N. Boudjlida. London: Springer London, 2006, pp. 311–322.

[77] C. Marques-Lucena, J. Sarraipa, J. Fonseca, A. Grilo, and R. Jardim-Gonçalves. “Framework for Customers’ Sentiment Analysis”. English. In: *Intelligent Systems 2014*. Ed. by P. Angelov, K. Atanassov, L. Doukovska, M. Hadjiski, V. Jotsov, J. Kacprzyk, N. Kasabov, S. Sotirov, E. Szmidt, and S. Zadrożny. Vol. 322. Advances in Intelligent Systems and Computing. Springer International Publishing, 2015, pp. 849–860.

[78] T. R. Gruber. “A translation approach to portable ontology specifications”. In: *Knowl. Acquis.* 5.2 (1993), pp. 199–220.

[79] Y. J. An, K.-c. Huang, and J. Geller. “Naturalness of Ontology Concepts for Rating Aspects of the Semantic Web”. In: (2006).

[80] J Sarraipa and R Jardim-Gonçalves. “Semantics Adaptability for Systems Interoperability”. In: *Control Engineering and Applied Informatics*. 2010.

[81] H. A. Simon. “The New Science of Management Decisions”. In: *The Shape of Automation for Men and Management*. Harper and Row, 1960.

[82] R. Djedidi and M.-A. Aufaure. “Ontological Knowledge Maintenance Methodology”. In: *Proceedings of the 12th international conference on Knowledge-Based Intelligent Information and Engineering Systems, Part I*. KES ’08. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 557–564.

[83] M. Greco, M. Grimaldi, and M. Hanandi. “How to select knowledge management systems: A framework to support managers regular paper”. In: *International Journal of Engineering Business Management* 5 (2013), pp. 1–11.

[84] T. H. Davenport, L. Prusak, and L. Prusak. *Working Knowledge: How Organizations Manage What They Know*. Boston, MA, USA: Harvard Business School Press, 1997.

[85] E. Tsui. “Tracking the Role and Evolution of Commercial Knowledge Management Software”. English. In: *Handbook on Knowledge Management*. Ed. by C. Holsapple. Vol. 2. International Handbooks on Information Systems. Springer Berlin Heidelberg, 2003, pp. 5–27.

[86] K. Y. Wong and E. Aspinwall. “An empirical study of the important factors for knowledge-management adoption in the SME sector”. In: *Journal of Knowledge Management* 9.3 (2005), pp. 64–82.

[87] E. Ngai and E. Chan. “Evaluation of knowledge management tools using AHP”. In: *Expert Systems with Applications* 29.4 (2005), pp. 889 –899.

[88] A. Silvestri, F. D. Felice, and A. Petrillo. “Multi-criteria risk analysis to improve safety in manufacturing systems”. In: *International Journal of Production Research* 50.17 (2012), pp. 4806–4821.

[89] S. Sipahi and M. Timor. “The analytic hierarchy process and analytic network process: an overview of applications”. In: *Management Decision* 48.5 (2010), pp. 775–808.

[90] T. Saaty. *The Analytic Hierarchy Process, Planning, Piority Setting, Resource Allocation*. New york: McGraw-Hill, 1980.

[91] D.-H. Byun. “The {AHP} approach for selecting an automobile purchase model”. In: *Information & Management* 38.5 (2001), pp. 289 –297.

[92] M. C. Tam and V. Tummala. “An application of the {AHP} in vendor selection of a telecommunications system”. In: *Omega* 29.2 (2001), pp. 171 –182.

[93] K. Muralidhar, R. Santhanam, and R. L. Wilson. “Using the analytic hierarchy process for information system project selection”. In: *Information & Management* 18.2 (1990), pp. 87–95.

[94] F. Mamaghani. “Evaluation and selection of an antivirus and content filtering software”. In: *Information Management & Computer Security* 10.1 (2002), pp. 28–32.

[95] P. Reyes and M. S. Raisinghani. “Integrating information technologies and knowledge-based systems: a theoretical approach in action for enhancements in production and inventory control”. In: *Knowledge and Process Management* 9.4 (2002), pp. 256–263.

[96] L. Davis and G. Williams. “Evaluating and Selecting Simulation Software Using the Analytic Hierarchy Process”. In: *Integrated Manufacturing Systems* 5.1 (1994), pp. 23–32.

[97] V. S. Lai, R. P. Trueblood, and B. K. Wong. “Software selection: a case study of the application of the analytical hierarchical process to the selection of a multimedia authoring system”. In: *Information & Management* 36.4 (1999), pp. 221 –232.

[98] T. Srikantaiah, T. Srikantaiah, M. Koenig, and A. S. for Information Science. *Knowledge Management for the Information Professional*. ASIS monograph series. American Society for Information Science, 2000, pp. 269–276.

[99] S. Conway and C. Sligar. *Unlocking Knowledge Assets*. Redmond, WA, USA: Microsoft Press, 2002.

[100] A. Tiwana and B. Ramesh. “Integrating Knowledge on the Web”. In: *IEEE Internet Computing* 5.3 (May 2001), pp. 32–39.

[101] W. Ossadnik and O. Lange. “AHP-based evaluation of AHP-Software”. In: *European Journal of Operational Research* 118.3 (1999), pp. 578 –588.

[102] D.-H. Byun and E.-H. Suh. “A Methodology for Evaluating EIS Software Packages”. en. In: *Journal of Organizational and End User Computing (JOEUC)* 8.2 (Apr. 1996), pp. 21–31.

[103] Y.-H. Yao, A. J. Trappey, and P.-S. Ho. “XML-based {ISO9000} electronic document management system”. In: *Robotics and Computer-Integrated Manufacturing* 19.4 (2003), pp. 355 –370.

[104] H. Zantout and F. Marir. “Document management systems from current capabilities towards intelligent information retrieval: an overview”. In: *International Journal of Information Management* 19.6 (1999), pp. 471 –484.

[105] *ISO 9000:2005 Quality management systems – Fundamentals and vocabulary*. Norm. 2009.

[106] *ISO/TR 10013:2001 Guidelines for quality management system documentation*. Norm. 2001.

[107] Y. Yao and A. J. C. Trappey. “Using ISO 10303 Data Standard and XML Standard Web Technology to Enable ISO 9000 Document Management”. In: *Int. J. Comput. Appl. Technol.* 18.1-4 (July 2003), pp. 117–132.

[108] P. Willett, ed. *Document Retrieval Systems*. London, UK, UK: Taylor Graham Publishing, 1988.

[109] F. W. Lancaster and A. J. Warner. *Information Retrieval Today*. 1st. Arlington, VA, USA: Information Resources Press, 1993.

[110] G.-J. Hwang, C.-H. Wu, J. C. R. Tseng, and I. Huang. “Development of a ubiquitous learning platform based on a real-time help-seeking mechanism”. In: *British Journal of Educational Technology* 42.6 (2011), pp. 992–1002.

[111] P. C. ROGERS, S. W. LIDDLE, P. CHAN, A. DOXEY, and B. ISOM. *Web 2.0 Learning Platform: Harnessing Collective Intelligence*. en. 2007.

[112] H. Ogata and Y. Yano. “Context-aware support for computer-supported ubiquitous learning”. In: *Wireless and Mobile Technologies in Education, 2004. Proceedings. The 2nd IEEE International Workshop on*. 2004, pp. 27–34.

[113] K. Yatani, M. Sugimoto, and F. Kusunoki. “Musex: a system for supporting children’s collaborative learning in a museum with PDAs”. In: *Wireless and Mobile Technologies in Education, 2004. Proceedings. The 2nd IEEE International Workshop on*. 2004, pp. 109–113.

[114] M. El-Bishouty, H. Ogata, and Y. Yano. “A Model of Personalized Collaborative Computer Support Ubiquitous Learning Environment”. In: *Advanced Learning Technologies, 2008. ICALT ’08. Eighth IEEE International Conference on*. 2008, pp. 97–101.

[115] S. J. H. Yang. “Context Aware Ubiquitous Learning Environments for Peer-to-Peer Collaborative Learning”. In: *Educational Technology & Society* 9.1 (2006), pp. 188–201.

[116] D. P. Lepak and S. A. Snell. “Virtual HR: Strategic human resource management in the 21st century”. In: *Human Resource Management Review* 8.3 (1998). Special Issue: Strategic Human Resource Management Research, pp. 215 –234.

[117] C. Ichniowski and K. Shaw. “Beyond Incentive Pay: Insiders’ Estimates of the Value of Complementary Human Resource Management Practices”. In: *Journal of Economic Perspectives* 17.1 (2003), pp. 155–180.

[118] K Harris, M Fleming, R Hunter, B Rosser, and A Cushman. “The knowledge management scenario: trends and directions for 1998-2003”. In: *Gartner Group* (1999).

[119] “Corporate portal: a tool for knowledge management synchronization”. In: *International Journal of Information Management* 24.3 (2004), pp. 201 –220.

[120] M. White. “Enterprise information portals”. In: *The Electronic Library* 18.5 (2000), pp. 354–362.

[121] Y. J. Kim, A. Chaudhury, and H. R. Rao. “A knowledge management perspective to evaluation of enterprise information portals”. In: *Knowledge and Process Management* 9.2 (2002), pp. 57–71.

[122] J. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson. “Past, present, and future of decision support technology”. In: *Decision Support Systems* 33.2 (2002). Decision Support System: Directions for the Nest Decade, pp. 111 –126.

[123] H. K. Bhargava, D. J. Power, and D. Sun. “Progress in Web-based decision support technologies”. In: *Decision Support Systems* 43.4 (2007). Special Issue Clusters, pp. 1083 –1095.

[124] R Ling and D. Yen. “Customer relationship management: An analysis framework and implementation strategies”. In: *The Journal of Computer Information Systems* (2001).

[125] R. Swift. *Accelerating Customer Relationships: Using Crm and Relationship Technologies™* First. Upper Saddle River, NJ, USA: Prentice Hall Press, 2000.

[126] A. Parvatiyar and J. N. Sheth. “Customer Relationship Management: emerging practice, process and discipline”. In: *Journal of Economic and Social Research* 3 (2002), pp. 6–23.

[127] A. Kracklauer, D. Mills, and D. Seifert. “Customer Management as the Origin of Collaborative Customer Relationship Management”. English. In: *Collaborative Customer Relationship Management*. Ed. by A. Kracklauer, D. Mills, and D. Seifert. Springer Berlin Heidelberg, 2004, pp. 3–6.

[128] “Cross-functional issues in the implementation of relationship marketing through customer relationship management”. In: *European Management Journal* 19.5 (2001), pp. 534 –542.

[129] D. J. Thomas and P. M. Griffin. “Coordinated supply chain management”. In: *European Journal of Operational Research* 94.1 (1996), pp. 1 –15.

[130] D. N. Chorafas. *Integrating Erp, Supply Chain Management, and Smart Materials*. Boca Raton, FL, USA: CRC Press, Inc., 2001.

[131] M. Schrage. *Shared Minds: The New Technologies of Collaboration*. New York, NY, USA: Random House Inc., 1990.

[132] E. Churchill and D. Snowdon. “Collaborative virtual environments: An introductory review of issues and systems”. English. In: *Virtual Reality* 3.1 (1998), pp. 3–15.

[133] V. Odumuyiwa, E. Site-loria, N. Université, C. Scientifique, and B. P. Vandœuvre. “Collaborative Watch : towards an effective intelligence gathering”. In: (2011), pp. 5–7.

[134] D. D'Amour, L. Goulet, J.-F. Labadie, L. San Martín-Rodriguez, and R. Pineault. "A model and typology of collaboration between professionals in healthcare organizations". In: *BMC health services research* 8.1 (2008), p. 1.

[135] S. L. Elrod and M. M. Somerville. "Literature-Based Scientific Learning: A Collaboration Model". In: *The Journal of Academic Librarianship* 33.6 (2007), pp. 684 –691.

[136] L. S. Thompson, M. Story, and G. Butler. "Use of a university-community collaboration model to frame issues and set an agenda for strengthening a community". In: *Health Promotion Practice* 4.4 (2003), pp. 385–392.

[137] H. Patel, M. Pettitt, and J. R. Wilson. "Factors of collaborative working: A framework for a collaboration model". In: *Applied Ergonomics* 43.1 (2012), pp. 1 –26.

[138] G. Bafoutsou and G. Mentzas. "Review and functional classification of collaborative systems". In: *International Journal of Information Management* 22.4 (2002), pp. 281 –305.

[139] M. J. Pazzani. "A Framework for Collaborative, Content-Based and Demographic Filtering". In: *Artif. Intell. Rev.* 13.5-6 (Dec. 1999), pp. 393–408.

[140] W. Tolone, G.-J. Ahn, T. Pai, and S.-P. Hong. "Access Control in Collaborative Systems". In: *ACM Comput. Surv.* 37.1 (Mar. 2005), pp. 29–41.

[141] D. Coleman and L. Ward. "Taking Advantage of Real-Time Collaboration Tools". In: *IT Professional* 1.4 (1999), pp. 25–30.

[142] S. E. Poltrock and G. Engelbeck. "Requirements for a virtual collocation environment". In: *Information and Software Technology* 41.6 (1999), pp. 331 –339.

[143] R. BENTLEY, W. APPELT, U. BUSBACH, E. HINRICHS, D. KERR, K. SIKKEL, J. TREVOR, and G. WOETZEL. "Basic Support for Cooperative Work on the World Wide Web". In: *Int. J. Hum.-Comput. Stud.* 46.6 (June 1997), pp. 827–846.

[144] T. L. Friedman. *The World is flat : a brief history of the twenty-first century*. New York: Farrar, Straus and Giroux, 2006.

[145] A. Gunasekaran and E. Ngai. "Build-to-order supply chain management: a literature review and framework for development". In: *Journal of Operations Management* 23.5 (2005). The Build to Order Supply Chain (BOSC)A Competitive Strategy for the 21st Century, pp. 423 –451.

[146] B. Pine and J. Gilmore. *The Experience Economy: Work is Theatre & Every Business a Stage*. Harvard Business School Press. Harvard Business School Press, 1999.

[147] R. Jardim-Goncalves, C. Agostinho, P. Malo, and A. Steiger-Garcao. "Harmonising technologies in conceptual models representation". In: *International Journal of Product Lifecycle Management* 2.2 (2007), pp. 187–205.

[148] S. Ray and A. Jones. "Manufacturing interoperability". English. In: *Journal of Intelligent Manufacturing* 17.6 (2006), pp. 681–688.

[149] IEEE. *IEEE Standard Glossary of Software Engineering Terminology*. Tech. rep. 1990, pp. 1+.

[150] V. Peristeras and K. A. Tarabanis. "The Connection, Communication, Consolidation, Collaboration Interoperability Framework (C4IF) For Information Systems Interoperability." In: *IBIS* 1.1 (2006), pp. 61–72.

[151] "IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries". In: *IEEE Std 610* (1991), pp. 1–217.

[152] A. Ojo, T. Janowski, and E. Estevez. "Semantic Interoperability Architecture for Electronic Government". In: *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections Between Citizens, Data and Government*. dg.o '09. Puebla, Mexico: Digital Government Society of North America, 2009, pp. 63–72.

[153] M. Gadu and N. El-Khameesy. "A Knowledge Management Framework Using Business Intelligence Solutions". In: *International Journal of Computer Science Issues (IJCSI)* 11.5 (2014), p. 102.

[154] C. Cohen. *Business Intelligence: The Effectiveness of Strategic Intelligence and Its Impact on the Performance of Organizations*. John Wiley & Sons, 2013.

[155] R. Shehzad, M. Khan, and M Naeem. "Integrating knowledge management with business intelligence processes for enhanced organizational learning". In: *International Journal of Software Engineering and Its Applications* 7.2 (2013), pp. 83–91.

[156] G. K. Rao and R. Kumar. "Framework to integrate business intelligence and knowledge management in banking industry". In: *arXiv preprint arXiv:1109.0614* (2011).

[157] G. Radonić and K. Ćurko. "A review of business intelligence approaches to key business factors in banking". In: *Journal of Knowledge Management Practice* 8.1 (2007), pp. 1–10.

[158] R. A. Khan and S. Quadri. "Business intelligence: an integrated approach". In: *Business Intelligence Journal* 5.1 (2012), pp. 64–70.

[159] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.

[160] R. A. Khan and S. Quadri. "Dovetailing of business intelligence and knowledge management: An integrative framework". In: *Information and Knowledge Management*. Vol. 2. 4. 2012, p. 2.

[161] P. Vassiliadis, A. Karagiannis, V. Tziovara, A. Simitsis, and I. Hellas. "Towards a benchmark for etl workflows". In: (2007).

[162] R. Sabherwal and I. Becerra-Fernandez. *Business Intelligence: Practices, Technologies, and Management*. Wiley, 2011.

[163] R. T. Herschel and N. E. Jones. "Knowledge management and business intelligence: the importance of integration". In: *Journal of knowledge management* 9.4 (2005), pp. 45–55.

[164] Z. Weidong, D. Weihui, and Y. Kunlong. "The relationship of business intelligence and knowledge management". In: *Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on*. IEEE. 2010, pp. 26–29.

[165] E. Ngai, L. Xiu, and D. Chau. "Application of data mining techniques in customer relationship management: A literature review and classification". In: *Expert Systems with Applications* 36.2, Part 2 (2009), pp. 2592 –2602.

[166] J. Hipp, U. Güntzer, and G. Nakhaeizadeh. "Algorithms for Association Rule Mining &Mdash; a General Survey and Comparison". In: *SIGKDD Explor. Newsl.* 2.1 (June 2000), pp. 58–64.

[167] P. Tan and V Kumar. "Chapter 6. Association Analysis: Basic Concepts and Algorithms". In: *Introduction to Data Mining. Addison-Wesley. ISBN* (2005).

[168] F. Cavaco. *Ontologies learn by searching*. run.unl.pt, 2011.

[169] C. C. Aggarwal. *Data Classification: Algorithms and Applications*. 1st. Chapman & Hall/CRC, 2014.

[170] A. Abraham and C. Grosan. "Engineering Evolutionary Intelligent Systems: Methodologies, Architectures and Reviews". English. In: *Engineering Evolutionary Intelligent Systems*. Ed. by A. Abraham, C. Grosan, and W. Pedrycz. Vol. 82. Studies in Computational Intelligence. Springer Berlin Heidelberg, 2008, pp. 1–22.

[171] Y Tang, K Sycara, M Sensoy, and J. Pan. "Survey on clustering methods for ontological knowledge". In: *ITACS International Technology Aliance Collaboration Systems* (2012).

[172] G. Paliouras, C. Papatheodorou, V. Karkaletsis, and C. D. Spyropoulos. "Clustering the Users of Large Web Sites into Communities". In: *Proceedings of the Seventeenth International Conference on Machine Learning*. ICML '00. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 719–726.

[173] T. Breaux and J. Reed. "Using Ontology in Hierarchical Information Clustering". In: *System Sciences, 2005. HICSS '05. Proceedings of the 38th Annual Hawaii International Conference on*. 2005, 111b–111b.

[174] D. Godoy and A. Amandi. "Modeling user interests by conceptual clustering". In: *Information Systems* 31.4–5 (2006). The Semantic Web and Web Services, pp. 247 –265.

[175] J. Völker, P. Haase, and P. Hitzler. "Learning Expressive Ontologies". In: *Proceedings of the 2008 Conference on Ontology Learning and Population: Bridging the Gap Between Text and Knowledge*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2008, pp. 45–69.

[176] T Rey, A Kordon, and C Wells. *Applied data mining for forecasting using SAS*. books.google.com, 2012.

[177] J. John B. Guerard. *Introduction to Financial Forecasting in Investment Analysis [electronic resource]/by John B. Guerard, Jr.* Springer New York, 2013.

[178] A. Berson, S. Smith, and K. Thearling. *Building Data Mining Applications for CRM*. 1st. McGraw-Hill Professional, 1999.

[179] S. Mitra, S. Pal, and P. Mitra. "Data mining in soft computing framework: a survey". In: *Neural Networks, IEEE Transactions on* 13.1 (2002), pp. 3–14.

[180] M. J. Shaw, C. Subramaniam, G. W. Tan, and M. E. Welge. "Knowledge management and data mining for marketing". In: *Decision Support Systems* 31.1 (2001). Knowledge Management Support of Decision Making, pp. 127 –137.

[181] E. Turban, R. Sharda, and D. Delen. *Decision Support and Business Intelligence Systems*. 9th. Upper Saddle River, NJ, USA: Prentice Hall Press, 2010.

[182] E. Rolls. "Memory systems in the brain". In: *Annual review of psychology* (2000).

[183] K. Patterson, P. J. Nestor, and T. T. Rogers. "Where do you know what you know? The representation of semantic knowledge in the human brain". In: *Nature Reviews Neuroscience* 8.12 (Dec. 2007), pp. 976–987.

[184] C. Lucena, J. a. Sarraipa, and R. Jardim-Gonçalves. "Semantic Adaptation of Knowledge Representation Systems". In: *Technological Innovation for the Internet of Things*. Ed. by L. Camarinha-Matos, S. Tomic, and P. Graça. Vol. 394. IFIP Advances in Information and Communication Technology. 2013, pp. 88–95.

[185] A. M. Meystel and J. S. Albus. *Intelligent Systems: Architecture, Design, and Control*. 1st. New York, NY, USA: John Wiley & Sons, Inc., 2000.

[186] L. A. Zadeh. "Fuzzy logic, neural networks, and soft computing". In: *Commun. ACM* 37.3 (Mar. 1994), pp. 77–84.

[187] N Kasabov and D Filev. "Evolving intelligent systems: methods, learning, & applications". In: *Evolving Fuzzy Systems* (2006).

[188] J. Carbonell, R. Michalski, and T. Mitchell. "An Overview of Machine Learning". English. In: *Machine Learning*. Ed. by R. Michalski, J. Carbonell, and T. Mitchell. Symbolic Computation. Springer Berlin Heidelberg, 1983, pp. 3–23.

[189] H. A. Simon. "Search and Reasoning in Problem Solving". In: *Artif. Intell.* 21.1-2 (Mar. 1983), pp. 7–29.

[190] R. Schapire. *Theoretical Machine Learning. In Princeton courses @ONLINE*. 2008. URL: http://www.cs.princeton.edu/courses/archive/spr08/cos511/scribe_notes/0204.pdf.

[191] I. Guyon. *An Introduction to Machine Learning @ONLINE*. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.217.8944&rep=rep1&type=pdf.

[192] Z. Xi-Hu and L. Yan-Fei. "Building Ontology Automatically Based on Bayesian Network and PART Neural Network". In: *Proceedings of the 2009 WRI Global Congress on Intelligent Systems - Volume 04*. GCIS '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 563–566.

[193] L. Fu. "Rule generation from neural networks". In: *Systems, Man and Cybernetics* (1994).

[194] G. G. Towell and J. W. Shavlik. "The Extraction of Refined Rules from Knowledge-Based Neural Networks". In: *Machine Learning*. 1993, pp. 71–101.

[195] G. Nagypál and B. Motik. "A fuzzy model for representing uncertain, subjective and vague temporal knowledge in ontologies". In: *In Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics, (ODBASE), volume 2888 of LNCS*. Springer, 2003, pp. 906–923.

[196] B Zadjabbari, S Mohseni, and P Wongthongtham. "Fuzzy logic based model to measure knowledge sharing". In: *3rd IEEE International Conference on Digital Ecosystems and Technologies* (2009).

[197] P. Cimiano. *Ontology Learning and Population from Text: Algorithms, Evaluation and Applications*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[198] T. M. Mitchell. *Machine Learning*. 1st ed. New York, NY, USA: McGraw-Hill, Inc., 1997.

[199] I. Johnson, J. Abécassis, B. Charnomordic, S. Destercke, and R. Thomopoulos. "Making Ontology-Based Knowledge and Decision Trees Interact: An Approach to Enrich Knowledge and Increase Expert Confidence in Data-Driven Models". English. In: *Knowledge Science, Engineering and Management*. Ed. by Y. Bi and M.-A. Williams. Vol. 6291. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2010, pp. 304–316.

[200] J. Zhang, A. Silvescu, and V. Honavar. "Ontology-Driven Induction of Decision Trees at Multiple Levels of Abstraction". English. In: *Abstraction, Reformulation, and Approximation*. Ed. by S. Koenig and R. Holte. Vol. 2371. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2002, pp. 316–323.

[201] N. Fanizzi, C. d'Amato, and F. Esposito. "Towards the Induction of Terminological Decision Trees". In: *Proceedings of the 2010 ACM Symposium on Applied Computing*. SAC '10. Sierre, Switzerland: ACM, 2010, pp. 1423–1427.

[202] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. 3rd. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.

[203] R. Pan, Z. Ding, Y. Yu, and Y. Peng. "A Bayesian Network Approach to Ontology Mapping". English. In: *The Semantic Web – ISWC 2005*. Ed. by Y. Gil, E. Motta, V. Benjamins, and M. Musen. Vol. 3729. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, pp. 563–577.

[204] G. S. kumar and G. Zayaraz. "Concept relation extraction using Naïve Bayes classifier for ontology-based question answering systems". In: *Journal of King Saud University - Computer and Information Sciences* 27.1 (2015), pp. 13 –24.

[205] F. Colace and M. De Santo. "Ontology for E-Learning: A Bayesian Approach". In: *Education, IEEE Transactions on* 53.2 (2010), pp. 223–233.

[206] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification (2Nd Edition)*. Wiley-Interscience, 2000.

[207] O Sutton. "Introduction to k Nearest Neighbour Classification and Condensed Nearest Neighbour Data Reduction". In: *University lectures, University of Leicester* (2012).

[208] A. Haratian Nezjadi, B. Shadgara, and A. Osareh. "Ontology Alignment using Machine Learning Techniques". In: *International Journal of Computer Science and Information Technology* 3.2 (2011).

[209] N. Fanizzi, C. d'Amato, and F. Esposito. "Semantic Nearest Neighbor Search in OWL Ontologies." In: *SWAP*. Ed. by G. Semeraro, E. D. Sciascio, C. Morbidoni, and H. Stoermer. Vol. 314. CEUR Workshop Proceedings. CEUR-WS.org, June 9, 2008.

[210] G. P. Basharin, A. N. Langville, and V. A. Naumov. "The life and work of A.A. Markov". In: *Linear Algebra and its Applications* 386 (2004). Special Issue on the Conference on the Numerical Solution of Markov Chains 2003, pp. 3 –26.

[211] Z. Kastrati and A. Imran. "Adaptive Concept Vector Space Representation Using Markov Chain Model". English. In: *Knowledge Engineering and Knowledge Management*. Ed. by K. Janowicz, S. Schlobach, P. Lambrix, and E. Hyvönen. Vol. 8876. Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 203–208.

[212] S. Albagli, R. Ben-Eliyahu-Zohary, and S. E. Shimony. "Markov network based ontology matching". In: *Journal of Computer and System Sciences* 78.1 (2012). {JCSS} Knowledge Representation and Reasoning, pp. 105 –118.

[213] M. Mitchell. *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1998.

[214] M. Paulinas and A. Ušinskas. "A survey of genetic algorithms applications for image enhancement and segmentation". In: *Information Technology and control* 36.3 (2015).

[215] C. Sharma, S. Sabharwal, and R. Sibal. “A survey on software testing techniques using genetic algorithm”. In: *arXiv preprint arXiv:1411.1154* (2014).

[216] K. Man, K. Tang, and S. Kwong. “Genetic algorithms: concepts and applications [in engineering design]”. In: *Industrial Electronics, IEEE Transactions on* 43.5 (1996), pp. 519–534.

[217] D.-z. Yang, C.-y. Liang, G.-y. Si, and W.-j. Cheng. “An Intelligent Knowledge Discovery Method Based on Genetic Algorithms and Conditional Probability”. English. In: *Network Computing and Information Security*. Ed. by J. Lei, F. Wang, M. Li, and Y. Luo. Vol. 345. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2012, pp. 281–286.

[218] J. Du and R. Rada. “Knowledge-guided genetic algorithm for financial forecasting”. In: *Computational Intelligence for Financial Engineering Economics (CIFEr), 2012 IEEE Conference on*. 2012, pp. 1–8.

[219] L. Baron, S. Achiche, and M. Balazinski. “Fuzzy decision support system knowledge base generation using a genetic algorithm”. In: *International Journal of Approximate Reasoning* 28.2–3 (2001), pp. 125 –148.

[220] M. R. Girardi and B Ibrahim. “Using English to Retrieve Software”. In: *The Journal of Systems and Software* 30 (1995), pp. 249–270.

[221] S. E. Middleton, N. R. Shadbolt, and D. C. De Roure. “Ontological User Profiling in Recommender Systems”. In: *ACM Trans. Inf. Syst.* 22.1 (2004), pp. 54–88.

[222] D. Vallet, M. Fernández, and P. Castells. “An Ontology-Based Information Retrieval Model”. In: *The Semantic Web: Research and Applications*. Ed. by A. Gómez-Pérez and J. Euzenat. Vol. 3532. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, pp. 455–470.

[223] L. Drumond and R. Girardi. “A Survey of Ontology Learning Procedures.” In: *WONTO*. Ed. by F. L. G. de Freitas, H. Stuckenschmidt, H. S. Pinto, A. Malucelli, and O. Corcho. Vol. 427. CEUR Workshop Proceedings. CEUR-WS.org, 2008.

[224] A. Maedche and S. Staab. “Ontology Learning for the Semantic Web”. In: *IEEE Intelligent Systems* 16.2 (2001), pp. 72–79.

[225] W. Wong, W. Liu, and M. Bennamoun. “Ontology Learning from Text: A Look Back and into the Future”. In: *ACM Comput. Surv.* 44.4 (2012), 20:1–20:36.

[226] C. Biemann. “Ontology Learning from Text: A Survey of Methods.” In: *LDV forum*. Vol. 20. 2. 2005, pp. 75–93.

[227] A. Gómez-Pérez and D. Manzano-Macho. *A survey of ontology learning methods and techniques*. Deliverable 1.5. OntoWeb Consortium, 2003.

[228] L. Zhou. “Ontology learning: state of the art and open issues”. In: *Inf. Technol. and Management* 8.3 (Sept. 2007), pp. 241–252.

[229] D. Benz. “Collaborative Ontology Learning”. PhD thesis. Albert-Ludwigs-Universit{"a}t Freiburg, Department of Computer Science, 2007.

[230] C Jacinto and C Antunes. “User-Driven Ontology Learning from Structured Data”. In: *Computer and Information Science (ICIS), 2012 IEEE/ACIS 11th International Conference on*. 2012, pp. 184–189.

[231] P Johannesson. “A method for transforming relational schemas into conceptual schemas”. In: *Data Engineering, 1994. Proceedings.10th International Conference*. 1994, pp. 190–201.

[232] V. Kashyap. "Design and Creation of Ontologies for Environmental Information Retrieval". In: *Twelfth Workshop on Knowledge Acquisition, Modelling and Management Voyager Inn*. Banff, Alberta, Canada, 1999.

[233] D. L. Rubin, M. Hewett, D. E. Oliver, T. E. Klein, and R. B. Altman. "Automating Data Acquisition into Ontologies from Pharmacogenetics Relational Data Sources Using Declarative Object Definitions and XML". In: *Proceedings of the 7th Pacific Symposium on Biocomputing, PSB 2002, Lihue, Hawaii, USA, January 3-7, 2002*. 2002, pp. 88–99.

[234] L. Stojanovic, N. Stojanovic, and R. Volz. "Migrating Data-Intensive Web Sites into the Semantic Web". In: *Proceedings of the 2002 ACM Symposium on Applied computing*. 2002, pp. 1100–1107.

[235] M. Li, X.-Y. Du, and S. Wang. "Learning ontology from relational database". In: *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*. Vol. 6. 2005, 3410–3415 Vol. 6.

[236] K. S. Esmaili and H Abolhassani. "A Categorization Scheme for Semantic Web Search Engines". In: *Computer Systems and Applications, 2006. IEEE International Conference on*. 2006, pp. 171–178.

[237] H. Alani. "Position paper: Ontology construction from online ontologies". In: *In Proc. of 15th International World Wide Web Conference (WWW2006)*. 2006.

[238] H. Suryanto and P. Compton. "Discovery of Ontologies from Knowledge Bases". In: *Proceedings of the 1st International Conference on Knowledge Capture*. K-CAP '01. New York, NY, USA: ACM, 2001, pp. 171–178.

[239] H. Suryanto and P. Compton. *Learning Classification Taxonomies From a Classification Knowledge Based System*. 2000.

[240] M. A. Hearst. "Automatic Acquisition of Hyponyms from Large Text Corpora". In: *Proceedings of the 14th Conference on Computational Linguistics - Volume 2*. COLING '92. Stroudsburg, PA, USA: Association for Computational Linguistics, 1992, pp. 539–545.

[241] G. Rigau. "Automatic Acquisition of Lexical Knowledge from MRDs". PhD Thesis. Universitat Politècnica de Catalunya, 1998.

[242] J. Jannink. "Thesaurus Entry Extraction from an On-line Dictionary". In: *Proceedings of Fusion*. Sunnyvale CA, 1999.

[243] L. Bozzato, M. Ferrari, and A. Trombetta. "Building a domain ontology from glossaries: A general methodology". In: *CEUR Workshop Proceedings* 426 (2008).

[244] J.-U. Kietz, J.-U. Kietz, A. Mädche, A. Mädche, E. Maedche, E. Maedche, R. Volz, and R. Volz. "A Method for Semi-Automatic Ontology Acquisition from a Corporate Intranet". In: *EKAW-2000 Workshop "Ontologies and Text"* (2000), pp. 2–6.

[245] G. A. Miller. "WordNet: A Lexical Database for English". In: *Commun. ACM* 38.11 (1995), pp. 39–41.

[246] N. Aussenac-Gilles, B. Biébow, and S. Szulman. *Corpus Analysis for Conceptual Modelling*. 2000.

[247] L Karoui, M. Aufaure, and N Bennacer. "Ontology discovery from web pages: Application to tourism". In: *Workshop of Knowledge Discovery and Ontologies* (2004).

[248] H. Davulcu, S. Vadrevu, and S. Nagarajan. "OntoMiner: Bootstrapping Ontologies from Overlapping Domain Specific Web Sites". In: *Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers &Amp; Posters*. WWW Alt. '04. New York, NY, USA: ACM, 2004, pp. 500–501.

[249] M. Hazman, S. R. El-Beltagy, and A. Rafea. "Ontology Learning from Domain Specific Web Documents". In: *Int. J. Metadata Semant. Ontologies* 4.1/2 (May 2009), pp. 24–33.

[250] N. Aussenac-Gilles, B. Biebow, and S. Szulman. "Revisiting ontology design: a method based on corpus analysis". In: *Knowledge Engineering and Knowledge Management Methods, Models, and Tools*. Springer, 2000, pp. 172–188.

[251] B. Bachimont, A. Isaac, and R. Troncy. "Semantic commitment for designing ontologies: a proposal". In: *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web*. Springer, 2002, pp. 114–121.

[252] C. H. Hwang. "Incompletely and Imprecisely Speaking: Using Dynamic Ontologies for Representing and Retrieving Information." In: *KRDB*. Vol. 21. Citeseer. 1999, pp. 14–20.

[253] L. Khan and F. Luo. "Ontology construction for information selection". In: *Tools with Artificial Intelligence, 2002. (ICTAI 2002). Proceedings. 14th IEEE International Conference on*. 2002, pp. 122–127.

[254] M. Missikoff, R. Navigli, and P. Velardi. "The Usable Ontology: An Environment for Building and Assessing a Domain Ontology". English. In: *The Semantic Web — ISWC 2002*. Ed. by I. Horrocks and J. Hendler. Vol. 2342. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2002, pp. 39–53.

[255] F. Xu, D. Kurz, J. Piskorski, and S. Schmeier. "A Domain Adaptive Approach to Automatic Acquisition of Domain Relevant Terms and their Relations with Bootstrapping." In: *LREC*. 2002.

[256] E. Agirre, O. Ansa, E. Hovy, and D. Martínez. "Enriching very large ontologies using the WWW". In: *Proceedings of the Workshop on Ontology Construction of the European Conference of AI (ECAI-00)*. 2000.

[257] E. Alfonseca and S. Manandhar. "An unsupervised method for general named entity recognition and automated concept discovery". In: *Proceedings of the 1st international conference on general WordNet, Mysore, India*. 2002, p. 34.

[258] A. Faatz and R. Steinmetz. "Ontology enrichment with texts from the WWW". In: *Semantic Web Mining* 20 (2002).

[259] D. Lonsdale, Y. Ding, D. W. Embley, and A. Melby. "Peppering knowledge sources with SALT: Boosting conceptual content for ontology generation". In: *Proceedings of the AAAI Workshop: Semantic Web Meets Language Resources*. 2002, pp. 30–36.

[260] C. Roux, D. Proux, F. Rechenmann, and L. Julliard. "An Ontology Enrichment Method for a Pragmatic Information Extraction System gathering Data on Genetic Interactions". In: *In position paper in Proceedings of the ECAI2000 Workshop on Ontology Learning(OL2000*. 2000.

[261] N Ben-Mustapha, H Baazaoui-Zghal, A Marie-Aude, and H Ben-Ghzala. "Survey on ontology learning from web and open issues". In: *Third International Symposium on Innovation in Information and Communication Technology (ISIICT'2009), Amman Jordan*. 2009.

[262] CEN. "WORKSHOP AGREEMENT European Guide to good Practice in Knowledge Management - Part 1 : Knowledge Management Framework". In: (2004), pp. 1–33.

[263] J. Hahn and M. R. Subramani. "A framework of knowledge management systems: issues and challenges for theory and practice". In: *Proceedings of the twenty first international conference on Information systems*. Association for Information Systems. 2000, pp. 302–312.

[264] P. Heisig. "Harmonisation of knowledge management-comparing 160 KM frameworks around the globe". In: *Journal of knowledge management* 13.4 (2009), pp. 4–31.

[265] R. S. Bhagat, B. L. Kedia, P. D. Harveston, and H. C. Triandis. "Cultural variations in the cross-border transfer of organizational knowledge: An integrative framework". In: *Academy of management review* 27.2 (2002), pp. 204–221.

[266] R. Maier and U. Remus. "Implementing process-oriented knowledge management strategies". In: *Journal of knowledge management* 7.4 (2003), pp. 62–74.

[267] T. H. D. Varun Grover. "General perspectives on knowledge management: Fostering a research agenda". In: *Journal of management information systems* 18.1 (2001), pp. 5–21.

[268] R. Maier. *Knowledge Management Systems: Information and Communication Technologies for Knowledge Management*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2002.

[269] K. M. Wiig. "Knowledge management methods". In: *Arlington (TX)* (1995).

[270] M. Stankosky, F Clabrese, and C Baldanza. *A Systems Approach to Engineering A Knowledge Management System*. Paper presented at the 2003 Knowldge Management: Employing Proven Tools for Results, Washington, DC, 1999.

[271] M. Stankosky. *Creating the discipline of knowledge management: The latest in university research*. Routledge, 2005.

[272] G. Probst, S. Raub, and K. Romhardt. *Managing Knowledge: Building Blocks for Success*. Wiley, 2000.

[273] I. Smeureanu and N. Isaila. "3rd World Conference on Educational Sciences - 2011 Information technology, support for innovation in education sciences". In: *Procedia - Social and Behavioral Sciences* 15 (2011), pp. 751 –755.

[274] P. Fettke and P. Loos. "Classification of reference models: a methodology and its application". In: *Information Systems and e-Business Management* 1.1 (2003), pp. 35–53.

[275] H. Pirkkalainen, S. Thalmann, J. Pawlowski, M. Bick, P. Holtkamp, and K.-H. Ha. "Internationalization processes for open educational resources". In: *Workshop on Competencies for the Globalization of Information Systems in Knowledge-Intensive Settings, ICSOB*. 2010.

[276] C. Park. "In other (people's) words: Plagiarism by university students–literature and lessons". In: *Assessment & evaluation in higher education* 28.5 (2003), pp. 471–488.

[277] D. Weber-Wulff. *False feathers: A perspective on academic plagiarism*. Springer Science & Business, 2014.

[278] J. Bao, C. Lyon, P. C. Lane, W. Ji, and J Malcolm. "Comparing different text similarity methods". In: (2007).

[279] C.-H. Leung and Y.-Y. Chan. "A Natural Language Processing Approach to Automatic Plagiarism Detection". In: *Proceedings of the 8th ACM SIGITE Conference on Information Technology Education*. SIGITE '07. Destin, Florida, USA: ACM, 2007, pp. 213–218.

[280] K. Thackston. *Matt Cutts on Google's Updated SEO Copywriting Strategy*. 2012.

[281] I. Jacobson, M. Ericsson, and A. Jacobson. *The Object Advantage: Business Process Reengineering with Object Technology*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1994. ISBN: 0-201-42289-1.

[282] M. A. Ould and M. Ould. *Business processes: modelling and analysis for re-engineering and improvement*. Vol. 598. Wiley Chichester, 1995.

[283] PNMSoft. *Business Process*. 2014. URL: `http://link.springer.com/article/10.1057/palgrave.jibs.8400102` (visited on 10/15/2016).

[284] A. Delgado, F. Ruiz, I. García-Rodríguez de Guzmán, and M. Piattini. "MINERVA: Model drIveN and sErvice oRiented Framework for the Continuous Business Process improVement and relAted Tools". In: *Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops: International Workshops, ICSOC/ServiceWave 2009, Stockholm, Sweden, November 23-27, 2009, Revised Selected Papers*. Ed. by A. Dan, F. Gittler, and F. Toumani. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 456–466.

[285] M. Weske. *Business process management: concepts, languages, architectures*. Springer Science & Business Media, 2012.

[286] Y. Ducq, D. Chen, and T. Alix. "Principles of Servitization and Definition of an Architecture for Model Driven Service System Engineering". In: *Enterprise Interoperability*. Ed. by M. van Sinderen, P. Johnson, X. Xu, and G. Doumeingts. Vol. 122. Lecture Notes in Business Information Processing. Springer Berlin Heidelberg, 2012, pp. 117–128.

[287] B Nuseibeh and S Easterbrook. "Requirements Engineering: A Roadmap". In: *ISCE '00 Proceedings of the Conference on the Future of Software Engineering*. 2000.

[288] A Finkelstein and W Emmerich. "The future of requirements management tools". In: *Information Systems in Public Administration and Law* (2000).

[289] V. Rogstrand and T. Kjellberg. "The representation of manufacturing requirements in model-driven parts manufacturing". In: vol. 22. February 2015. 2009, 1065–1072.

[290] C. Bozarth and S. Edwards. "The impact of market requirements focus and manufacturing characteristics focus on plant performance". In: *Journal of Operations Management* 15.3 (1997), 161–180.

[291] W. Hausman and D. Montgomery. "Market Driven Manufacturing". In: *Journal of Market-Focused Management* 2.1 (1997), 27–47.

[292] H. L. Tavares, G. G. Rezende, V. M. dos Santos, R. S. Manhães, and R. A. de Carvalho. "A tool stack for implementing Behaviour-Driven Development in Python Language". In: *CoRR* abs/1007.1722 (2010).

[293] E. Keogh. "BDD: A Lean Toolkit". In: *Proceedings of Lean Software & Systems Conference*. 2010.

[294] I. Lazr, S. Motogna, and B. Pírv. "Behaviour-Driven Development of Foundational UML Components". In: *Electron. Notes Theor. Comput. Sci.* 264.1 (Aug. 2010), pp. 91–105.

[295] A. J. Berre, B. Elvesæter, N. Figay, C. Guglielmina, S. G. Johnsen, D. Karlsen, T. Knothe, and S. Lippe. "The ATHENA Interoperability Framework". In: *Enterprise Interoperability II: New Challenges and Approaches*. Ed. by R. J. Gonçalves, J. P. Müller, K. Mertins, and M. Zelm. London: Springer London, 2007, pp. 569–580.

[296] R. Grangel, M. Bigand, and J.-P. Bourey. "A UML profile as support for transformation of business process models at enterprise level". In: *Proceedings of the First International Workshop on Model Driven Interoperability for Sustainable Information Systems (MDISIS 2008)*. Vol. 340. 2008, pp. 73–87.

[297] Y. Lemrabet, M. Bigand, D. Clin, N. Benkeltoum, and J.-P. Bourey. "Model Driven Interoperability in Practice: Preliminary Evidences and Issues from an Industrial Project". In: *Proceedings of the First International Workshop on Model-Driven Interoperability*. MDI '10. Oslo, Norway: ACM, 2010, pp. 3–9.

[298] C. Agostinho, H. Bazoun, G. Zacharewicz, Y. Ducq, and H. Boye. "Information models and transformation principles applied to servitization of manufacturing and service systems design". In: *Model-Driven Engineering and Software Development (MODELSWARD), 2014 2nd International Conference on*. 2014, pp. 657–665.

[299] S. M. Dent. "Partnering intelligence". In: *Creating Value for Your Business by Building strong Alliances. Davies-Black Publishing, CA. USA* (1999).

[300] N. Daclin, D. Chen, and B. Vallespir. "Methodology for enterprise interoperability". In: *Proceedings of the 17th World Congress: The International Federation of Automatic Control Seoul, Korea*. 2008, pp. 12873–12878.

[301] C Agostinho, R Costa, P Malo, and R Jardim-Goncalves. "Product Data integration in the demand of interoperability in e-Business". In: *Intelligent Systems, 2006 3rd International IEEE Conference on*. 2006, pp. 807–812.

[302] A. Sø lvberg. "Data and What They Refer to". In: *Conceptual Modeling*. Ed. by G Goos, J Hartmanis, J van Leeuwen, P. Chen, J. Akoka, H. Kangassalu, and B. Thalheim. Vol. 1565. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1999, pp. 211–226.

[303] C. Agostinho, J. Sarraipa, D. Goncalves, and R. Jardim-Goncalves. "Tuple-Based Semantic and Structural Mapping for a Sustainable Interoperability". English. In: *Technological Innovation for Sustainability*. Ed. by L. Camarinha-Matos. Vol. 349. IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, 2011, pp. 45–56.

[304] M. K. Gold. *Debates in the digital humanities*. U of Minnesota Press, 2012.

[305] R. Jardim-Goncalves and A. Grilo. "Building information modeling and interoperability". In: *Automation in Construction* 19.4 (2010). Building information modeling and interoperability, pp. 387 –.

[306] T. Nasukawa and J. Yi. "Sentiment Analysis: Capturing Favorability Using Natural Language Processing". In: *Proceedings of the 2Nd International Conference on Knowledge Capture*. K-CAP '03. Sanibel Island, FL, USA: ACM, 2003, pp. 70–77.

[307] L. Lizhen, S. Wei, W. Hanshi, L. Chuchu, and L. Jingli. "A novel feature-based method for sentiment analysis of Chinese product reviews". In: *Communications, China* 11.3 (2014), pp. 154–164.

[308] M. Gamon. "Sentiment Classification on Customer Feedback Data: Noisy Data, Large Feature Vectors, and the Role of Linguistic Analysis". In: *Proceedings of the 20th International Conference on Computational Linguistics*. COLING '04. Geneva, Switzerland: Association for Computational Linguistics, 2004.

[309] S Youn, A Arora, P Chandrasekhar, P Mestry, and A Sethi. “Survey about Ontology Development Tools for Ontology-based Knowledge Management”. In: *University of South California* (2009).

[310] D. Dicheva and C. Dichev. “C.: TM4L: Creating and Browsing Educational Topic Maps”. In: *British Journal of Educational Technology - BJET* 37 (2006), pp. 391–404.

[311] M. Krötzsch, S. Schaffert, and D. Vrandecic. “Reasoning in Semantic Wikis”. In: *Proceedings of the 3rd Reasoning Web Summer School, Dresden, Germany*. Ed. by G. Antoniou, U. Aßmann, C. Baroglio, S. Decker, N. Henze, P.-L. Patranjan, and R. Tolksdorf. Vol. 4636. LNCS. Springer, 2007, 310–329.

[312] M. Völkel, ed. *Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics*. Workshop on Semantic Wikis. ESWC2006, 2006.

[313] M. Krötzsch, D. Vrandecic, and M. Völkel. “Semantic MediaWiki.” In: *International Semantic Web Conference*. Vol. 4273. Lecture Notes in Computer Science. Springer, 2006, pp. 935–942.

[314] S Schaffert. “IkeWiki: A Semantic Wiki for Collaborative Knowledge Management”. In: *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006. WETICE ’06. 15th IEEE International Workshops on*. 2006, pp. 388–396.

[315] E. Oren. “SemperWiki: a semantic personal Wiki”. In: *Proceedings of the 1st Workshop on The Semantic Desktop at the ISWC 2005 Conference*. Galway, Ireland, 2005, pp. 107–122.

[316] D Vrandecic and M Krötzsch. “Reusing Ontological Background Knowledge in Semantic Wikis.” In: *SemWiki* (2006).

[317] K. Dello, E. P. B. Simperl, and R. Tolksdorf. “Creating and using Semantic Web information with Makna”. In: *Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics*. Workshop on Semantic Wikis. ESWC2006, 2006.

[318] R. Jardim-Goncalves, A. Grilo, C. Agostinho, F. Lampathaki, and Y. Charalabidis. “Systematisation of Interoperability Body of Knowledge: the foundation for Enterprise Interoperability as a science.” In: *Enterprise IS* 7.1 (2013), pp. 7–32.

[319] FInES Research Roadmap Force. *FInES Research Roadmap 2015*. 2012. URL: http://www.fines-cluster.eu/jm/Manage-Documents/Download-document/323-FInES-Research-Roadmap-2025-v3.0.html (visited on 2014).

[320] C. Agostinho and R. Jardim-Goncalves. “Dynamic Business Networks: A Headache for Sustainable Systems Interoperability”. English. In: *On the Move to Meaningful Internet Systems: OTM 2009 Workshops*. Ed. by R. Meersman, P. Herrero, and T. Dillon. Vol. 5872. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2009, pp. 194–204.

[321] F. Lampathaki, S. Koussouris, C. Agostinho, R. Jardim-Goncalves, Y. Charalabidis, and J. E. Psarras. “Infusing scientific foundations into Enterprise Interoperability.” In: *Computers in Industry* 63.8 (2012), pp. 858–866.

[322] C. Agostinho, R. Jardim-Goncalves, and A. Steiger-Garcao. “Underpinning EISB with Enterprise Interoperability Neighboring Scientific Domains”. English. In: *Revolutionizing Enterprise Interoperability through Scientific Foundations*. IGI Global, Feb. 2014, pp. 41–76.

[323] ISO/TC 184/SC 5. *Advanced automation technologies and their applications — Part 1 : Framework for enterprise interoperability*. (ISO/DIS 11354-1:2011), 2011.

[324] K. R. Parker and J. T. Chao. “Wiki as a teaching tool”. In: *Interdisciplinary Journal of Knowledge and Learning Objects* 3 (2007), pp. 57–72.

[325] F. Dengler and H.-J. Happel. “Collaborative Modeling with Semantic MediaWiki”. In: *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*. WikiSym ’10. Gdansk, Poland: ACM, 2010, 23:1–23:2.

[326] F. Dengler, S. Lamparter, M. Hefke, and A. Abecker. “Collaborative Process Development using Semantic MediaWiki”. In: *Proceedings of the 5th Conference of Professional Knowledge Management*. Solothurn, Switzerland: GI, 2009.

[327] J. Sarraipa, P. Gomes-de Oliveira, C. Marques-Lucena, R. Jardim-Goncalves, and J. M. da Silva. “E-training development approach for enterprise knowledge evolution”. In: *ASME 2013 International Mechanical Engineering Congress and Exposition*. American Society of Mechanical Engineers. 2013, V005T05A014–V005T05A014.

[328] V. Gupta, D. S. Chauhan, and K. Dutta. “Incremental development & revolutions of E-learning software systems in education sector: a case study approach”. In: *Human-centric Computing and Information Sciences* 3.1 (2013), p. 1.

[329] J. G. Ruiz, M. J. Mintzer, and R. M. Leipzig. “The impact of e-learning in medical education”. In: *Academic medicine* 81.3 (2006), pp. 207–212.

[330] R. R. Thomas Jr. *Beyond race and gender: Unleashing the power of your total workforce by managing diversity*. AMACOM Div American Mgmt Assn, 1992.

[331] T. Cox Jr. *Creating the multicultural organization: A strategy for capturing the power of diversity.* Jossey-Bass, 2001.

[332] S. Maxwell, P. Engel, and R. J. Leiteritz. *European development cooperation to 2010*. Overseas Development Institute London, 2003.

[333] Learning Technology Standards Committee (LTSC) of the IEEE. *Draft standard for learning technology - Learning Object Metadata*. Tech. rep. New York: IEEE Standards Department, July 2002. URL: http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf.

[334] A. H. Abdul Karim, A. S. Chaudhry, and C. S. G. Khoo. “Learning Objects Application Profile for Granularity and Reusability: Integrating Dublin Core with IEEE-LOM”. In: *Proceedings of the 2007 International Conference on Dublin Core and Metadata Applications: Application Profiles: Theory and Practice*. DCMI ’07. Singapore: Dublin Core Metadata Initiative, 2007, pp. 116–122.

[335] S. Baldiris, C. Ávila, P. A. Rivera, J. C. Guevara, and R. Fabregat. “Web editing module for tagging metadata of the Fedora Commons repository learning objects under DRD and LOM standards”. In: *2011 Frontiers in Education Conference (FIE)*. 2011, S2E–1–S2E–5.

[336] J. Sarraipa, R. Jardim-Goncalves, and A. Steiger-Garcao. “MENTOR: an enabler for interoperable intelligent systems”. In: *International Journal of General Systems* 39.5 (2010), pp. 557–573.

[337] N. F. Noy, M. Sintek, S. Decker, M. Crubézy, R. W. Fergerson, and M. A. Musen. “Creating Semantic Web Contents with ProtéGé-2000”. In: *IEEE Intelligent Systems* 16.2 (Mar. 2001), pp. 60–71.

[338] G. S. Altshuller. *Creativity as an exact science: the theory of the solution of inventive problems*. Gordon and Breach Science Publishers, 1984.

[339] C Marques-Lucena, C Agostinho, E Marcelino-Jesus, J Sarraipa, and R Jardim-Goncalves. "Collaborative Management of Requirements Using Semantic Wiki Modules". In: *Control Systems and Computer Science (CSCS), 2015 20th International Conference on*. 2015, pp. 665–672.

[340] P. a. Heslin. "Better than brainstorming? Potential contextual boundary conditions to brainwriting for idea generation in organizations". In: *Journal of Occupational and Organizational Psychology* 82 (2009), 129–145.

[341] INCOSE. *Tools Survey: Requirements Management (RM) Tools*. 2002. URL: http://www.incose.org/productspubs/products/setools/tooltax.html.

[342] A. Beal. *The benefits of using a wiki to manage requirements*. 2014. URL: http://www.bridging-the-gap.com/wiki-requirements-management-benefits/.

[343] N. Shadbolt, T. Berners-Lee, and W. Hall. "The Semantic Web Revisited". In: *IEEE Intelligent Systems* 21.3 (2006), pp. 96–101.

[344] M. M. Taye. "Understanding Semantic Web and Ontologies: Theory and Applications". In: *CoRR* abs/1006.4 (2010).

[345] J. Kaminski. "EDITORIAL Folksonomies boost Web 2.0 Functionality". In: 13.2 (2009), 1–8.

[346] D. Fensel. *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*. 2nd ed. Springer-Verlag New York, Inc., 2003.

[347] B. Decker, E. Ras, J. Rech, P. Jaubert, and M. Rieth. "Wiki-Based Stakeholder Participation in Requirements Engineering". In: *IEEE Software* 24.2 (2007), pp. 28–35. ISSN: 0740-7459. DOI: 10.1109/MS.2007.60.

[348] D. Yang, D. Wu, S. Koolmanojwong, A. W. Brown, and B. W. Boehm. "WikiWinWin: A Wiki Based System for Collaborative Requirements Negotiation". In: *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*. 2008, pp. 24–24. DOI: 10.1109/HICSS.2008.502.

[349] D. 51.2. *Methodology for Requirements Engineering & Evaluation in Manufacturing Service Ecosystems - MSEE - Manufacturing Services Ecosystem*. 2012.

[350] Software Engineering. *Requirements Engineering = Elicitation + Analysis and Negotiation + Specification or Documentation + Validation !* 2010. URL: http://se-thoughtograph.blogspot.pt/2010/01/requirement-engineering-key-aspect.html.

[351] D. 51.1. *Applicability of existing user requirements approaches in MSEE. MSEE – Manufacturing SErvices Ecosystem*. 2011.

[352] C. Marques-Lucena, C. Agostinho, S. Koussouris, and J. a. Sarraipa. "Knowledge Management Framework using wiki-based front-end modules". In: *International Conference in Enterprise Information Systems 17th , ICEIS 2015*. Barcelona, Spain: SCITEPRESS, 2015.

[353] F. Cluster. "FINES Research Roadmap". In: *Retrieved, August* 24 (2010), p. 2013.

[354] G. Santucci, C. Martinez, and D. Vlad-Câlcic. "The sensing enterprise". In: *FInES workshop at FIA*. 2012, pp. 1–14.

[355] C. Agostinho, M. Sesana, R. Jardim-Gonçalves, and S. Gusmeroli. "Model-Driven Service Engineering Towards the Manufacturing Liquid-Sensing Enterprise". In: *4th International Cce on Model-Driven and Software Development*. 2015.

[356] C. Agostinho, E. Jesus, J. a. Sarraipa, and C. Lucena. *Final User Requirement & PoC Specification*. Deliverable 2.22. OSMOSE, 2016.

[357] F. Lampathaki, H. P. Athens, P. Kokkinakos, C. Tucci, I. Alvertis, J. Psarras, S. Koussouris, and G. Viscusi. "FutureEnterprise, A Roadmap for Sensing Enterprise". In: (2015). URL: http://ceur-ws.org/Vol-1414/paper9.pdf.

[358] S. Gusmeroli, C. Agostinho, C. Lucena, M. Sesana, A. Felic, and K. Fischer. "OSMOSE: a Paradigm for the Liquid-Sensing Enterprise". In: *Proceeedings of the Workshops of the IWEI 2015 Conference co-located with the 6th International IFIP Working Conference on Enterprise Interoperability IWEI 2015, Nîmes, France, May 27, 2015*. 2015.

[359] Y. Ducq, C. Agostinho, D. Chen, G. Zacharewicz, and R. Jardim-Goncalves. "Generic Methodology for Service Engineering based on Service Modelling and Model Transformation." In: *Manufacturing Service Ecosystem: Achievements of the European 7th Framework Programme FoF-ICT Project* (2014).

[360] MSEE. *Specifications and Design of Service Lifecycle Management ToolBox*. 2013.

[361] H. Bazoun, G. Zacharewicz, Y. Ducq, and H. Boye. "Transformation of Extended Actigram Star to BPMN2.0 and Simulation Model in the Frame of Model Driven Service Engineering Architecture". In: *Proceedings of the Symposium on Theory of Modeling & Simulation - DEVS Integrative M&S Symposium*. DEVS 13. San Diego, CA, USA: Society for Computer Simulation International, 2013, 20:1–20:8.

[362] J. Bézivin, G. Dupé, F. Jouault, G. Pitette, and J. E. Rougui. "First experiments with the ATL model transformation language: Transforming XSLT into XQuery". In: *2nd OOPSLA Workshop on Generative Techniques in the context of Model Driven Architecture*. Vol. 37. 2003.

[363] S. a. White. "Introduction to BPMN". In: *BPTrends* (2004), pp. 1–11.

[364] M. Nagarajan, K. Verma, A. Sheth, J. Miller, and J. Lathem. "Semantic Interoperability of Web Services - Challenges and Experiences". In: *Web Services, 2006. ICWS '06. International Conference on*. 2006, pp. 373–382.

[365] A. Brogi, S. Corfini, and R. Popescu. "Composition-Oriented Service Discovery". English. In: *Software Composition*. Ed. by T. Gschwind, U. Aßmann, and O. Nierstrasz. Vol. 3628. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, pp. 15–30.

[366] K. Arabshian and H. Schulzrinne. "An Ontology-Based Hierarchical Peer-to-Peer Global Service Discovery System". In: *Journal of Ubiquitous Computing and Intelligence* 1.2 (2007), pp. 133–144.

[367] C. Schmidt and M. Parashar. "A Peer-to-Peer Approach to Web Service Discovery". In: *World Wide Web* 7.2 (June 2004), pp. 211–229.

[368] P. Haase, S. Agarwal, and Y. Sure. "Service-Oriented Semantic Peer-to-Peer Systems". English. In: *Web Information Systems – WISE 2004 Workshops*. Ed. by C. Bussler, S.-k. Hong, W. Jun, R. Kaschek, Kinshuk, S. Krishnaswamy, S. Loke, D. Oberle, D. Richards, A. Sharma, Y. Sure, and B. Thalheim. Vol. 3307. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2004, pp. 46–57.

[369] D. Martin, M. Paolucci, S. McIlraith, M. Burstein, D. McDermott, D. McGuinness, B. Parsia, T. Payne, M. Sabou, M. Solanki, N. Srinivasan, and K. Sycara. "Bringing Semantics to Web Services: The OWL-S Approach". English. In: *Semantic Web Services and Web Process Composition*. Ed. by J. Cardoso and A. Sheth. Vol. 3387. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, pp. 26–42.

[370] L. A. Kamaruddin, J. Shen, and G. Beydoun. "Evaluating Usage of WSMO and OWL-S in Semantic Web Services". In: *Proceedings of the Eighth Asia-Pacific Conference on Conceptual Modelling - Volume 130*. APCCM '12. Melbourne, Australia: Australian Computer Society, Inc., 2012, pp. 53–58.

[371] E. Tambouris, S. Gorilas, G. Kavadias, D. Apostolou, A. Abecker, L. Stojanovic, and G. Mentzas. "Ontology-Enabled E-gov Service Configuration: An Overview of the OntoGov Project". English. In: *Knowledge Management in Electronic Government*. Ed. by M. Wimmer. Vol. 3035. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2004, pp. 122–127.

[372] S. Vicente, M. Pérez, X. García, A. Gimeno, and J. Naval. "Terregov: eGoverment Interoperability on a Semantically Driven World". English. In: *Interoperability of Enterprise Software and Applications*. Ed. by D. Konstantas, J.-P. Bourrières, M. Léonard, and N. Boudjlida. Springer London, 2006, pp. 461–463.

[373] SemanticGov Consortium. *SemanticGov project website*. 2007. URL: http://islab.uom.gr/semanticgov/index.php?name=Web_Links\&req=viewlink\&cid=2\&orderby=titleA (visited on 12/28/2015).

[374] M. Sroga. "Access-eGov-personal assistant of public services". In: *Computer Science and Information Technology, 2008. IMCSIT 2008. International Multiconference on*. 2008, pp. 421–427.

[375] FUSION Consortium. *FUSION project website*. 2008. URL: http://www.seerc.org/fusion/semanticregistry/fusion.html (visited on 12/28/2015).

[376] Talend. *Advanced Mapping and Transformation*. 2015. URL: https://www.talend.com/advanced-mapping-and-transformation (visited on 12/28/2015).

[377] SAP. *SAP NetWeaver Composition Environment Library*. 2013. URL: https://help.sap.com/saphelp_nwce72/helpdata/en/66/5455600ef74868b62d0bdc2e4c9842/content.htm (visited on 12/28/2015).

[378] IBM. *WebSphere Application Server V8.5.5*. 2013. URL: http://www.ibm.com/developerworks/websphere/techjournal/1312_borley/1312_borley.html (visited on 12/28/2015).

[379] M. S. Abdullah, I. Benest, A. Evans, and C. Kimble. "Knowledge modelling techniques for developing knowledge management systems". In: *Third European Conference on Knowledge Management: Trinity College Dublin, Ireland*. 2002, p. 17.

[380] INTEROP NoE. *Deliverable MoMo. 2-TG MoMo Roadmap*. InterOp.

[381] C. Agostinho, J. Sarraipa, F. D'Antonio, and R. Jardim-Gonçalves. "Enhancing STEP-based Interoperabity Using Model Morphisms". English. In: *Enterprise Interoperability II*. Ed. by R. Gonçalves, J. Müller, K. Mertins, and M. Zelm. Springer London, 2007, pp. 817–828.

[382] M. Delgado, C. Agostinho, P. Malo, and R. Jardim-Goncalves. "A framework for STEP-based harmonization of conceptual models". In: *Intelligent Systems, 2006 3rd International IEEE Conference on*. 2006, pp. 781–785.

[383] INTEROP NoE. *Problem Analysis and Roadmap - I-VLab Platform.* InterOp.

[384] J. Dauben. *Georg Cantor: His Mathematics and Philosophy of the Infinite*. History of science. Princeton University Press, 1990.

[385] P. A. Bernstein. "Applying Model Management to Classical Meta Data Problems". In: *Proceedings of the 1st Biennial Conference on Innovative Data Systems Research (CIDR).* 2003.

[386] J. Sarraipa, N. Zouggar, D. Chen, and R. Jardim-Goncalves. "Annotation for enterprise information management traceability". In: *ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers. 2007, pp. 893–899.

[387] J. Bézivin, G. Dupé, F. Jouault, G. Pitette, and J. E. Rougui. "First experiments with the ATL model transformation language: Transforming XSLT into XQuery". In: *2nd OOPSLA Workshop on Generative Techniques in the context of Model Driven Architecture*. 2003.

[388] OMG. *Meta Object Facility (OMG) 2.0 Query/View/Transformation Specification*. 2008. URL: http://www.omg.org/spec/QVT/1.0/PDF/ (visited on 11/22/2015).

[389] J. Gharajedaghi. *Systems Thinking: Managing Chaos and Complexity: A Platform for Designing Business Architecture*. Elsevier Science, 2011.

[390] C. Lucena, J. a. Sarraipa, and R. Jardim-Goncalves. "A Knowledge Management Framework to Support Online Communities Creation". In: *Technological Innovation for Collective Awareness Systems*. Vol. 423. IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, 2014, pp. 29–36.

[391] C. Agostinho, J. Černý, and R. Jardim-Goncalves. "MDA-Based Interoperability Establishment Using Language Independent Information Models". English. In: *Enterprise Interoperability*. Ed. by M. van Sinderen, P. Johnson, X. Xu, and G. Doumeingts. Vol. 122. Lecture Notes in Business Information Processing. Springer Berlin Heidelberg, 2012, pp. 146–160.

[392] CRESCENDO IP. *Collaborative and Robust Engineering using Simulation Capability Enabling Next Design Optimisation (FP7-234344)*. 2009. URL: http://cordis.europa.eu/publication/rcn/13781_en.html (visited on 12/02/2015).

[393] W3C. *OWL 2 Web Ontology Language MOF-Based Metamodel*. 2008. URL: http://www.w3.org/2007/OWL/wiki/MOF-Based_Metamodel (visited on 12/02/2015).

[394] R. L. Ackoff. "From data to wisdom". In: *Journal of applied systems analysis* 16.1 (1989), pp. 3–9.

[395] K. Simon, T. Hornung, and G. Lausen. "Learning Rules to Pre-process Web Data for Automatic Integration." In: *RuleML*. Ed. by T. Eiter, E. Franconi, R. Hodgson, and S. Stephens. IEEE Computer Society, 2006, pp. 107–116.

[396] T. Wilson, J. Wiebe, and P. Hoffmann. "Recognizing Contextual Polarity in Phrase-level Sentiment Analysis". In: *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*. HLT '05. Vancouver, British Columbia, Canada: Association for Computational Linguistics, 2005, pp. 347–354.

[397] K. Dave, S. Lawrence, and D. M. Pennock. "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews". In: *Proceedings of the 12th International Conference on World Wide Web*. WWW '03. Budapest, Hungary: ACM, 2003, pp. 519–528.

[398] P. A. Torres-carrasquillo, E. Singer, M. A. Kohler, and J. R. Deller. "Approaches to language identification using Gaussian mixture models and shifted delta cepstral features". In: *Proc. ICSLP 2002*. 2002, pp. 89–92.

[399] E. Tromp and M. Pechenizkiy. "Graph-based n-gram language identification on short texts". In: *Proc. 20th Machine Learning conference of Belgium and The Netherlands*. 2011, pp. 27–34.

[400] A. H. Branco and J. Silva. *Tokenization of Portuguese: resolving the hard cases*. Tech. rep. 2003. URL: http://hdl.handle.net/10451/14199.

[401] D. P. Metzler and S. W. Haas. "The Constituent Object Parser: Syntactic Structure Matching for Information Retrieval". In: *ACM Trans. Inf. Syst.* 7.3 (July 1989), pp. 292–316.

[402] E. von Glasersfeld and P. P. Pisani. "The Multistore Parser for Hierarchical Syntactic Structures". In: *Commun. ACM* 13.2 (Feb. 1970), pp. 74–82.

[403] H. Gonçalo Oliveira and P. Gomes. "Revisiting the ontologising of semantic relation arguments in wordnet synsets". In: *Natural Language Engineering* (2015), (published online on July 2015).

[404] A Branco and J Silva. "Evaluating Solutions for the Rapid Development of State-of-the-Art POS Taggers for Portuguese." In: *LREC* (2004).

[405] J Silva, R Reis, P Gonçalves, and A Branco. "LX-Parser and LX-DepParser: Online Services for Constituency and Dependency Parsing". In: *inf.pucrs.br* ().

[406] J Silva, A Branco, S Castro, and R Reis. "Out-of-the-box robust parsing of Portuguese". In: *Processing of the Portuguese* (2010).

[407] M. S. Chaves, L. A. de Freitas, and R. Vieira. "Hontology: A Multilingual Ontology for the Accommodation Sector in the Tourism Industry." In: *KEOD*. Ed. by J. Filipe and J. L. G. Dietz. SciTePress, 2012, pp. 149–154.

[408] M. Silva, P. Carvalho, and L. Sarmento. "Building a Sentiment Lexicon for Social Judgement Mining". In: *Computational Processing of the Portuguese Language*. Ed. by H. Caseli, A. Villavicencio, A. Teixeira, and F. Perdigão. Vol. 7243. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 218–228.

[409] C. Marques-Lucena, C. Agostinho, S. Koussouris, J. Sarraipa, and R. Jardim-Gonçalves. "Collaborative Knowledge Management Using Wiki Front-End Modules". In: *Enterprise Information Systems: 17th International Conference, ICEIS 2015, Barcelona, Spain, April 27-30, 2015, Revised Selected Papers*. Ed. by S. Hammoudi, L. Maciaszek, E. Teniente, O. Camp, and J. Cordeiro. Cham: Springer International Publishing, 2015, pp. 69–86.

[410] C. Lucena, J. a. Sarraipa, and R. Jardim-Goncalves. "A Knowledge Management Framework to Support Online Communities Creation". English. In: *Technological Innovation for Collective Awareness Systems*. Ed. by L. Camarinha-Matos, N. Barrento, and R. Mendonça. Vol. 423. IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, 2014, pp. 29–36.

[411] C. Marques-Lucena, J. Sarraipa, C. Agostinho, and R. Jardim-Gonçalves. "Model-driven Approach for the Interoperability of Enterprises' Services Information Exchange". In: *Proceedings of the 4th International Conference on Model-Driven Engineering and Software Development*. 2016, pp. 789–799.

[412] C. Marques-Lucena, J. Ferreira, M. Sesana, K. Fischer, and A. Carlos. “Process Modelling Approach for the Liquid-Sensing Enterprise”. In: *Conference: I-ESA 2016 - Interoperability for Enterprise Systems and Applications*. InterOP-VLab. 2016.

[413] C. Agostinho, C. Marques-Lucena, M. Sesana, A. Felic, K. Fischer, C. Rubattino, and J. Sarraipa. “Osmosis Process Development for Innovative Product Design and Validation”. In: *ASME 2015 International Mechanical Engineering Congress and Exposition*. American Society of Mechanical Engineers. 2015.

[414] J. Sarraipa, M. Ferro-Beca, C. Marques-Lucena, and R. Jardim-Goncalves. “Knowledge Management support in Sensing Enterprises Establishment”. In: *IFAC Proceedings Volumes* 47.3 (2014), pp. 839–844.

[415] C. Marques-Lucena, C. Agostinho, S. Koussouris, and J. Sarraipa. “Knowledge Management Framework using Wiki-based Front-end Modules”. In: *Proceedings of the 17th International Conference on Enterprise Information Systems*. 2015, pp. 79–86.

[416] A. Felic, S. Alexakis, C. Agostinho, C. Marques-Lucena, K. Fischer, and M. Sesana. “Osmotic Event Detection and Processing for the Sensing-Liquid Enterprise”. English. In: *On the Move to Meaningful Internet Systems: OTM 2015 Workshops*. Ed. by I. Ciuciu, H. Panetto, C. Debruyne, A. Aubry, P. Bollen, R. Valencia-García, A. Mishra, A. Fensel, and F. Ferri. Vol. 9416. Lecture Notes in Computer Science. Springer International Publishing, 2015, pp. 147–156.

[417] F. Luís-Ferreira, C. Marques-Lucena, J. Sarraipa, and R. Jardim-Goncalves. “Framework for Management of Internet Objects in Their Relation With Human Sensations and Emotions”. In: *ASME 2013 International Mechanical Engineering Congress and Exposition*. American Society of Mechanical Engineers. 2013, V03AT03A078–V03AT03A078.

[418] C. Lucena, L. Palma, A. Cardoso, and P. Gil. “Adaptive quasi-optimal gains tuning of PI-fuzzy controllers”. In: *2012 IEEE 16th International Conference on Intelligent Engineering Systems (INES)*. 2012, pp. 399–404.

[419] C. Lucena, L. Palma, A. Cardoso, and P. Gil. “Optimal gains tuning of PI-fuzzy controllers”. In: *Control Automation (MED), 2012 20th Mediterranean Conference on*. 2012, pp. 824–829.

[420] C. Lucena, A. Cardoso, L Palma, and P. Gil. “Gains tuning of fuzzy controllers based on a cost function optimization”. In: Controlo. 2012.

[421] F Januário, A Santos, C Lucena, L Palma, A Cardoso, and P Gil. “Outliers Accommodation in Fuzzy Control Systems over WSAN”. In: *Intelligent Decision Technologies: Proceedings of the 5th KES International Conference on Intelligent Decision Technologies (KES-IDT 2013)*. Vol. 255. IOS Press. 2013, p. 334.

[422] P. Gil, C. Lucena, A. Cardoso, and L. Palma. “Fuzzy controllers gains tuning: a constrained nonlinear optimization approach”. In: *Neural Computing and Applications* 23.3 (2013), pp. 617–624.

[423] P. Gil, C. Lucena, A. Cardoso, and L. Brito Palma. “Gain Tuning of Fuzzy PID Controllers for MIMO Systems: A Performance-Driven Approach”. In: *Fuzzy Systems, IEEE Transactions on* 23.4 (2015), pp. 757–768.

[424] A. Sebastião, C. Lucena, L Palma, A Cardoso, and P. Gil. “Optimal tuning of scaling factors and membership functions for mamdani type PID fuzzy controllers”. In: *Control, Automation and Robotics (ICCAR), 2015 International Conference on*. IEEE. 2015, pp. 92–96.

[425] A. Sebastião, C. Lucena, L. Palma, and P. Gil. "Fuzzy-PID gains tuning: Differential evolution versus second order analytical algorithms". In: *Control Conference (ASCC), 2015 10th Asian*. 2015, pp. 1–6. DOI: 10.1109/ASCC.2015.7244378.

2016

Framework for collaborative knowledge management in organizations

Catarina Lucena