Designing a Method for Discovering Expertise in Cyber Security Communities: an ontological approach

Marcelo Paiva Fontenele

Submitted in partial fulfilment of the requirements of
University of Reading for
the degree of Doctor of Philosophy

Department of Computer Science, University of Reading

October 2016

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

This work used a third party proofreading service for quality improvement.

Marcelo Paiva Fontenele

Abstract

Cyber security aims to protect our connected society from threats affecting services that rely on cyberspace. The pervasive nature of those threats requires a collaborative engagement in which a heterogeneous set of stakeholders request or provide security services. One of the major challenges in current cyber security initiatives is to place skilled people wherever needed whilst reducing the overall knowledge gap. Thus, in order to orchestrate roles in such a complex and dynamic environment, a novel approach to discover talent within the cyber security community is required.

This PhD research addresses this challenge by devising a conceptual model and an ontological methodology, which aids a robust discovery of the fittest expertise driven by the specific needs of cyber security projects, as well as benchmarking expertise shortages. Talent management, knowledge management and organisational modelling theories provide the theoretical foundations upon which the cyber security community is articulated. Mixed methods were performed within a cyber security community to triangulate findings in the literature, test the method and appraise the solution.

The method for discovering expertise in cyber security communities (DECYSE) is capable of delivering a seamless solution for processes involving expertise discovery. This method enables learning from previous projects; supports selection, ranking and assessment of experts according to specified requirements in a project profile; and provides indicators to measure knowledge gaps and shortages in the cyber security community. The DECYSE method is robust and underpinned by analytical techniques, considering complex interactions and perspectives from the actors involved. In order to promote ongoing improvement on the method itself, this thesis also details the conceptual model which articulates the requirements for developing DECYSE.

A round of experiments was successfully conducted, where a team of three experts, out of sixty-six participant profiles, met the criteria in a cyber security project. The method was also positively appraised by a board of experts working with strategic CS projects. DECYSE enables ongoing improvement and contributes to both theory and the cyber security community.

Acknowledgements

I would like to express my gratitude to my supervisor Dr. Lily Sun for her constant support during my research and even before my arrival in the UK. Her patience, continuous guidance and inspiring commitment as a supervisor and as a research professional were crucial for unveiling new knowledge and understanding the challenges of an academic career. It was certainly a true privilege. I would also like to thank my co-supervisor Dr. Frederic Stahl who provided enlightening answers and constructive advice.

To my wife Gláucia, I would like to express my most profound feelings for her love, encouragement and invaluable support, which provided the motivating environment to carry on with my work.

To my family in Brazil, I am grateful for their love, emotional support and positive thoughts throughout my life.

I would like to acknowledge that this work was realised with support from CNPq, National Council for Scientific and Technological Development – Brazil, as well as my colleagues in the Brazilian Army for creating this lifetime opportunity and providing the requested feedback that helped me develop my thesis.

To my friends and colleagues in UK, thank you for your company and fruitful discussions.

List of Publications

Fontenele, M.P., Sampaio, R.B., Silva, A.I.B., Fernandes, J.H.C., Sun, L. (2014). Performing Knowledge Requirements Analysis for Public Organisations in a Virtual Learning Environment: a Social Network Analysis approach. Journal of Information Technology & Software Engineering. ISSN: 2165-7866. doi:10.4172/2165-7866.1000134

Fontenele, M.P., Sun, L. (2015). Talent discovery for cyber security: an organisational perspective. In Proceedings of International Conference on Cyber Situational Awareness, Data Analytics and Assessment, London, pp. 253–259. doi:10.1109/CyberSA.2015.7166122

Fontenele, M.P., Sun, L. (2016). Knowledge management of cyber security expertise: an ontological approach to the talent discovery. In Proceedings of International Conference on Cyber Security and Protection of Digital Services, London, pp. 466-490. doi:10.1109/CyberSecPODS.2016.7502356

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter presents the research background and introduces the importance of discovering talented individuals for the cyber security community and addressing the skills gap. The discussion introduces the main research areas and the motivations to carry out this research. The research problem is discussed under theoretical and methodological perspectives. Then, the research questions, aims and objectives are unveiled. In conclusion, the structure of the thesis is presented.

1.1 Research Background and Motivation

This research is based on the context of expertise discovery within cyber security communities and considers a particular interest in articulating information requirements to promote the user's continuous and collaborative engagement. The research background of this thesis is set on socio-technical systems, bridging the fields of cyber security, talent management, knowledge management and organisational theory. This section briefly introduces the research background with some key concepts and the motivations that guided this research work.

Prior to introducing those key concepts, a remark on the research theme is worth noting. The concept of expertise relates to knowledge based on experience (Davenport and Prusak, 2000), while discovery encompasses learning, innovation (Oxford Dictionaries, 2009), interactions and change (Brown and Duguid, 1991). Thus, this thesis defines expertise discovery in the context of cyber security communities as the dynamic and self-improved allocation of the most capable individuals for cyber security projects. Our approach also implies that the expertise requested as a criterion in a project can be structured as a set of capabilities.

Cyber security (CS) refers to securing information and non-information assets that rely on cyberspace (for storage or communication), cyberspace users (either playing an individual or a collective role) and even cyberspace itself from cyber attacks (von Solms and van Niekerk, 2013). The adopted approach on CS in this thesis is social-oriented rather than technical-oriented, because it investigates the need for people and skills within a web of trust.

Talent refers to people (either the whole population or a select group of high performers / high potentials) and their capabilities. Such capabilities result from a combination of natural ability, commitment, mastery of skills and proper allocation (Gallardo-Gallardo et al., 2013). Talent Management (TM) provides additional practices which include the development of a specific architecture, performance evaluation and integration between internal and external talent pools (Collings and Mellahi, 2009).

Knowledge can be “viewed as a mix of information, understanding, capability, experience, skills and values” (Rowley, 2007: p. 174). Knowledge management (KM) plays a key role in communities of innovation, where stakeholders co-evolve through seamless knowledge flows across organisational borders. Related research areas include multi-criteria decision-making, reputation systems and analytics. Moreover, KM can help in understanding the identification and development of skills in organisations. Therefore, organisational modelling approaches and proper techniques for knowledge representation are discussed (Alavi and Leidner, 2001; Owen, 1999).

Trust relies on positive expectations upon the behaviour of others (Lewicki et al., 1998). It is an important subject to support the decision-making process and knowledge exchange within organisational partnerships (Harris, 2011), especially among CS stakeholders (Choo, 2011).

Society is growing dependent on digital services built on cyberspace. The underlying fast changing technology in cyberspace requires capable operators to combat emerging threats that exploit digital service consumers and providers. Hence, CS has become a common goal for society. In the past few years, attracting talented individuals, developing skills and retaining such individuals have become some of the main concerns in national CS (Caldwell, 2013; ENISA, 2014). In contrast, initiatives and collaborative efforts in national CS have been investing in a set of practices and perspectives, which have not been successfully integrated and optimised. Thus, managing capabilities and ensuring trusted information for such a complex and heterogeneous set of stakeholders have become some of the major challenges for national CS (Choo, 2011). This research was devised while visualising national CS as a system that nurtures itself on the interactions of its members and as a seamless community of innovation guided by a common goal. Aligned with this approach, stakeholders co-evolve by collaborating with different types of resources (e.g. services and knowledge) and by gaining benefits of such resource exchanges in manifold ways. Hence, the choice of modelling the expertise discovery scope within the national CS context led to the investigation of theories, methods and techniques normally employed in organisational environments.

This research is motivated by having witnessed permanent loss of misplaced talented people in different agencies, which eventually sought for better work conditions, during the time that this researcher was involved with Brazilian CS projects. In contrast, it became noticeable how a slight change of position in given opportunities could substantially enhance motivation, optimise self-improvement and create a trans-organisational network of committed experts. Such a change of position was focused on addressing the experts' needs whenever possible. Notwithstanding, searching for individuals to participate in CS projects or activities remains a cumbersome task, either due to the lack of structured search mechanisms or due to regular bureaucratic barriers between institutions. Moreover, sources used during this research for discovering talent to work in CS projects were found to either lack reliability, relevant information or a combination of those options. Although motivation for this research work originated within a Brazilian CS community, special care has been taken to present guidelines based on common issues across different societies. These issues seem widespread across societies, despite their maturity levels or budgets for investing in CS initiatives.

1.2 Problem Statement

The relevance of CS is determined on its scale, reaching different sectors of society, and on how these sectors rely on information and services in cyberspace. One of the major concerns that drive CS is discovering proper expertise in an effective fashion in such a complex environment. The literature has shown that such concern is influenced by a number of factors that have not been resolved. For example, Caldwell (2013) acknowledges a shortage of CS skills should be dealt with by attracting the right talent and developing capabilities, which in this thesis include academic areas, certifications and competencies. Moreover, such a shortage does not seem to be properly benchmarked. In line with these perspectives, there has been a lack of awareness and review mechanisms for expertise, which are common pitfalls while deploying CS strategies. These perspectives also include the need to identify critical services for the CS community (Klimburg, 2012).

This research has found that the current methodology for discovering talents is still unstructured, manual or empiric. One of the issues that contribute to adopting such a methodology is a lack of mechanisms to integrate, measure and process relevant and trustworthy information about experts, since CS is such a multidisciplinary and sensitive topic. Despite some collaborative initiatives focused on education (e.g. e-skills UK, 2014) and competence frameworks (e.g. US Government, 2014), no mechanisms to effectively discover those talents and understand their shortage were found in the literature. Even these initiatives do not ensure that the expertise claimed by individuals is reliable. In fact, Kouttis (2016) argues that the shortage of skills persists even with high investments in education. In addition, no evidence of systematic learning from the rich and dynamic knowledge environment within the CS domain was found. Therefore, the research problem involves a lack of integration and continuous improvement of the expertise discovery processes for the national CS community. In order to provide better understanding on the subjects and approaches involved, the research problem is henceforth presented according to theoretical and methodological perspectives.

The need for CS talent (Caldwell, 2013) has led to an investigation on a number of theories. The Talent Management (TM) theory reveals that managing talent in a multi-organisational and dynamic CS environment is a task that has not been efficiently solved or grounded on research. Hence, TM alone is not capable of addressing the entire research problem.

The current scenario to address the skills shortage in CS lacks a systemic view and research underpinning. This can lead to a biased understanding of stakeholders' needs and to static solutions that lose alignment with the dynamic CS environment over time. In addition, supporting expertise discovery is still a challenge due to the wide variety of sources, absence of consensual understanding on concepts and unreliable information. A review on the fields of Knowledge Management (e.g. decision support and trust) and organisational modelling was conducted in this thesis in order to provide the means for a novel and comprehensive solution to address such issues. Findings on this research, however, revealed that existing approaches and methods convey modelling techniques suited only particular aspects of the research problem. Hence, those approaches had to be combined and adapted to address the research problem as a whole. For example, reputation systems are used in this thesis as a means to ensure reliability on profiled information through co-creation and to benchmark the performance of experts.

The challenge of developing a method to perform discovery of expertise on such a scale requires a novel combination of approaches providing scalability and flexibility.

Knowledge requirements in CS change over time. Recognising these changes as well as retaining the proper knowledge needs more systematic processing. This work contributes to the aforementioned fields by presenting a knowledge-driven, self-evolving and robust methodology comprising techniques to profile criteria, measure the expertise shortage and enhance expertise discovery for CS projects.

1.3 Research Questions

The problem is widespread, requiring a multidisciplinary approach due to its complexity, and has not yet been solved, leading to the following research questions:

- How can a methodology be developed and maintained to aid a robust discovery of expertise within a collaborative CS environment, where requirements for expertise are dynamic and evolving?
- What criteria and metrics can be formulated in performing the expertise discovery workflow within the complex CS environment?
- How can an expertise shortage be methodologically described in order to target skill development and to satisfy the fast-changing CS environment?

1.4 Aims and Objectives

This research aims to contribute to the fields of CS, Talent Management and Knowledge Management and to stakeholders in the context of national CS to support an effective and systematic discovery of expertise for CS projects. The contribution adopts a methodological perspective focused on integrating talent practices and views from distinct stakeholders (e.g. government agencies, the private sector and individuals/the Public). The following objectives strive to achieve this goal:

- To determine the main expertise discovery challenges of the CS communities based on the findings in academic publications and in governmental documents.
- To explore the best practices in expertise discovery and identify the viewpoints to overcome the problems which prevent CS communities from involving the right expertise for the right projects and understanding the expertise shortage.
- To examine techniques for representing, measuring and analysing capability requirements for projects in CS and identify limitations on selected techniques in the context of CS expertise discovery.

- To establish a research design guided by replicable meta-processes in order to ensure continuous improvement of the research solution.
- To create a method of DECYSE which methodologically facilitates discovering expertise in CS communities through a pluralistic, seamless and self-evolving information flow based on a unique combination of consolidated techniques.
- To test DECYSE by applying its techniques to an experiment, while validating its acceptance in the CS community in terms of generalisability, applicability and robustness.
- To evaluate the results, the implications to research and to the CS community, and the development process for their rigorousness and critically assess the DECYSE limitations for future work.

1.5 The Thesis Structure

The remainder of this thesis is divided as illustrated in Figure 1.1. Some chapters present findings in the literature and summaries that refer back to the research questions. A list of abbreviations (Appendix A) is provided to facilitate quick reference on the acronyms used throughout this work.

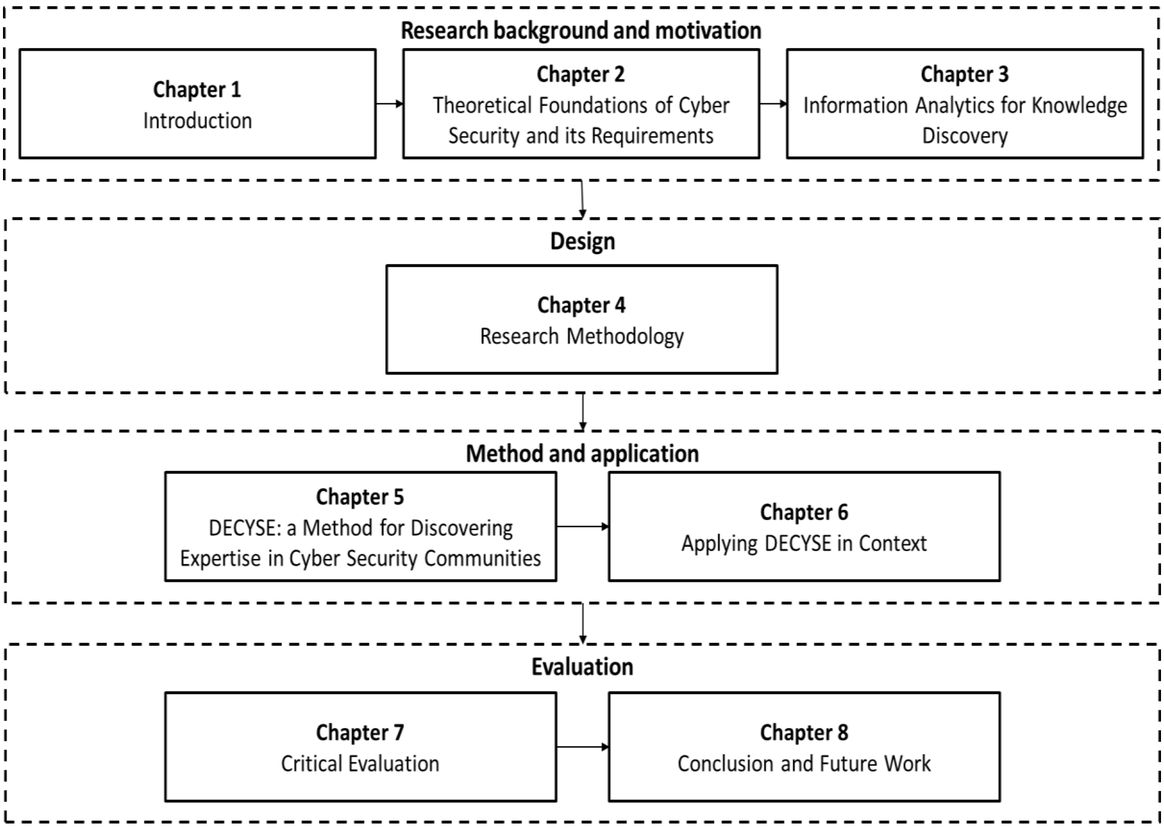


Figure 1.1: The thesis outline

Chapter 2 investigates the context of this research, which comprises the fields of CS, Talent Management (TM), Knowledge Management (KM) and organisational theories. The CS scenario is outlined based upon recurring initiatives drawn from research and government documents. In compliance with findings in CS, TM concepts are introduced with particular emphasis on talent practices suitable to address the research questions. In addition, knowledge definitions and processes are presented along with a discussion on trust concepts. The role of organisational modelling is briefly introduced with the discussion of community of practice, business ecosystems, requirements engineering, service-oriented theory and the organisational semiotics approach.

Chapter 3 builds on the theoretical foundations presented in Chapter 2 by examining representation approaches, information sources and processing techniques in order to guide the development of DECYSE. Information representation is discussed based on systems thinking approaches, ontology and user profiling. Social media is discussed as an information source. Analytical techniques (e.g. decision-making and performance evaluation) are further introduced regarding their role in data processing as a means to support selected talent practices. Reliability of profiled information is discussed through an overview on reputation systems.

Chapter 4 performs a thorough review on the research methodology leading to the adopted paradigms, methods and techniques that drive the development of this work. The chapter also discusses the context "as is" and draw features for DECYSE. The investigation enabled to triangulate the CS literature and scope the actors, services and types of criteria for expertise discovery within the domain context.

Chapter 5 introduces the articulation of the research solution and the DECYSE method, which represents an expertise discovery methodology underpinned by an ontological approach. DECYSE is capable of recommending criteria when defining projects, profiling relevant information on experts, rating the expertise; selecting and ranking the fittest candidates for CS projects, providing suitable feedback for members in the CS community; and benchmarking the expertise shortage by performing analytics.

Chapter 6 describes an empirical validation of DECYSE. The chapter introduces the datasets for articulating the research problem and for the experiments. Afterwards, results from experimenting DECYSE are presented in terms of project specification, registering participants, selecting candidates, providing feedbacks and performing analytics.

Chapter 7 critically evaluates the components of the research. Initially, DECYSE is empirically validated through an experiment. Following this appraisal, the implications and contributions to research fields and practice are evoked and then the adopted approach regarding the research design is justified. Finally, limitations of the adopted approach are discussed.

Chapter 8 concludes the activities and findings performed in this thesis based on the research questions and objectives and suggests future directions to extend this research work.

Chapter 2

Theoretical Foundations of Cyber Security and its Requirements

This chapter explores theories that underline the issues and challenges in developing and managing talents for cyber security. Research areas, such as cyber security (CS), talent management (TM), knowledge management (KM) and organisational modelling lay the theoretical foundations to this study. Critiques are drawn upon some of these research areas to contextualise findings within the research problem. Each critique presents a framework which highlights theoretical aspects in those research areas.

2.1 Cyber Security

CS has become a global concern in the information era and a central challenge for government, business, society and even the international community. Hence, it requires a set of regulations, a joint effort built upon knowledge, skilled personnel and a trusted environment (Caldwell, 2013; Cebula and Young, 2010; Klimburg, 2012). This section provides an overview of CS aspects drawn from both academic papers and governmental initiatives in order to contextualise the research problem. Existing gaps in the literature are addressed with a pilot questionnaire (further introduced in Section 4.2.1 and discussed in Section 6.1), which is briefly referred to in this section whenever required. First, some of the main ideas from CS definitions are drawn. Then, the growing importance of CS is discussed, introducing the main players involved, the motivations that drive those players and how they are dealing with the issue. Afterwards, some of the educational initiatives introduce the types of knowledge that are relevant to the CS domain. Finally, some findings on CS are presented.

2.1.1 Conceptual Aspects of Cyber Security

CS is an inter-disciplinary field comprising a wide variety of topics (Julisch, 2013), such as information security (von Solms and van Niekerk, 2013) and risk management (Cebula and Young, 2010). It seems that just a few sources clearly distinguish CS and related concepts such as the information security. While information security is achieved by improving confidentiality, integrity and availability of information (ISO/IEC 27000:2009, 2009; von Solms and van Niekerk, 2013), CS involves applying those elements to cyberspace (Bayuk, 2012; ISO/IEC 27032:2012, 2012; Klimburg, 2012). Cyberspace is the virtual and automated information network supported by information technology (IT) infrastructures that is connected on a global scale. It includes the Internet, computer systems and telecommunications (French Government, 2011; German Federal Ministry of the Interior, 2011; Oxford Dictionaries, 2009; US Government, 2009).

Conceptually, CS involves protecting information systems (French Government, 2011; German Federal Ministry of the Interior, 2011; Klimburg, 2012; US Government, 2009) and technology against threats from cyberspace. Technology also includes non-information based assets (e.g. an individual's personal image or data used to control SCADA[1] systems), as depicted in Figure 2.1. Likewise, CS does not include protecting information that is not within the extent of cyberspace (e.g. a document stored in a safe), although this is still an information security issue.

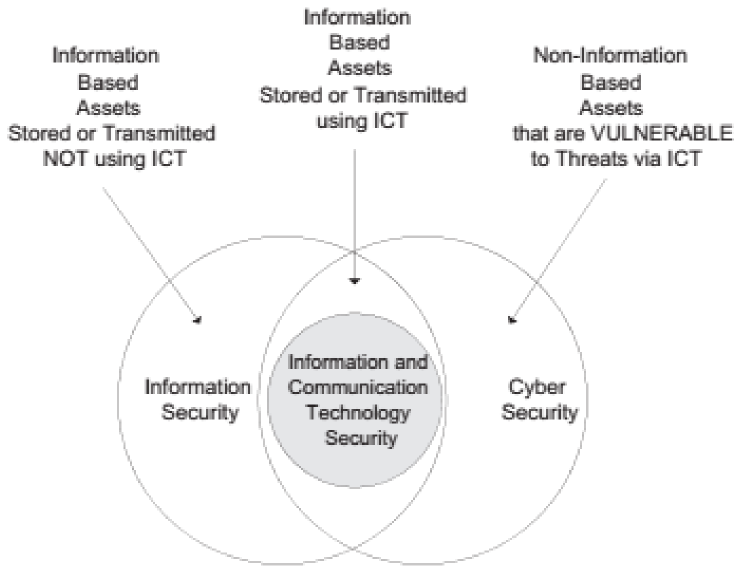


Figure 2.1: The relationship between information security and cyber security (von Solms and van Niekerk, 2013: p. 101)

There are some other concepts closely related to CS, such as cyber defence and cyber warfare. Cyber defence is "the set of all technical and non-technical measures allowing a State to defend in cyberspace information systems that it considers to be critical" (French Government, 2011: p. 21). According to Clarke and Knake (2012: p. 6), cyber war comprises "actions by a nation-state to penetrate another nation's computers or networks for the purposes of causing damage or disruption". Klimburg (2012) argues that both

[1] "supervisory control and data acquisition (SCADA) [are] systems used to manage large-scale industrial control systems (ICS) at industrial facilities" (Choo, 2011: p. 724).

concepts relate to military cyber operations within CS. Due to the use of different concepts across countries, in this work, the term “cyber defence” relates to the more general concept of CS. Appendix B presents a list with CS related definitions for further reference. In order to present a comprehensive overview, CS is discussed according to the recurring initiatives and its main actors.

2.1.2 Strategic Initiatives

One of the most significant challenges to coordinating CS is deploying regulations, guidance and priorities that can effectively raise awareness and motivate engagement (Klimburg, 2012). CS requires a set of regulations (Cebula and Young, 2010), such as policies, strategies, guidelines, best practices, standards and security safeguards (ITU, 2014). These norms are deployed through a series of methods that involve people, processes and technologies (Bayuk, 2012). Due to the complexity of the domain, however, von Solms and van Niekerk (2013) argue that the current norms are not comprehensive enough to secure cyberspace. There is difficulty in addressing fuzzy stakeholders, setting and following-up an agenda for them and complying with regulations, but there is actually a lack of review mechanisms which eventually leads to obsolete or inconsistent regulations (Klimburg, 2012). The research solution provides feedback to keep relevant expertise up-to-date.

Due to the sudden emergence of the role of cyberspace in national security, more than 50 countries have already addressed this issue by publishing a CS strategy or some other form of official document delineating their stance (Klimburg, 2012; von Solms and van Niekerk, 2013). Examples of these initiatives are seen in several governmental reports (e.g. Australian Government, 2009; Brazilian Government, 2015; French Government, 2011; German Federal Ministry of the Interior, 2011; Public Safety Canada, 2014; UK Cabinet Office, 2011; US Government, 2009). Other countries that already deployed CS strategies include the Czech Republic, Estonia, India, Japan, Lithuania, Luxembourg, the Netherlands, New Zealand, Romania, Slovakia, South Africa, South Korea, Spain, Switzerland and Uganda. One of the common themes across those strategies involves promoting trust and social prosperity (Klimburg, 2012). Hence, the challenge for coordinating CS is worldwide and some examples of initiatives to improve compliance with regulation are hereby discussed.

CS strategies are also important for addressing funding, drawing trends for the private sector and highlighting objectives for education (Caldwell, 2013). The UK published their CS Strategy in 2011 and committed funding of £650m over four years to support the National Cyber Security Programme (UK Cabinet Office, 2011). After two years, the UK Science Minister argued that an extra investment was needed, producing a total of £860 million until 2016 (UK Cabinet Office, 2013). The growth of initiatives for protecting business, investment for training and the creation of new business opportunities was stressed during the Infosecurity Europe Conference (2014).

The Brazilian Government (2008a) published the National Strategy of Defence (END) in 2008. It outlined activities and measures to increase security in the “cyber sector”. The Brazilian Government (2008b) also proposed standards (some of which are based on known technical standards) and guidelines stating the main goals for public organisations and the private sector. Such guidelines addressed training, certifications and general information security subjects.

Most countries acknowledge the impacts that cyber threats might pose to national critical systems. Therefore, protecting critical infrastructures is a common issue addressed in different countries’ CS documents (e.g. Australian Government, 2009; Brazilian Government, 2008a; French Government, 2011; German Federal Ministry of the Interior, 2011; Klimburg, 2012; Public Safety Canada, 2014; UK Cabinet Office, 2011; US Government, 2009).

The growing sophistication, scale and persistent nature of recent incidents involving government agencies have been a major concern for most governments. Cyber threats affect individual end users, businesses, government and society as a whole (Choo, 2011). Cyber threats may undermine society by provoking failure on basic services provided by critical infrastructures, exploiting financial services and decreasing systemic economic value due to intellectual property issues (von Solms and van Niekerk, 2013).

Estonia suffered a massive cyber attack in 2007 (Geers, 2010) with long term impacts (Choo, 2011) and is considered to be the first country that suffered the effects of cyberwar. Estonia’s government, financial and computer networks were paralysed by a series of cyber attacks (Gjelten, 2011). The aftermath led to the creation of NATO Cooperative Cyber Defence Centre of Excellence in Estonia that was supported and sponsored by a group of nations (Klimburg, 2012).

The growing dependence of information and communication technology (ICT) in electricity grids is a matter of concern (Ananda Kumar et al., 2014; Pearson, 2011; Wang and Lu, 2013). SCADA systems support many critical infrastructures, from public transport to industrial manufacturing systems and are vulnerable to cyber attacks (Nicholson et al., 2012). The consequences of attacks to such systems can vary from temporary outages to a collapse of an entire power grid (Choo, 2011). Stuxnet became a classical example of malware that targeted a SCADA system (Nicholson et al., 2012) and is known as the “first deployed cyber weapon in history” (Ananda Kumar et al., 2014: p. 129).

Due to the sensitive nature of the CS context, trust that personnel have in information within cyberspace is critical for nations (von Solms and van Niekerk, 2013). A trusted information-sharing environment for individuals and organisations is required to enable partnerships between public and private sectors (Choo, 2011; ENISA, 2014; Klimburg, 2012).

Since CS affects society as a whole, it requires a joint effort between various sectors of society and individuals. Most CS strategies and related documents acknowledge the need for partnerships with the private sector in order to protect their critical infrastructures against cyber threats (e.g. Australian Government, 2009; Brazilian Government, 2015; French Government, 2011; German Federal Ministry of the Interior, 2011; Klimburg, 2012; Public Safety Canada, 2014; UK Cabinet Office, 2011; US Government, 2009). Some initiatives illustrate successful cases of such partnerships, such as “Get Safe Online” (2016), that aim to provide awareness and training for the general population. Other initiatives focus on cyber defence activities (Choo, 2011; Gjelten, 2011).

In order to develop state of the art technologies, there must be investments in research and development (Choo, 2011). In fact, the US Government (2009) acknowledges opportunities for academia, industry and government with commercial opportunities, training and other incentive mechanisms.

International collaboration is also a recurring subject among CS strategies. Sponsored by more than 10 countries, the North Atlantic Treaty Organisation Cooperative Cyber Defence Centre of Excellence (NATO CCD COE) supports research, conferences, education, training and consultancy (Klimburg, 2012). It seems, therefore, that much of the effort in CS has been placed in partnerships, regulation and education. This thesis assumes that the ultimate goal of those CS initiatives is having capable experts to occupy given positions or to perform tasks that promote safety in cyberspace whenever requested. However, this goal has not been properly addressed in the literature.

2.1.3 Actors involved in Cyber Security

CS comprises the triad people-process-technology. This means that security cannot be achieved simply by implementing technology or optimising processes: it requires capable operators (Bayuk, 2012). These people are also referred to as talent (Caldwell, 2013; Conrad, 2012; Conti and Easterly, 2010; Klimburg, 2012; UK Cabinet Office, 2011; US Government, 2009), although without a clear conceptual distinction, which leads to an investigation on such topic. von Solms and van Niekerk (2013) argue that individuals are playing an increasing role in CS, because they are either threats, vulnerabilities or assets to be protected. Indeed, skills, knowledge and proper human resource management (e.g. availability, staffing, training and development) are operational CS risks (Cebula and Young, 2010). Hence, there is a need to structure the employment of those experts and identify the actors involved in such tasks. Notwithstanding, discovering the right talents for CS projects is still an inefficient and manual task, since the major expertise sources are highly unstructured and lack common analytical processes. For example, a pilot questionnaire indicates that the majority of individuals looking for expertise rely on networking in conferences and recommendations from peers.

The key players are CS experts; however, there are other players involved. Among the other players, there are those who need to discover such expertise or those who can collaborate to improve the expertise (Caldwell, 2013). Klimburg (2012) introduces three dimensions of activity in CS (i.e. governmental, national and international). Each dimension of activity is associated with a group of stakeholders performing a specific role (i.e. coordination, cooperation and collaboration, respectively). For example, the government is in charge of coordinating the CS effort through its agencies in various forms (e.g. military, law enforcement, intelligence and others). The national dimension of activity involves different sectors of society (e.g. critical infrastructures, private sector, academia and society as a whole) performing cooperation roles. For example, the private sector on the one hand needs those experts to protect their assets. On the other hand, the private sector can provide educational services for the community. Because there are no boundaries in cyberspace, some strategies also recognise the need for international collaboration through agreements, politics or diplomacy. As previously stated, however, those stakeholders do not have well defined responsibilities within the CS community (Klimburg, 2012). Due to such a complex set of stakeholders affecting the discovery of expertise, a broader perspective is required rather than strictly focusing on CS experts. Hence, this thesis considers these issues by further determining and analysing the roles that those stakeholders perform, which contribute to expertise discovery (as further discussed in Section 5.1).

To organise the coordination effort, countries are either creating their own CS centres or assigning CS authority to governmental agencies (e.g. French Government, 2011; German Federal Ministry of the Interior, 2011; Klimburg, 2012; Lewis and Timlin, 2011; UK Cabinet Office, 2011). This indicates that CS communities may rely on an organisation capable of providing infrastructure to mediate expertise discovery processes.

2.1.4 Required Knowledge and Skills Working in Cyber Security

Understanding how the CS community is addressing the skills shortage can provide valuable cues on how to find those individuals holding such knowledge. This section introduces some of those initiatives and investigates recurring subjects in CS courses and frameworks in order to map the expertise.

Despite the aforementioned initiatives to promote CS, authors still recognise a "shortage of information security savvy *talent*" (Conti and Easterly, 2010: p. 2) and a skills shortage in CS. Some of the issues that contribute to this shortage are outdated curricula in different degrees of education and experience requirements (Caldwell, 2013). Choo (2011) adds that governments cannot have all CS expertise at hand without an adequate public-private partnership (PPP). The military is also concerned with "creating a career path to effectively recruit, manage and retain cyber talent" (Conti and Easterly, 2010: p. 3). Hence, there is a need to share the expertise for the benefit of the whole CS community.

Due to such skills shortage, recruiting, developing and retaining skilled personnel has become critical for CS (ENISA, 2014). Another option to addressing the skills shortage is sharing resources, whether by pooling talent, information / knowledge or services (Caldwell, 2013; French Government, 2011; German Federal Ministry of the Interior, 2011; Klimburg, 2012; UK Cabinet Office, 2011; US Government, 2009).

Cyber threats change over time (UK Cabinet Office, 2011) and so does the nature of the required skills (US Government, 2009). Thus, "cyber warriors" must have "up-to-date knowledge, skills and experiences" (Choo, 2011: p. 727). Besides creating opportunities and attracting skilled personnel, user awareness and education are critical to mitigate cyber threats (Klimburg, 2012). Hence, there is a need to invest in education and research as a long-term measure (Choo, 2011). However, a race for skills development along with the lack of regulation may lead into a confused patchwork of qualifications that may be difficult to choose from (Caldwell, 2013). This research work considers awareness about up-to-date recommended CS capabilities, as well as systematic accreditation for education providers.

Some PPP initiatives address such issues by defining the recommended certifications for government professionals dealing with different subjects within the CS domain (Brazilian Government, 2013a). Other PPP initiatives intend to focus on research, development (Australian Government, 2009; UK Cabinet Office, 2011; US Government, 2009) and educational activities (Brazilian Government, 2014; French Government, 2011; German Federal Ministry of the Interior, 2011; Public Safety Canada, 2014).

New opportunities for individuals, schools and private companies arise from governmental initiatives, such as capture the flag exercises (Dodge Jr. et al., 2007), massive open online courses and Cyber Security Challenge competitions (UK Cabinet Office, 2013). Such challenges also aim to identify and bring "new *talent*" into the business and support a "community of ethical hackers" (UK Cabinet Office, 2011: p. 28). Some initiatives evolved into portals comprising career opportunities, courses and frameworks supporting the CS effort. Examples of these initiatives can be seen in e-skills UK (2014), SFIAplus (BCS, 2015) and NICCS (US Government, 2014). The latter example presents a framework directly related with topics in CS. Such a framework has seven major categories and each comprises several distinct specialty areas. Each specialty area, in its turn, shares some of the 65 available competencies and is generally associated to a set of job titles. In addition, each specialty area shares some of the KSA's (knowledge[2], skill[3] and ability[4]) and is associated with some of the available CS tasks. Likewise, the Brazilian Government (2013a: p. 2) links recommended certifications to the development and validation of "skills, knowledge, competencies, abilities and professional fitness".

The academy also offers masters-level courses (Caldwell, 2013). Competencies for CS can also be drawn from IT skills databases (e.g. ACM, 2015) or academic and governmental initiatives (Table 2.1). In this thesis, these competencies are further regarded as recommended capabilities in order to test the research solution.

Table 2.1: Subjects for cyber security

Subjects	**References**
Information security	Management: (Brazilian Government, 2013a), (Klimburg, 2012), (Lancaster University, 2014), (NYU Polytechnic School of Engineering,

[2] "a body of information applied directly to the performance of a function" (US Government, 2014)
[3] "an observable competence to perform a learned psychomotor act" (US Government, 2014)
[4] "competence to perform an observable behavior or a behavior that results in an observable product" (US Government, 2014)

Subjects	References
	2014), (University of London, 2014), (University of Surrey, 2014); architecture: (Caldwell, 2013), (University of Warwick, 2014); governance: (Cranfield University, 2014), (The National Skills Academy, 2014), (University of Oxford, 2014); using ICT: (von Solms and van Niekerk, 2013), (US Government, 2014)
Risk management	(Brazilian Government, 2013a), (Caldwell, 2013), (Cranfield University, 2014), (The National Skills Academy, 2014), (University of Oxford, 2014), (University of Warwick, 2014), (US Government, 2014)
Intelligence / threat analysis / surveillance	(Caldwell, 2013), (Cranfield University, 2014), (NYU Polytechnic School of Engineering, 2014), (University of Warwick, 2014), (US Government, 2014)
Network security	(Brazilian Government, 2013a), (Cranfield University, 2014), (Edinburgh Napier University, 2014), (Lancaster University, 2014), (NYU Polytechnic School of Engineering, 2014), (Tallinn University of Technology, 2014), (University of London, 2014), (University of Surrey, 2014), (University of Warwick, 2014), (US Government, 2014)
Incident handling	(Brazilian Government, 2013a), (Cranfield University, 2014), (The National Skills Academy, 2014), (University of Oxford, 2014), (US Government, 2014)
Digital forensics	(Brazilian Government, 2013a), (Edinburgh Napier University, 2014), (Lancaster University, 2014), (University of Dallas, 2014), (University of Oxford, 2014), (University of London, 2014), (University of Surrey, 2014), (University of Warwick, 2014), (US Government, 2014)
Software / systems developing / architecture security	(Brazilian Government, 2013a), (NYU Polytechnic School of Engineering, 2014), (The National Skills Academy, 2014), (University of London, 2014), (University of Surrey, 2014), (US Government, 2014)
Business continuity	(Brazilian Government, 2013a), (The National Skills Academy, 2014), (University of Surrey, 2014)
Audit and compliance / systems testing (e.g. penetration testing)	(Brazilian Government, 2013a), (Cranfield University, 2014), (Lancaster University, 2014), (NYU Polytechnic School of Engineering, 2014), (Tallinn University of Technology, 2014), (The National Skills Academy, 2014), (University of Dallas, 2014), (University of London, 2014), (US Government, 2014)
Digital certification and accreditation	(Brazilian Government, 2013a), (Cranfield University, 2014), (US Government, 2014)
Law and regulations	(Lancaster University, 2014), (University of Dallas, 2014), (University of London, 2014), (US Government, 2014)
Cryptography	(Edinburgh Napier University, 2014), (NYU Polytechnic School of Engineering, 2014), (Tallinn University of Technology, 2014), (University of Dallas, 2014), (University of London, 2014), (University of Surrey, 2014), (US Government, 2014)
Cloud computing	(Brazilian Government, 2013a), (Edinburgh Napier University, 2014), (University of Oxford, 2014), (University of Surrey, 2014)
Mobile security	(Brazilian Government, 2013a), (Edinburgh Napier University, 2014), (University of Oxford, 2014)
Social networks / multimedia technologies	(Brazilian Government, 2013a), (Cranfield University, 2014), (US Government, 2014)
Operational security management	(Cranfield University, 2014), (The National Skills Academy, 2014), (University of Dallas, 2014)
Information assurance methodology and testing	(Cranfield University, 2014), (The National Skills Academy, 2014), (University of Surrey, 2014), (US Government, 2014)

Subjects	References
Critical infrastructures (SCADA, resilience)	(Cranfield University, 2014)
Change management	(Cranfield University, 2014)
Service-Oriented Architectures	(Edinburgh Napier University, 2014)

Although some fields relate to information security, this work considers some subjects in information security as part of CS (cf. Figure 2.1). Caldwell (2013) argues that the first three fields presented in Table 2.1 are the top demanded fields for CS professionals.

Besides attracting and developing skills, Choo (2011) argues that the average government salaries and the absence of a career for those individuals can result in a continuous depletion of skilled professionals. In contrast, professionalisation efforts seem not to be fully embraced by practitioners. The reasons for this attitude include a resistance on current certifications and competence schemes imposed by governments which do not respect the dynamic nature of CS professionals. These schemes could exclude experts qualified by experience which cannot be replaced with certifications and academic degrees (Reece and Stahl, 2015). The research solution considers all the aforementioned aspects (i.e. certifications, competencies, academic degrees and experience) to discover the right expert.

2.1.5 Findings about Expertise Discovery within CS Initiatives

The major critique for expertise discovery within CS initiatives is that such a topic has not been properly addressed so far, despite all available means. Knowing the right expertise and how to find it is critical to optimising CS and, therefore, all the investments that are made in those initiatives. The following paragraphs highlight the issues affecting such assumption and consolidate some of the findings with regard to the current CS expertise discovery scenario.

Some of the findings from the literature and from a pilot questionnaire indicate the need to integrate the national CS initiatives, stakeholders and their needs efficiently in a trusted and systemic informational environment. According to these perspectives and to the research questions, some of the deliverables required for integrating Talent Management to support CS are presented in Figure 2.2. Arrows outbound from stakeholders (written in bold text) represent the resources and services provided. On the other hand, inbound arrows represent the current benefits that stakeholders gain with those CS initiatives, which reside in a fuzzy environment. This thesis assumes that such an environment resembles an ecosystem where the stakeholders contribute and benefit (i.e. co-evolve) with regard to the common goal of expertise discovery. Hence, further discussion on this topic is conducted in Section 2.4.1. Solid shapes represent the situation as is, while dashed shapes represent the context to be. Hence, it seems that modelling such an environment is crucial to effectively improve and structure national CS in order to address the research problem. The theoretical support for this approach is discussed in the following sections.

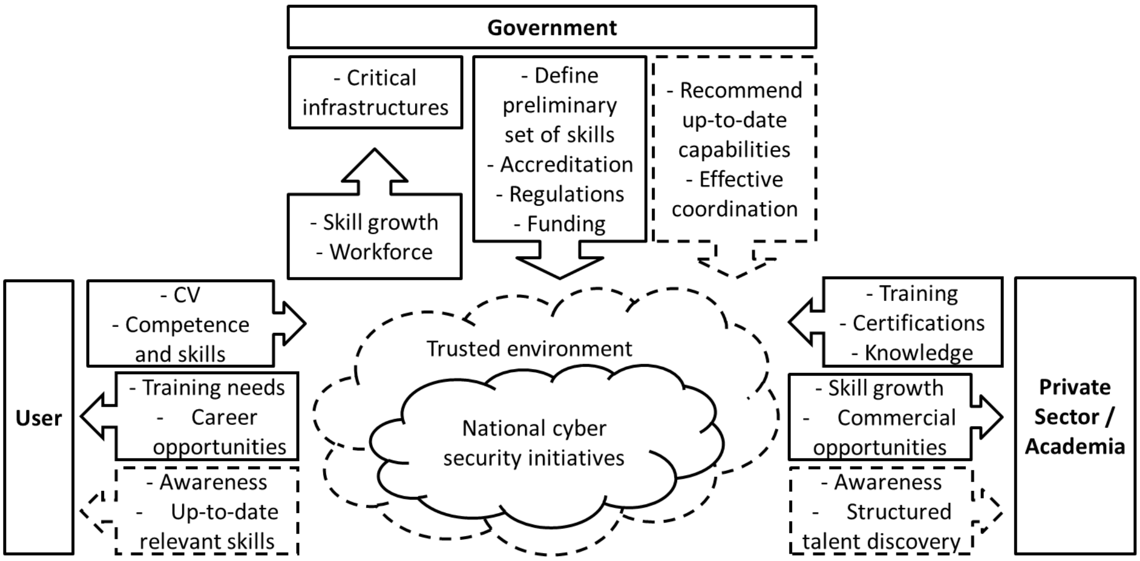


Figure 2.2: Cyber security coordination and cooperation framework

From national CS framework manuals (ENISA, 2014; Klimburg, 2012) to standards (ISO/IEC 27032:2012, 2012), much of the legislation regarding CS is already available and can serve as a starting point. Most of the current national CS strategies already indicate that the main procedures to increase CS and some countries are already developing compliant initiatives. Such concepts and regulations still lack proper dissemination, however, affecting the awareness of stakeholders. In addition, despite the existing CS websites for education and awareness and others containing regulations, no single portal or framework integrating such initiatives with expertise discovery mechanisms was found. A pilot questionnaire, however, indicates that expertise discovery for CS projects is still highly unstructured and empirical. It is not even clear to what extent the roles that stakeholders play in CS contribute to expertise discovery. This research adopts a method to identify those roles and responsibilities, which promote CS expertise discovery.

There is a shortage of personnel with the necessary knowledge (Caldwell, 2013; Kouttis, 2016) and, therefore, individuals need to be properly recruited (ENISA, 2014). This research work assumes that skills databases with trustworthiness features where people could register and update their CVs should support mapping the available expertise and some initiatives (e.g. e-skills UK, 2014) have been providing such a service. Search mechanisms, however, do not address the information needs of stakeholders and no mechanisms to ensure trust in such information were found. It is interesting to stress that defining the desired skills is a step prior to mapping the individuals who have such skills, which in turn makes it possible to determine the skills shortage. In fact, the wide variety and complexity of competences (some of which are presented in Table 2.1) and certifications still poses a challenge to identify and to develop career paths for CS professionals. Even some of the most advanced frameworks that define competences and skills do not share the same concepts with knowledge sources (e.g. academic courses and certifications). Moreover, mechanisms to identify emerging CS competences, skills and certifications over time were not found in the literature and still require further investigation. Those concepts seem to be the type of criteria by which expertise discovery is normally conducted. Therefore, the definition and recommendation of up-to-date skills and certifications falls into the context of this thesis.

The e-skills UK (2014) is an example of responding to some major issues in terms of accreditation, IT skills improvement and career development in CS. Developed as a comprehensive and pragmatic solution for several stakeholders, such initiative addresses career development through profiles. Such a solution, however, does not represent a model that can be replicated in other societies, but was rather tailored with a static set of capabilities to address the problem in one nation. In addition, no integrated and systematic solution to address the shortage of skills and to select the relevant human resource was found in the literature or similar initiatives. Besides providing career opportunities, it seems that mapping the expertise supply and demand is a necessary challenge to effectively generate metrics and improve the discovery of expertise over time. Hence, it is believed that such a task is capable of reducing the skills shortage by addressing issues laid out by Caldwell (2013), such as defining experience requirements. This thesis not only identifies the types of CS capabilities, but also improves awareness on expertise by delivering a mechanism to discover the most relevant capabilities over time. Relevance, in this thesis, is determined by benchmarking the expertise supply and demand. Talent Management and Knowledge Management fields are further discussed to shed some light on such issues.

2.2 Talent and Talent Management

Managing skilled personnel is one of the main concerns of CS (Choo, 2011; Caldwell, 2013; US Government, 2009). This section provides an overview of managing talents by defining "who" is considered talented, "what" is Talent Management (TM) and "why" is it so important. Then, the discussion continues to define "how" to deploy TM features by using frameworks and processes. Afterwards, typical TM aspects and requirements, such as performance, talent pool and selected human resource (HR) practices, are discussed. Finally, some viewpoints on TM are presented to connect the subject with the research questions.

2.2.1 Definitions of Talent

Understanding who and what can be considered as talent can help identify the kind of people and the features demanded in the CS domain. Whether with sports athletes, musicians, gifted children or in work, the term talent is being widely used nowadays (Gallardo-Gallardo et al., 2013). Business leaders believe that the search for talented personnel is the most important organisational concern for the first decade of this millennia (Deloitte, 2011). However, because talent is a relative term (Iles, 2013) and due to its different definitions (see Appendix C), it is not clear what it means to be talented (Gallardo-Gallardo et al., 2013; Thunnissen et al., 2013). Table 2.2 summarises some talent attributes drawn from the aforementioned definitions.

Table 2.2: Terms associated with talent in the literature (extended from Gallardo-Gallardo et al., 2013: p. 293)

Associated terms	Sources
Ability	(Gagne, 2000); (Gallardo-Gallardo et al., 2013); (Michaels et al., 2001); (Oxford Dictionaries, 2009); (Schiemann, 2014); (Silzer and Dowell, 2009); (Tansley et al., 2006)
Attitude	(Michaels et al., 2001)
Behaviour	(Cheese, 2007); (Schiemann, 2014)
Capacity	(Hinrichs, 1966)
Character	(Festing and Schäfer, 2014); (Michaels et al., 2001)
Commitment	(Gallardo-Gallardo et al., 2013); (Ulrich and Smallwood, 2012)
Competence / competency	(Festing and Schäfer, 2014); (Silzer and Dowell, 2009); (Ulrich and Smallwood, 2012)
Contribution	(Ulrich and Smallwood, 2012)
Creativity	(Hinrichs, 1966)
Drive	(Festing and Schäfer, 2014); (Michaels et al., 2001)
Effectiveness	(Hinrichs, 1966)
Experience	(Cheese, 2007); (Festing and Schäfer, 2014); (Michaels et al., 2001); (Schiemann, 2014)
Gift	(Michaels et al., 2001)

Associated terms	Sources
Intelligence	(Festing and Schäfer, 2014); (Hinrichs, 1966); (Michaels et al., 2001)
Judgement	(Michaels et al., 2001)
Knowledge	(Cheese, 2007); (Festing and Schäfer, 2014); (Gagne, 2000); (Lin, 2010); (Michaels et al., 2001); (Tansley et al., 2006); (Schiemann, 2014); (Ulrich and Smallwood, 2012)
Mastery	(Gagne, 2000); (Gallardo-Gallardo et al., 2013)
Potential	(Tansley et al., 2006)
Skills	(Cheese, 2007); (Festing and Schäfer, 2014); (Gagne, 2000); (Hinrichs, 1966); (Lin, 2010); (Michaels et al., 2001); (Oxford Dictionaries, 2009); (Schiemann, 2014); (Silzer and Dowell, 2009); (Tansley et al., 2006); (Ulrich and Smallwood, 2012)
Values	(Schiemann, 2014), (Tansley et al., 2006), (Ulrich and Smallwood, 2012)

Talent can be perceived under two complementary dimensions. The first one comprehends the object dimension (which focuses on talent attributes) and the second one is known as the subject dimension (focusing on individuals as talents). The attributes in Table 2.2 are normally associated with the "talent as object" dimension (Gallardo-Gallardo et al., 2013; Thunnissen et al., 2013). Figure 2.3 summarises the different perceptions of talent.

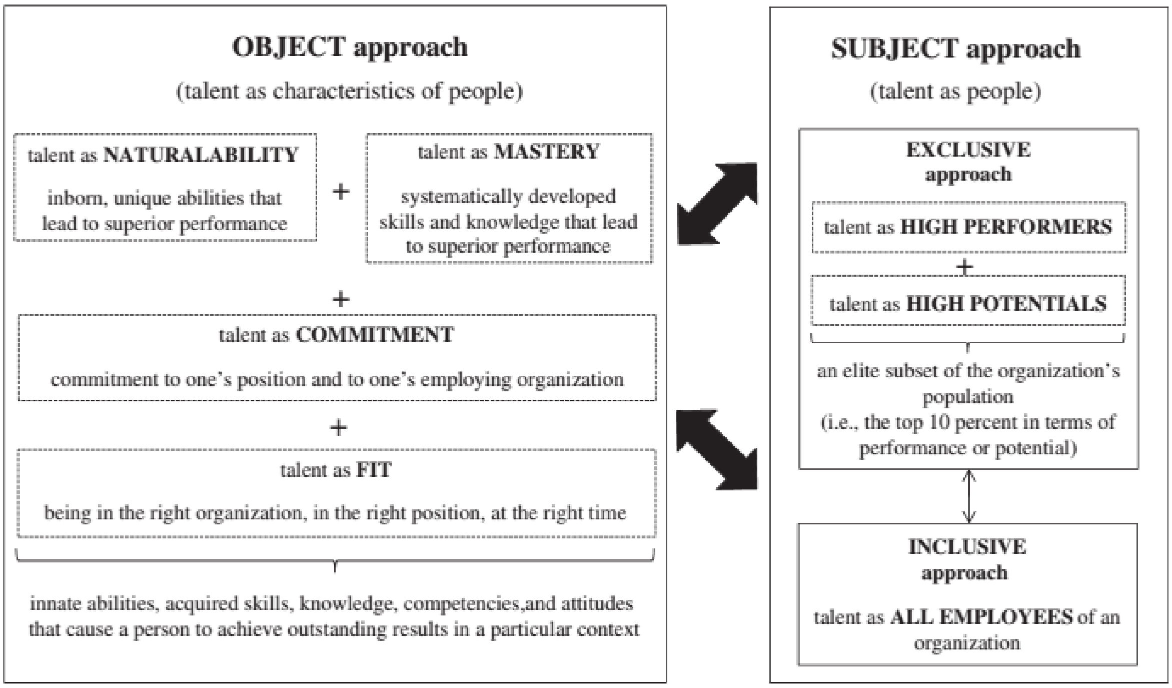


Figure 2.3: Summary of talent perception (Gallardo-Gallardo et al., 2013: p. 297)

According to the object dimension, talent aspects (such as those in Table 2.2) can be a result of either natural ability, mastery of developed skills, commitment or proper work assignment (Gallardo-Gallardo et al., 2013). Some authors consider talent as natural ability because talented personnel are rare, innate, inimitable and difficult to replace (Barney, 1991). Others argue that talent may be mastered through self-development, experience and improving required knowledge and skills (Ulrich and Smallwood, 2012). The commitment approach relates to determination and passion to work. Indeed, differential treatment and rewards can create a competitive and challenging environment where all employees are stimulated to develop and apply useful skills. Finally, talent is also seen as the proper fit or personnel allocation, because talents bloom when people are in the right place, performing the right tasks at the right time (Gallardo-Gallardo et al., 2013).

The subject dimension of talent offers both an exclusive approach, in which talented personnel constitute a minority, and an inclusive approach, in which every person is considered a talent (Gallardo-Gallardo et al., 2013). Ulrich and Smallwood (2012) argue that every employee should be considered as a talent. However, a drawback for this approach is that managing talent turns out to be basically the same as strategic human resource management (HRM), and "talent" could turn into just a synonym for people (Lewis and Heckman, 2006). The exclusive subject approach (talent as some people) relies on identifying who can be considered talent. Some authors argue that talented people are the high performers and have high potential among the whole group (Gallardo-Gallardo et al., 2013; Hor et al., 2010). According to Ulrich and Smallwood (2012), technical experts are generally those with high potential.

While there are advantages and drawbacks for both inclusive and exclusive approaches (Gallardo-Gallardo et al., 2013), some authors argue that they are all equally viable and depend on what is the best fit for the organisation (Dries, 2013). Others recommend a balanced approach that benefits both talented and non-talented individuals (Hughes and Rog, 2008; Thunnissen et al., 2013). In order to clarify the meaning of talent and what it represents for CS, the field of Talent Management is introduced according to a set of relevant practices, challenges and requirements.

2.2.2 Characteristics of Talent Management

Authors argue that there is no consensus in the concept of Talent Management (TM) (Collings and Mellahi, 2009; Festing and Schäfer, 2014; Tansley et al., 2013; Thunnissen et al., 2013) and that TM may just be applying HRM in a faster way (Lewis and Heckman, 2006). However, there are recurring concepts in TM research, some of which are drawn from TM definitions (Appendix C), such as those presented in Figure 2.4. These practices are discussed in the following sections.

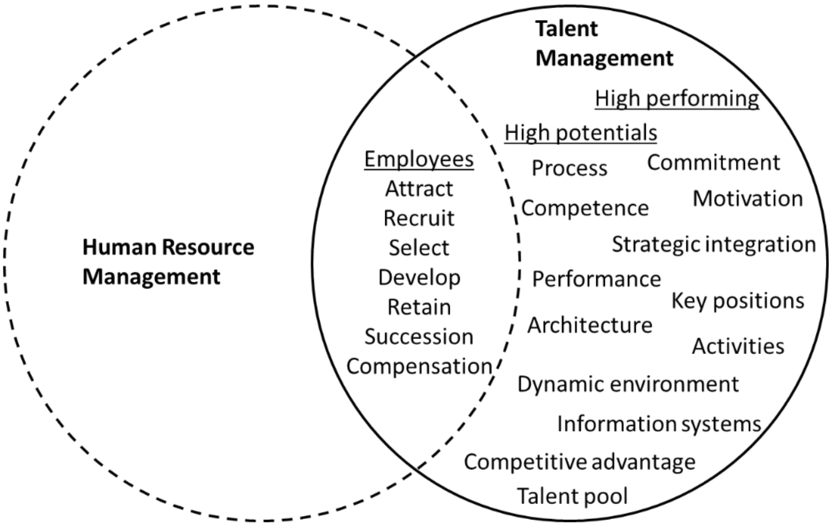


Figure 2.4: Typical TM ideas and common HR practices

Besides the application of sound HR practices, Talent Management implies strategic integration between internal and external factors and requires information systems and architecture (Collings and Mellahi, 2009; Hughes and Rog, 2008; Lewis and Heckman, 2006).

The expected major outcome of Talent Management is increasing individual and organisational performance (Collings and Mellahi, 2009). Some authors agree that performance is the function of Ability (A), Motivation (M) and Opportunity (O) to perform, also known as the AMO framework (Boselie et al., 2005). Indeed, there is a similarity among elements from the AMO framework, the object approach of talent (cf. Figure 2.3) and the talent definition (Ulrich and Smallwood, 2012), as presented in each row of Table 2.3. The latter approach introduces the competence as a set of skills, knowledge and other values required for a position in an organisation.

Table 2.3: The AMO framework and perceptions of talent

AMO framework	Perceptions of talent	Talent definition
Ability	Talent as mastery (of skills and knowledge)	Competence
Motivation	Talent as commitment	Commitment
Opportunity	Talent as fit (being in the right position at the right time)	Contribution

In addition, it is interesting to stress that in the "talent equation, the three terms are multiplicative, not additive. If anyone is missing, the other two will not replace it" (Ulrich and Smallwood, 2012: p. 60). In contrast with the CS initiatives focused on improving skills, Table 2.3 sheds light on other perspectives to create talent, such as creating opportunities for individuals to show their talents by proper placement and improving motivation.

2.2.3 Challenges for Managing Talents

Since McKinsey consultants created the expression "War for Talent" in 1997 (Michaels et al., 2001), Talent Management (TM) has received a remarkable degree of interest among practitioners and researchers and is one of the most debated subjects in HRM (Collings and Mellahi, 2009). An adequate TM is the key to organisational success (Gallardo-Gallardo et al., 2013) and, according to the approach in this thesis, it is crucial for the CS community.

There are some reasons for the growing importance of Talent Management (TM). First of all, TM is needed in a fast changing business environment (Hor et al., 2010; Lewis and Heckman, 2006) and has a positive impact in organisational performance (Gelens et al., 2013; Hiltrop, 1999; Hughes and Rog, 2008). Lewis and Heckman (2006) argue that change management is required to support decision-making. Second, combining traditional HRM practices with strategic decision-making increases organisational performance (Huselid, 1995; Schalk et al., 2013). Finally, talented people are responsible to provide a competitive advantage to organisations (Collings and Mellahi, 2009; Farndale et al., 2014; Hor et al., 2010; Iles, 2013; Tarique and Schuler, 2010; Thunnissen et al., 2013). Organisations in general, however, are facing a critical shortage of talented personnel (Deloitte, 2011; Festing and Schäfer, 2014; Gelens et al., 2013; Hiltrop, 1999; Hughes and Rog, 2008; Michaels et al., 2001).

In contrast to the aforementioned arguments, Talent Management (TM) is not immune to criticism. Some authors warn that TM is resource consuming (Collings and Mellahi, 2009), eventually might become the same as HRM (Lewis and Heckman, 2006) and normally adopts an unitarist and economic-oriented approach. The unitarist approach means that the organisation is perceived as a unified actor without considering individual needs (Thunnissen et al., 2013). By adopting a pluralistic approach, however, TM can provide not only mutual benefits for organisations and individuals (Farndale et al., 2014), but even for society, as presented in Table 2.4. It seems that the CS community permeates these three levels.

Table 2.4: Pluralistic approach to talent management (adapted from Tansley et al., 2013: p. 338; Thunnissen et al., 2013: p. 331)

Value	Individual (micro level)	Organisation (meso level)	Society (macro level)
Economic	Financial rewards; job security	Profitability; flexibility; efficiency and effectiveness	Economic condition and competitive position of industry, region or country
Non-economic	Meaningful work; growth and social needs	Legitimacy	Social responsibility (social / moral development)

Worldwide senior managers argue that major challenges in Talent Management involve career paths, leadership, compensation, succession planning, attracting, retaining, training and recruiting talent (Deloitte, 2011).

2.2.4 Practices and Requirements for Managing Talent

Managing talents requires an architecture that comprises a set of HR practices (Lewis and Heckman, 2006) and literature has already structured such practices into frameworks and processes (e.g. Altınöz et al., 2013; Collings and Mellahi, 2009; Festing and Schäfer, 2014; Gümüş et al., 2013; Lewis and Heckman, 2006; Schiemann, 2014). Combining the aforementioned approaches results in some requirements and general steps for managing talents, which are discussed in the following subsections. Requirements involve defining strategies, creating profiles, designing a differentiated architecture and determining performance assessment. Those requirements deliver valuable signals to address the expertise discovery within the CS domain.

2.2.4.1 Defining the Strategy, Key Positions and Profiles

The first step to developing Talent Management (TM) is defining the targets and strategies (Altınöz et al., 2013; Gümüş et al., 2013). The strategy should provide sustainable competitive advantage (e.g. identifying relevant organisational resources) and define the implications for talent (Lewis and Heckman, 2006). When establishing the strategy, special care should be taken with regard to the interactions between the organisations and talent. With regard to the sustainable aspect of TM, Schiemann (2014) depicts the main interactions and talent practices between the organisation and talent as an iterative talent lifecycle (Figure 2.5).

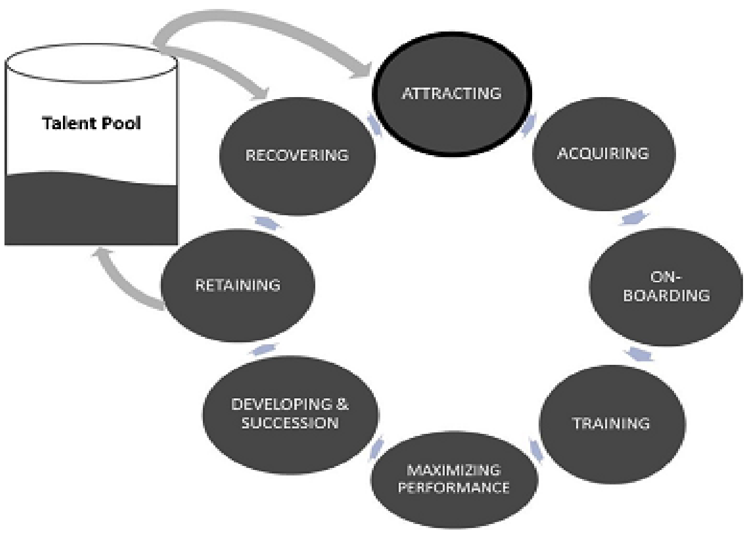


Figure 2.5: Talent lifecycle (Schiemann, 2014: p. 282)

Based on their own definition of strategic Talent Management (TM), Collings and Mellahi (2009) propose a framework (Figure 2.6) comprising four major steps. The first step is to identify the key positions that affect organisational performance. Developing a talent pool that combines internal development and external recruiting is the second step. The third step is creating a differentiated HR architecture, by adopting selected HR practices and developing commitment. Finally, the outcomes depend on individual performance, which is a function of ability, motivation and opportunity to perform (Collings and Mellahi, 2009). The private sector has been already developing their corporate solutions for integrated TM[5].

Key positions that may have the potential to gain a competitive advantage for the organisation and impact outcomes should be identified even before talent is determined (Altınöz et al., 2013; Collings and Mellahi, 2009; Gallardo-Gallardo et al., 2013; Gümüş et al., 2013). This thesis focuses in determining such key positions as CS projects.

After defining the key positions, talent profiles should be designed containing the required skills, knowledge, experience and other qualities for key positions (Altınöz et al., 2013). This becomes particularly important when selecting individuals to occupy these roles (Thunnissen et al., 2013). After all, one of the major challenges in Talent Management is allocating the right people at the right place at the right time (Tarique and Schuler, 2010; Ulrich and Smallwood, 2012). Profiles are further discussed in Chapter 3.

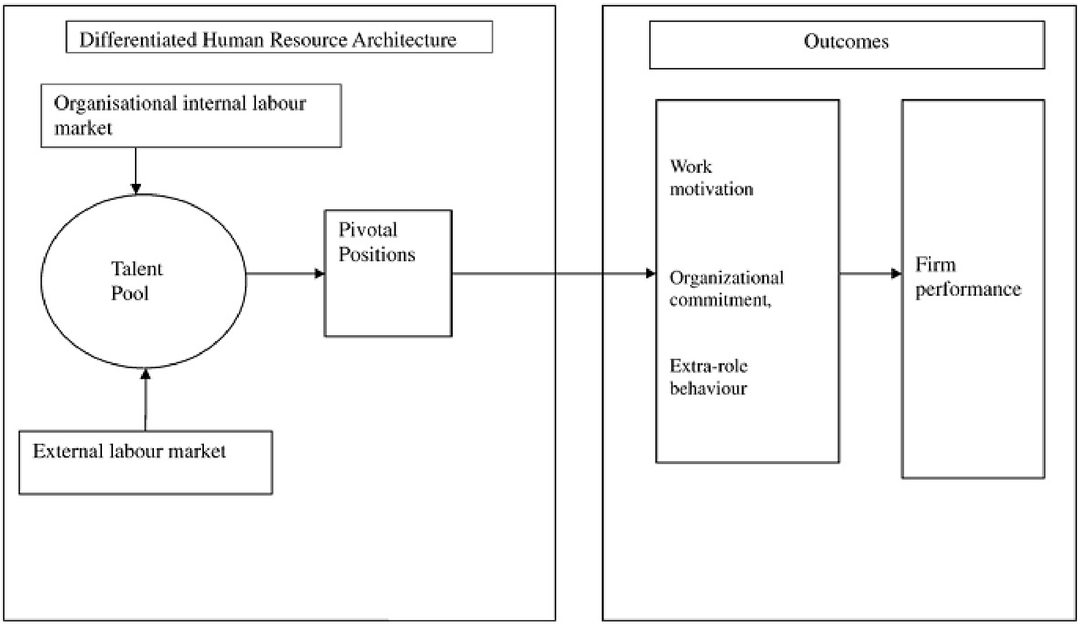


Figure 2.6: Strategic talent management framework (Collings and Mellahi, 2009: p. 306)

[5] A list of some products is available in <http://www.fosway.com/9-grid/talent-management/>

2.2.4.2 Talent Pool and Skills Gap Analysis

In order to fill key positions, the development of a talent pool with high potential and high performing individuals is required (Altınöz et al., 2013; Collings and Mellahi, 2009). However, Ulrich and Smallwood (2012) argue that the talent culture should reach all employees.

Developing a talent pool involves mapping incumbents' skills (Collings and Mellahi, 2009). The first step is gaining a deep understanding of the internal workforce (Lewis and Heckman, 2006). Filling the talent pool from within as much as possible provides several benefits (e.g. commitment, opportunities and lower costs); however, external recruitment is advisable as a second step. Under the individual perspective, protean careers support the search for self-fulfilment to incumbents themselves (Hiltrop, 1999). Instead of solely relying on organisational careers, investing on boundaryless careers adds flexibility to the talent supply (under the organisational perspective) (Collings and Mellahi, 2009). Therefore, talent pools should combine developing internal personnel with external recruitment. Nevertheless, it is worth noting that agencies should find the proper balance between spending effort in talent identification versus talent development (Altınöz et al., 2013; Gallardo-Gallardo et al., 2013). Both approaches are taken into account for the research solution.

Comparing the qualities of the individual already occupying a key position against the respective talent profile helps to identify shortcomings. These shortcomings can be addressed by applying HR practices, such as providing complementary training (Altınöz et al., 2013). It is important to stress that even after properly selecting from a talent pool, incumbents still need lifelong training to be compliant with the changing requirements of key positions (Thunnissen et al., 2013). The research solution provides feedback for expertise improvement to the actors involved in the selection.

2.2.4.3 Differentiated HR Architecture and Talent Management Practices

A differentiated HR architecture is required to add some features to the traditional HR practices and improve organisational performance and commitment, which are clearly objectives of Talent Management (Collings and Mellahi, 2009). Hence, one of the key aspects to the successful deployment of a strategy involves selecting the proper talent practices (Lewis and Heckman, 2006). Some of the HR practices for the architecture are presented in Table 2.5. This research project "differs" from traditional HR architecture because it considers how selected practices affect not just talent, but also other stakeholders involved in CS expertise discovery as well.

Table 2.5: HR practices for Talent Management

HR practice	References
Identification	(Adobor, 2004; Collings and Mellahi, 2009; Gelens et al., 2013; Hor et al., 2010; Tarique and Schuler, 2010)
Attract	(Festing and Schäfer, 2014; Gelens et al., 2013; Hiltrop, 1999; Michaels et al., 2001; Schiemann, 2014; Tarique and Schuler, 2010; Thunnissen et al., 2013)
Select	(Adobor, 2004; Collings and Mellahi, 2009; Festing and Schäfer, 2014; Gallardo-Gallardo et al., 2013; Gelens et al., 2013; Hor et al., 2010; Lewis and Heckman, 2006; Tarique and Schuler, 2010; Thunnissen et al., 2013)
Reputation management	(Tarique and Schuler, 2010)
Recruitment	(Collings and Mellahi, 2009; Deloitte, 2011; Gallardo-Gallardo et al., 2013; Hiltrop, 1999; Hughes and Rog, 2008; Lewis and Heckman, 2006; Tarique and Schuler, 2010; Thunnissen et al., 2013)
Develop	(Adobor, 2004; Collings and Mellahi, 2009; Festing and Schäfer, 2014; Gallardo-Gallardo et al., 2013; Gelens et al., 2013; Hiltrop, 1999; Hor et al., 2010; Hughes and Rog, 2008; Iles, 2013; Lewis and Heckman, 2006; Michaels et al., 2001; Schiemann, 2014; Tarique and Schuler, 2010; Thunnissen et al., 2013; Ulrich and Smallwood, 2012)
Train	(Adobor, 2004; Deloitte, 2011; Gallardo-Gallardo et al., 2013; Hiltrop, 1999; Hor et al., 2010; Lewis and Heckman, 2006; Schiemann, 2014; Tarique and Schuler, 2010; Thunnissen et al., 2013; Ulrich and Smallwood, 2012)
Retain	(Collings and Mellahi, 2009; Deloitte, 2011; Festing and Schäfer, 2014; Gallardo-Gallardo et al., 2013; Gelens et al., 2013; Hiltrop, 1999; Hughes and Rog, 2008; Michaels et al., 2001; Schiemann, 2014; Tarique and Schuler, 2010; Thunnissen et al., 2013)
Career	(Collings and Mellahi, 2009; Deloitte, 2011; Farndale et al., 2014; Gubler, 2011; Hiltrop, 1999; Hor et al., 2010; Hughes and Rog, 2008; Lepak and Snell, 2002; Lewis and Timlin, 2011; Tarique and Schuler, 2010; Thunnissen et al., 2013)
Performance management	(Altınöz et al., 2013; Lewis and Heckman, 2006; Schiemann, 2014; Tarique and Schuler, 2010)
Succession planning	(Collings and Mellahi, 2009; Deloitte, 2011; Festing and Schäfer, 2014; Hor et al., 2010; Lewis and Heckman, 2006; Schiemann, 2014; Thunnissen et al., 2013; Ulrich and Smallwood, 2012)
Compensation	(Deloitte, 2011; Lepak and Snell, 2002; Lewis and Heckman, 2006; Tarique and Schuler, 2010)
Analysis of talented worker gap	(Altınöz et al., 2013)

Although *attracting* talent is still a challenge, research tells us that people are more attracted to larger firms, training and development opportunities, teamwork, participation and autonomy. Other attraction factors include promotion, career development, clear intentions and outcomes, employment security, equal benefits and opportunities, proactive HRM and performance rewards (Hiltrop, 1999).

Tarique and Schuler (2010) include *reputation management* as a challenge to Talent

Management (TM). In fact, *trust* is one of the top drivers of commitment (Hughes and Rog, 2008) and should be developed through TM practices (Altınöz et al., 2013; Boselie et al., 2005). It is particularly required when selecting talent to manage partnerships between organisations (Adobor, 2004) and within the CS community.

When using the talent pool strategy to *select* personnel, the organisation "recruits the *best* people and then selects them for positions rather than trying to select *specific* people for *specific* positions" Tarique and Schuler (2010: p. 127). Collings and Mellahi (2009) also recognise that in terms of fit, it is the position and not the employee that dictates the requirements. This research relies on a set of expertise requirements describing a CS project, which may relate to a job position or a temporary task. However, selecting people on the scale of a CS community is a challenging task, which may require proper techniques.

Hughes and Rog (2008) argue that the organisational effort in the *development* of employees' skills and the extent of their participation are also seen as top drivers of commitment. The unequal allocation of training and skills development opportunities can cause positive effects on the high potentials and high performers and negative reactions on "non-high potentials". Because the negative reactions of the latter may outweigh positive effects on the former, the opportunities of developing skills and participation on training should adopt an inclusive approach (Gelens et al., 2013). *Compensation* practices support the motivation and commitment of individuals (Lepak and Snell, 2002). Finally, an analysis of the talented work gap leads to designing plans to improve the overall development of talents considering individual needs (Altınöz et al., 2013). The research solution encompasses a mechanism to benchmark such an expertise gap.

2.2.4.4 Assessment of the Candidates' Performance

Altınöz et al. (2013) argue that performance has to be benchmarked for evaluation purposes in order to support decision-making. In this thesis, the performance of candidates with regard to given competencies is benchmarked, determining experience, which supports selecting candidates to work on CS projects.

Performance variables, such as candidates' abilities, should be predetermined. The problem for such assessment, however, is that the exact relationship between variables that affect performance is not known (Collings and Mellahi, 2009). Measurement is important in order to adhere to a rigorous, science-based approach. Using standard metrics in a different way, instead of creating new ones, might prove to be a useful solution (Lewis and

Heckman, 2006). Hence, this thesis adapts an existing evaluation method to perform assessment of candidates' performance. Such evaluation considers perspectives of different actors in order to provide a more comprehensive coverage of candidates' behaviour in the CS community as a whole. Therefore, the assessment of candidates is not restricted to a particular employer, but rather shapes the collective perception of each candidate's expertise, which seems more aligned with a pluralistic approach.

Although organisations are already developing their own metrics and performance indicators (Deloitte, 2011), no literature has been found to discuss a scale to measure Talent Management practices (Festing and Schäfer, 2014) or feedback of such metrics to decision makers (Lewis and Heckman, 2006). However, recent studies use techniques such as Data Envelopment Analysis (DEA) or Analytic Hierarchy Process (AHP) to support decision-making for talent practices and performance (e.g. Feng et al., 2004; Lin, 2010; Tseng and Lee, 2009). This thesis further adopts measures to benchmark not only individuals, but also organisational (i.e. the CS community) performance in a given set of capabilities in order to improve expertise discovery.

2.2.4.5 Talent Retaining

Retaining talent in the context of this thesis is not about keeping an expert associated to a particular agency, but is about retaining as much expertise as possible for the CS community as a whole. This practice may require an intricate combination of factors rather than just raising payment. Investments on high performance work practices impact on employee retainment along with increasing organisational productivity and performance (Huselid, 1995). HR practices that support retaining talents are basically the same as those used to attract them (Hiltrop, 1999).

Increasing employee engagement (Tarique and Schuler, 2010) and motivation contributes immensely towards talent retention. Some researchers have reasoned that talented employees were willing to stay with their agencies and even give something in return to compensate for the investment and the trust that they have received. Therefore, it is important to manage the expectations of employees (Festing and Schäfer, 2014). Otherwise, delivering a flawed perception of organisational justice may lead to negative reactions by the employees (Gelens et al., 2013). Factors that particularly contribute to retain CS experts were investigated in a pilot questionnaire and agree with the literature. For example, skill development and courses along with an efficient human resource planning were deemed as the top drivers to retain CS experts. Other drivers include equal opportunities and feedback. This thesis considers those talent-retaining drivers when devising the research solution (as further described in Section 5.1.2).

Losing a valuable employee is not a definitive situation. *Talent recovery* is being implemented in some companies in order to recover lost talent or to use their network to attract new talent based on a good employer-employee relationship (Schiemann, 2014). Although such practice is part of the talent lifecycle and, therefore, requires at least a brief introduction, it is not crucial to this thesis since the research solution concentrates on mapped and up-to-date expertise and not on particular individuals. Once experts are mapped, however, the research solution delivers a mechanism for employers to recommend their former employees so that those CS talents can be recovered over time based on those evaluations.

2.2.5 Adapting Talent Practices for the CS Expertise

The findings in the Talent Management (TM) literature provide support and detail some features and requirements to the framework presented in Figure 2.7, which highlight the CS viewpoints for discovering expertise (cf. Figure 2.2). Although most issues are discussed in the TM literature (i.e. those in the solid shapes), other talent features require deeper analysis from other fields for effective implementation (i.e. those in the dashed shape). The stakeholders and their inputs and outputs described in Figure 2.2 are represented as the three leftmost shapes in Figure 2.7, whereas discussions in this section concentrate on the TM viewpoints for expertise discovery.

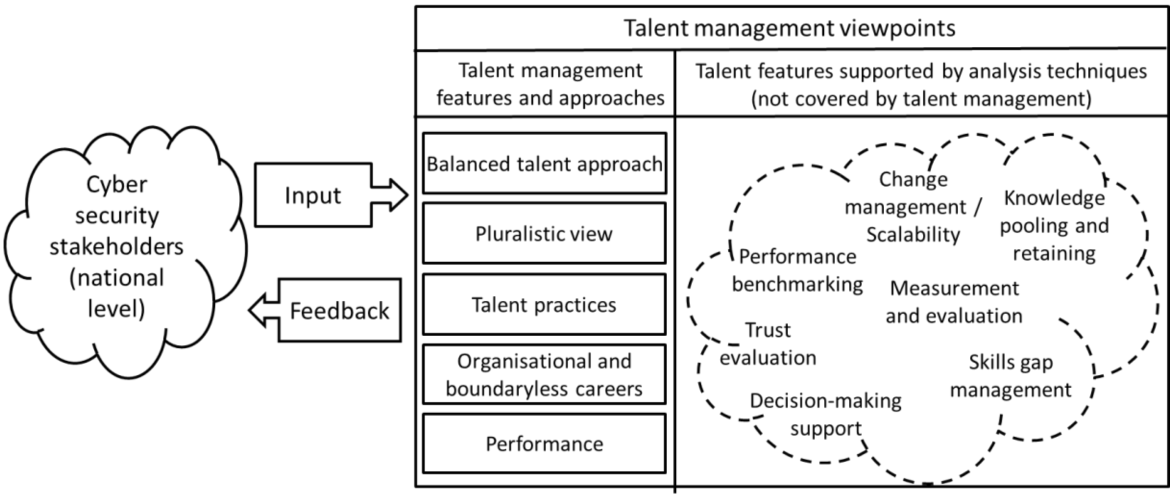


Figure 2.7: Talent management supporting cyber security

Since CS and strategic Talent Management (TM) are quite recent fields of study (less than 10 and 20 years, respectively) (Klimburg, 2012; Michaels et al., 2001), there are not many academic studies combining both areas. TM literature is rather focused on concepts and practices, leaving a gap for contextualised applications. For example, although performance is a crucial output of TM (Collings and Mellahi, 2009) for benchmarking expertise, the literature lacks a proper scale to measure such TM practice (Festing and Schäfer, 2014). Moreover, an investigation into strategic TM reveals limited applied works between organisations (Thunnissen et al., 2013). It seems, however, that the term talent and its related practices are appropriate to address the CS expertise discovery issue. Such shortcomings became some of the reasons why conceptualising talent practices in the CS domain has not yet been properly discussed, leaving a methodological gap. Hence, this work seeks to structure and underpin the discovery of CS expertise by sharing experts and talent practices across organisations. The adopted TM approaches are hereby discussed.

Talent Management (TM) relies on a differentiated architecture comprising internal and external talent pools and mapping skilled personnel and critical positions. Such positions or temporary contracts are hereby described through projects so that expertise requirements may be shared throughout the CS community. Individuals holding the desired capabilities can then be identified and selected to perform specific tasks. Finding the proper workforce, retaining and leveraging required skills are the core capabilities provided by TM to support the CS effort. Thus, some TM practices, processes and requirements discussed in the previous section (e.g. talent pool, selection, performance evaluation and analysis of worker gap) are adopted to underpin the solution design in this work. Aligned with the definitions in (Gallardo-Gallardo et al., 2013), the research solution concentrates on talent as “fit” and “commitment” to complement the CS initiatives, which are currently oriented to simply “master the skills”. The vast and dynamic CS knowledge environment facilitates the adoption of a balanced approach for talents where those highly reputed have the best chances to be selected for a project. On the other hand, our approach also aims to identify trending subjects, which can become opportunities for discovery of new high potentials.

Talent Management (TM) is a resource consuming activity deployed in a strategic (Collings and Mellahi, 2009) and even global level (Deloitte, 2011; Farndale et al., 2014; Tarique and Schuler, 2010). However, apart from some discussions about joint ventures (Adobor, 2004) and employment modes (Lepak and Snell, 2002), there is a lack of literature regarding partnerships among organisations for sharing TM resources. This occurs because partnerships seem to contrast with competitive environments with scarce talent assets. In addition, the average one-dimensional and economic oriented approach to TM focuses on multinational and private organisations, leaving public organisations, small enterprises and individuals out of the TM scope as indicated by Thunnissen et al. (2013). Some of the emerging challenges for the adopted pluralistic approach are sharing evaluations of those professionals to promote their own boundaryless/protean careers and how can organisations benefit from such knowledge exchange.

Because Talent Management (TM) is resource consuming (Collings and Mellahi, 2009), it is believed that inter-organisational TM can save expertise resources. In fact, sharing unique experts through alliances and partnerships can help organisations expand the talent pool without compromising competitive advantage and avoiding additional costs of internal employment (Lepak and Snell, 2002). We realise, however, that sharing expertise within the CS community should be encouraged under perspectives of the players involved. This can be done by understanding the needs of those players and developing processes under the understanding of selected talent practices. For example, in compliance with the adopted pluralistic approach, it is assumed that sharing a talent pool could benefit single individuals by increasing the opportunities for the most talented. In order to address both pivotal positions (Collings and Mellahi, 2009) as well as temporary contracts (Tansley et al., 2013; Thunnissen et al., 2013), this thesis henceforth adopts the term “project” comprising a set of expertise requirements demanded by an organisation.

Among several requirements (cf. Section 2.2.4) and talent practices identified in the literature (cf. Table 2.5), this research work embeds those that seem most relevant to the adopted pluralistic approach for the CS community as a whole by adapting the talent lifecycle introduced by Schiemann (2014). Therefore, in this thesis, a CS expertise discovery lifecycle is analysed according to each stakeholder’s perspective regarding the research problem instead of solely concentrating on talent. For such reasons, practices that are the concern of a single stakeholder (e.g. succession planning) are beyond the scope of the solution.

Inter-organisational Talent Management (TM) addresses some criticisms the reports by Lewis and Heckman (2006) as well as Tansley et al. (2013) and suits the discovery of CS experts for several reasons. First, TM becomes more complex than just a limited set of HR practices. Second, it increases the overall benefits (cf. Table 2.4) for multiple stakeholder levels (i.e. in the CS environment), by adopting a systemic approach rather than a unitarist one. Finally, TM becomes oriented by a wider set of purposes rather than just being economic driven. In addition, authors request evidence of how TM can be applied in a different context (Gallardo-Gallardo et al., 2013). In order to support such an approach for discovering expertise in the CS context, organisational modelling theories are further discussed.

Talent Management (TM) alone cannot properly address some remaining issues for expertise discovery. First, knowledge is one of the major building blocks of talent (cf. Table 2.2 and Figure 2.3) that cannot be retained within organisations using typical TM practices (cf. Table 2.5). Second, TM does not provide the mechanisms for change management (Lewis and Heckman, 2006), to support decision-making on expertise, to identify the skills gap and to deliver tailored knowledge for training individuals, let alone on the scale of a CS community. Third, besides pooling people, CS also requires pooling knowledge (Klimburg, 2012), which is also not supported by typical TM features. Finally, the TM literature does not provide reusable measurement mechanisms for evaluating effectiveness on the scale of a CS community. The following section delivers a theoretical foundation to address these shortcomings and support the solution to this research work.

2.3 Knowledge and Knowledge Management

The shortage of talented people (i.e. experts) is a common concern for both CS (Caldwell, 2013; Kouttis, 2016) and Talent Management fields (Deloitte, 2011; Festing and Schäfer, 2014; Gelens et al., 2013). There is already a great supply of talent, but no effective mechanisms to discover those talents have been found in the literature. Discovering talent is critical to supporting the CS joint effort, although there is still a risk of losing those skilled people and their knowledge to turnover. Notwithstanding this, knowledge is one of the core assets of talent (Cheese, 2007; Gagne, 2000; Michaels et al., 2001; Tansley et al., 2006; Ulrich and Smallwood, 2012) that can be created, retained and shared through proper Knowledge Management (KM) (Alavi and Leidner, 2001). Building on the contextual relevance of KM, the CS environment can be understood with adequate information exchange and knowledge pooling (German Federal Ministry of the Interior, 2011; Klimburg, 2012). In this section, some concepts related to knowledge and the applications of KM are introduced for the purpose of shedding some light on requirements for discovering the CS expertise.

2.3.1 Basic Concepts related to Knowledge

Knowledge is “the result of cognitive processing triggered by the inflow of new stimuli” (Alavi and Leidner, 2001: p. 109) and can be “viewed as a mix of information, understanding, capability, experience, skills and values" (Rowley, 2007: p. 174). According to these approaches, information becomes knowledge while in a person's mind, and knowledge becomes information when it is written or represented in some type of media (Alavi and Leidner, 2001). When experience promotes a deep knowledge of a subject, the person holding such knowledge is referred to as an "expert" (Davenport and Prusak, 2000). Likewise, this thesis refers to such a kind of knowledge as "expertise".

Scholars have increasingly recognised that information and knowledge are not interchangeable concepts (Nonaka et al., 2006), although the latter is much more complex to define. Table 2.6 summarises the main differences between knowledge and other related concepts using the data-information-knowledge-wisdom hierarchy with regard to other features discussed in Rowley (2007).

Table 2.6: The wisdom hierarchy mapping to types of systems, understanding and related features (adapted from Rowley, 2007: p. 167–178)

Wisdom hierarchy	System	Understanding	Features
Wisdom	Expert system	Know why	Evaluated understanding
Knowledge	Decision support system	Know how	Mix of information, understanding, capability, experience, skills and values
Information	Management information system	Know what	Explicit knowledge, structured and processed data
Data	Transaction processing system	Know nothing	Unorganised and unprocessed items

The hierarchy implies that when moving from the top (i.e. wisdom) towards the bottom concept (i.e. data), there is an increase in programmability in contrast to a decrease in meaning and value, and vice-versa. Because knowledge can be seen as "actionable information" (Rowley, 2007), knowledge-based systems support decision-making and are built upon management information systems (cf. Table 2.6). The UK Academy for Information Systems defines information systems (IS) as the following:

> *"(...) the means by which people and organisations, utilising technologies, gather, process, store, use and disseminate information. (...) The domain involves the study of theories and practices related to the social and technological phenomena, which determine the development, use and effects of information systems in organisations and society"* (UKAIS, 2014)

While management IS improve efficiency in information retrieval as well as in standard decision rules, decision support systems (DSS) extend support for managers to reason upon ill-structured problems and improve judgements. Hence, DSS provides flexibility to accommodate unexpected changes by combining analytic techniques with IS (Finlay, 1989). The authors agree that organisational knowledge can be tacit or explicit (Alavi and

Leidner, 2001; Rowley, 2007). Tacit knowledge is based on skills (Lee and Lan, 2011) and improved through personal experience (i.e. expertise). If tacit knowledge is not codified or shared, it is lost when skilled individuals leave their organisation (Teo, 2005). However, tacit knowledge can be learned through imitation or shared by using metaphors and analogies. Explicit knowledge is rule-based when codified into routines or procedures or object-based when formally expressed using artefacts (e.g. products, tools, databases) or symbols (Choo, 2000).

In addition to the common sense tacit/explicit classification of knowledge, the literature provides some other common perceptions to knowledge. These perceptions influence the Knowledge Management approach, which impacts the development of Knowledge Management Systems (KMS) and the role of ICT (Alavi and Leidner, 2001), as presented in Table 2.7.

Table 2.7: Perceptions of knowledge and implications for knowledge management systems (adapted from Alavi and Leidner, 2001: p. 111)

Knowledge perception	**KM approaches**	**Implications for KMS (role of ICT)**
Personalised information (or interpreted data)	Providing useful information and the means to assimilate it	Not significantly different from existing IS, but improves user's assimilation
State of mind (of knowing and understanding)	Learning / understanding through information provision	KMS provide access to sources of knowledge instead of knowledge itself
Object (can be stored and manipulated)	Building and managing knowledge stocks	Creating, storing and managing knowledge
Process (of applying expertise)	Developing expertise through knowledge processes (i.e. creation, sharing and distribution)	Linking knowledge sources and distribution to increase knowledge flow
A condition of access to information	Organising access and retrieval to content	Provide search and retrieval mechanisms to useful information
Capability (to influence action)	Potential to influence action	Building competencies, skills and managing intellectual capital

2.3.2 The Importance of Discovering Knowledge

Similar to Talent Management, effective Knowledge Management is crucial to developing a sustainable competitive advantage (Alavi and Leidner, 2001; Savvas and Bassiliades, 2009; Teo, 2005) and organisational performance (Kamal, 2011; Nonaka et al., 2006) in a dynamic environment (Lee and Lan, 2011; Owen, 1999; Sourouni et al., 2010).

Business processes are shifting from manual labour to knowledge work, which demands higher levels of knowledge and expertise (Owen, 1999). Therefore, keeping organisational knowledge up-to-date and mapping expertise constitute current organisational challenges, which are supported by Knowledge Management (KM) applications. For example, creating corporate directories (Alavi and Leidner, 2001) or CVs that are updated by the employees themselves (Choo, 2000) address some of these issues. Effective Knowledge Management systems improve decision-making (Wong and Aspinwall, 2006), which is required for talent selecting practices. Among other benefits, KM enhances core competencies, increases intellectual capital (Teo, 2005) and supports ongoing learning (Choo, 2000).

Knowledge Management (KM) is a twofold activity: 1) developing a KM solution through organisational modelling and 2) maintaining such solution through its own knowledge processes (Staab and Studer, 2004). With regard to the former activity, KM requires a system that supports creation, storage, retrieval, transfer and application of knowledge (Alavi and Leidner, 2001; Butler et al., 2008). Hence, ICT plays a fundamental role in supporting KM applications (Owen, 1999), such as communities of practice, discussion forums and user training (Butler et al., 2008). This thesis concentrates on delivering a solution for CS expertise discovery and sheds light on the meta-processes to maintain the research solution. Knowledge creation comprises four modes developed in specific common spaces or "ba" as presented on Table 2.8. These environments involve informal talks, knowledge cafes and online forums (Choo, 2000). It is worth noting that tacit knowledge is also created when learning from applying knowledge (Alavi and Leidner, 2001).

Table 2.8: Knowledge creation modes and environments (adapted from Alavi and Leidner, 2001: p. 116; Nonaka et al., 2006: p. 1185)

Mode	**Type of knowledge creation**	**"ba" (common space for creation)**
Externalisation	Tacit knowledge generates explicit knowledge (e.g. lessons learned)	Interacting ba (dialogue and collaboration in workgroups; skills and mental models are codified)
Internalisation	Explicit knowledge generates tacit knowledge (e.g. continuous learning and training)	Exercising ba (information provision[6] for lifelong learning)
Socialisation	Tacit knowledge generates tacit knowledge (through socialising)	Originating ba (face-to-face interactions, share experiences, feelings and mental models)
Combination	Explicit knowledge generates explicit knowledge (e.g. literature review)	Cyber ba (virtual space, use of data warehousing to search related knowledge)

The externalisation mode of knowledge creation is a difficult task because explicit knowledge may not be capable of capturing the richness and complexity of tacit knowledge accrued over a long period. Hence, an alternative to knowledge externalisation

[6] Examples of KM architectures providing internalisation are in Sun et al. (2010) and Sun and Mushi (2010)

involves locating the person with the right tacit knowledge (Davenport and Prusak, 2000), which is the main knowledge creation mode adopted in this thesis. The remaining knowledge creation modes drive further discussions on profiles, social media and performance measures in the following chapter. Additional challenges posed to Knowledge Management include proper information delivery (Teo, 2005), which refers to proper feedbacks and trust (Alavi and Leidner, 2001). Since expertise is a type of tacit knowledge, this thesis assumes that it can be discovered by applying the knowledge processes. Due to the difficulty of codifying the expertise, discovering the experts is crucial to the research solution.

2.4 Organisational Modelling Theories

Now that the processes for discovering knowledge have been presented, this section introduces the theories through which those processes and the CS community can be modelled. The complex CS environment involves a wide range of organisations and people exchanging knowledge, providing services and establishing partnerships (Klimburg, 2012). However, a partnership itself can be viewed as a single organisation, because it still comprises stakeholders, goals and their activities (Harris, 2011). In essence, organisations are simply information systems, which develop their products and services through business processes (Liu et al., 2003). Modelling organisations and their business processes enables a better understanding of organisational activities and knowledge processes and improve automation (Sani, 2011; Stamper et al., 2000). This research portrays the CS community as an overarching organisation with its own set of stakeholders, actors and processes. This section introduces some approaches for organisational modelling theories as a means to pave the theoretical foundations on which the CS community can be modelled. Initially, communities of innovation, requirements engineering and methods to identify the stakeholders are presented. Then, the service-oriented theory is briefly introduced, followed by a discussion on organisational semiotics. Due to its relevance for CS, information and knowledge reliability issues are addressed through the discussion of trust. Finally, some viewpoints on organisational modelling for CS expertise discovery are discussed.

2.4.1 From Communities of Practice to Business Ecosystems

CS communities normally rely on PPP for information sharing (Klimburg, 2012) and enable an environment for knowledge creation, such as RENASIC (Brazilian Government, 2013b). Findings about CS expertise discovery seem to resemble such a collaborative environment (cf. Figure 2.2) as an ecosystem. Hence, this section briefly discusses CS communities under the perspective of communities of innovation and business ecosystems in order to understand their possibilities and to draw viewpoints for the research solution.

Communities of practice (CoP) are self-organised and inter-organisational groups comprised of individuals that share common or complementary interests and goals (Brown and Duguid, 1991). CoP are important to developing group knowledge by enabling borderless interactions among people with similar interests (Alavi and Leidner, 2001). Innovation is created as a reflection on the interactions between a single organisation and its environment, which eventually leads to change (Brown and Duguid, 1991), and is a synonym to discovery (Oxford Dictionaries, 2009). When knowledge is created and shared across organisations of different natures, CoP may evolve into “communities of innovation”. Strategic partnerships among governments, universities and private sector can merge distinct knowledge pools and CoP into a richer knowledge environment. Such kinds of partnerships play an increasing role in national innovation systems (Carayannis et al., 2000). A prominent example of a community of innovation is the Linux project.

In his seminal work “Out of Control”, Kelly (1994) asserts: “as we make our machines and institutions more complex, we have to make them more biological in order to manage them”. In line with such an assertion, business ecosystems comprise aspects which are crucial to CS communities in the context of this work. Business ecosystems, coined by Moore in 1993, draw core ideas from natural ecosystems, which focus on adaptive, symbiotic and mutualistic relations. This thesis considers such relationships in terms of how members of the CS community provide each other expertise assets and services. Since no one organisation has all the required resources for survival and development, the concept of co-evolution emerges as collaboration and innovation beyond organisational borders. In fact, co-evolution lies in the core of complex adaptive systems and has been a matter of discussion in business ecosystems. In contrast to traditional hierarchical organisations, the literature has long recognised the application of ecosystems in business and economics, covering a spectrum of abstract visions to pragmatic approaches. The difficulty of managing complex business ecosystems can be overcome through articulated, structured and analytical processes. Some of the critical challenges for business ecosystems include an efficient coordination role to attract participants, ensure strategy alignment, support complementary capabilities, enable feedbacks, establish links for contributions and encourage innovation. Despite existing coordination roles, all ecosystem members share the stewardship of the community at different levels of responsibility. Business ecosystems comprise a diverse set of contributors exchanging different types of resources (e.g. services, knowledge and management) on a varied scale. Those exchanged resources benefit the stakeholders in a specific way and promote the growth of the ecosystem as a whole (Moore, 2006). Automobile companies along with the companies to which fabrication of car components is outsourced generally constitute a business ecosystem. The use of the term ecosystem associated to innovation has been criticised as a meme in the economic field, as a surrogate for "technopolis" and "innovation systems", as well as a faulty adaptation of natural ecosystems. The terminology is, however, still compelling and valid as a metaphor. The ideas similarly inspire innovation practices (Oh et al., 2016), some of which are adopted in this thesis.

In this research work, business ecosystem and community of innovation theories do not conflict, but rather inspire and drive the application of the other organisational modelling theories. This work recognises national CS as a developing community of innovation, with regard to the collaborative effort of the initiatives presented in Section 2.1. These initiatives, however, are normally originated in only a few proactively engaged CS stakeholders. The CS community is also moving towards the concept of a business ecosystem where a central CS authority coordinates and shares the co-evolution of its members (cf. Figure 2.2). Thus far, some business ecosystem challenges have been addressed in some countries (e.g. defining coordinating agencies, creating opportunities to attract talents and ensuring CS strategy alignment through frameworks). In contrast, other challenges are also adopted in this thesis, since they affect the research problem with regard to unstructured and empirical expertise discovery. These challenges include modelling capabilities, enabling innovation and feedbacks, establishing links among stakeholders and enabling joint stewardship.

2.4.2 Requirements Engineering for the CS Community

The CS community has a recurring set of stakeholders. However, there are specific roles performed by these stakeholders, business ecosystems challenges to address and information needs affecting CS expertise discovery, which require further analysis. Organisations (which include the CS Community) are generally the main users and providers of Knowledge Management (Alavi and Leidner, 2001). Prior to discovering the expertise, the concepts that underlie the CS community and the requirements that drive such organisation should be determined. Requirements engineering, as a process to elicit, analyse and document requirements, plays a major role in identifying and understanding the concepts related to expertise discovery within the CS community. These concepts include stakeholders, their roles and their informational needs (i.e. requirements). Requirements engineering supports both types of KM activities, i.e. determining requirements to develop a solution for expertise discovery and maintaining these requirements over time.

Developing organisational Knowledge Management starts with a feasibility study to identify the problem and the stakeholders involved (Staab and Studer, 2004), which in this thesis were outlined with findings in literature combined with a pilot questionnaire. There are numerous stakeholder identification approaches, such as the Stakeholder framework (Freeman et al., 2007) and the Power/Impact Matrix (BIS, 2010). Amongst those approaches, Liu et al. (2006) present a method for identification and analysis of stakeholders according to six possible roles called the stakeholder onion (Figure 2.8). Each role is driven by a set of norms describing their behaviour. The impact of each role in the problem domain is determined by their closeness (e.g. clients have a greater influence than facilitators do). In this model, the actor is able to perform activities, which directly affect the problem domain. Such activities are guided by norms and involve outputs and changing states of instances. The client is the beneficiary of the system outcomes. The provider is the role that controls the conditions or the resources to support the system. The facilitator ensures continuity by solving conflicts and driving the group to achieve organisational goals. The governing body establishes the strategies, objectives and monitors the group's progress. In contrast to the former role, the latter does not interact directly with the group. Bystanders influence the system's norms and outcomes, although they are not directly involved in decision-making (Liu, 2000). The stakeholder onion tool has its roots in the Organisational Semiotics theory, which is further discussed. In this thesis, due to the iterative nature of the research solution, the roles other than actors may eventually become actors to solve other problems related to CS expertise discovery.

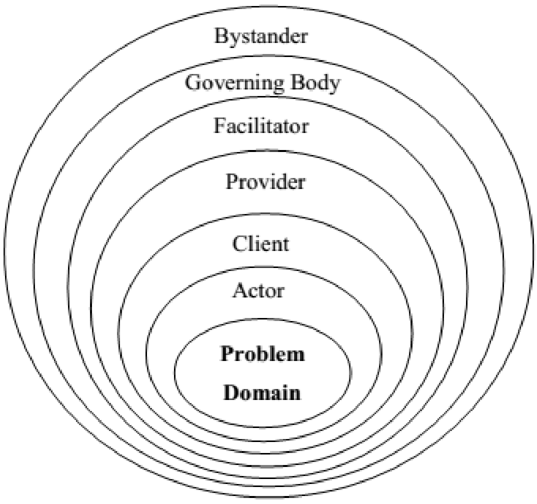


Figure 2.8: The stakeholder onion tool (Liu et al., 2006; Stamper and Kolkman, 1991)

Requirements engineering is critical for the development of successful IS and generally comprises the four processes (i.e. the central circle in Figure 2.9) and respective practices (i.e. the grey boxes in Figure 2.9). It is highly recommended that stakeholders participate in all processes due to the increasing complexity of the system. Combining different techniques (e.g. interviews, brainstorming and questionnaires) with graphical notations to approach stakeholders can increase consensual understanding (Ousmanou, 2007). This thesis relied on the iterative nature of requirements engineering for developing the research solution and detailing such development (Chapter 5) in order to facilitate the solution's further improvement. Therefore, aligned with Knowledge Management theory, requirements engineering guides the development of the meta-processes, which maintain the research solution.

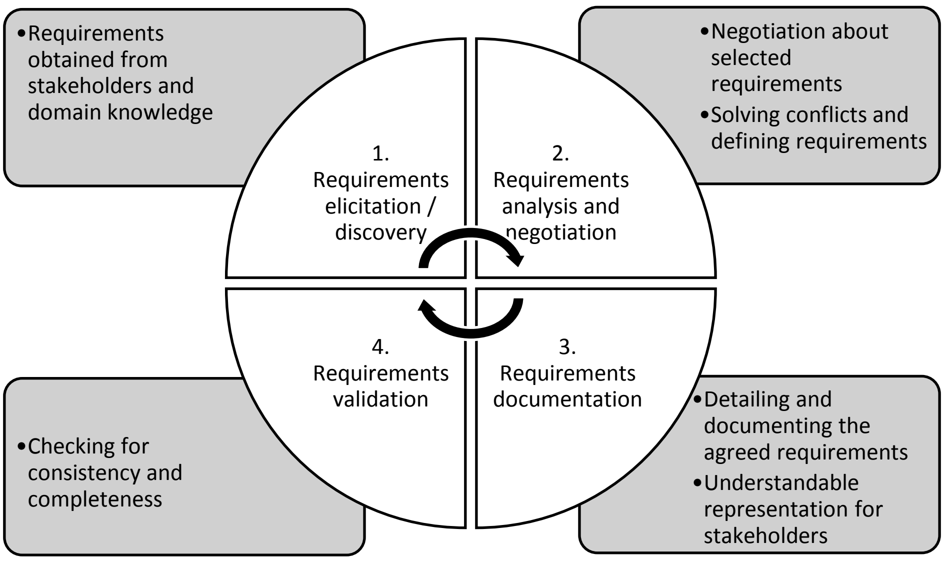


Figure 2.9: Requirements engineering processes (adapted from Kotonya and Sommerville, 1998; Liu, 2000)

Designing partnerships as organisations increases complexity and should take into consideration additional requirements in order to avoid collaboration inertia. Conflicting issues may arise when stakeholders from different organisations are involved. Thus, a leading organisation is a key factor for reconciling needs and partnership success. In addition, partnerships require consensual, clear and achievable goals for all involved stakeholders (Harris, 2011; The Open Group, 2013). The aim of stakeholder analysis is to identify the actors (whether people or groups) and their relevance according to a given problem situation (Sani, 2011). Actors and stakeholders are used interchangeably in some research fields. However, stakeholders are people or organisations that have interests or concerns relative to a system (Land et al., 2008), whereas an actor "specifies a role played by a user or any other system that interacts with the subject" (OMG, 2007: p. 586). The actor role can be fulfilled by a set of individuals (collectively or concurrently), and an individual can fulfil a number of actor roles concurrently (Liu et al., 2003). In the context of CS expertise discovery, the requirements engineering process is particularly relevant because it focuses on describing an entire community comprised by different organisations and partnerships in contrast to a single agency. Both cases differ in terms of scale and complexity.

2.4.3 Service-oriented Theory

Service-orientation plays a relevant role in connecting information sources and agrees with the adopted approach in this thesis where members in the CS community exchange and process information for each other in order to improve the overall expertise discovery (e.g. Figure 2.2). In service-oriented theory, business competencies are seen as services (Jambari, 2013), which basically consist of processes and pieces of information (Liu and Li, 2015). However, an organisation can be described through the set of services it provides, i.e. a service inventory (Erl, 2007). Numerous governments have been improving public service quality at an increasing rate due to advancements on ICT. This transformation of public services enables the adoption of a user-centred approach where the information flow reaches various platforms. However, service provision quality depends on integration and interoperability (Mushi, 2012) and the research scope concentrates on the business level. Hence, the CS community can be described as a service-providing organisation, and one of the challenges in this work is to map its service inventory.

Service-oriented thinking grasps opportunities offered by ICT to address flexibility and agility issues regarding the management of people, knowledge and organisations (Demirkan et al., 2008). Services enable access to capabilities, which are specified by a service description. Such a description comprises a set of constraints and policies afforded by the service provider. When invoked, services provide value by changing the state of a given entity, returning a piece of information or a combination of both responses (OASIS, 2006). For example, an individual can be evaluated with regard to the knowledge exchanged in a social media platform or his/her performance in a project. Services present a twofold approach. From the (external) point of view of a service consumer, a service expresses a functionality unit whereas hiding its technical details (The Open Group, 2013). From the (internal) perspective of service providers, services are actually pieces of software independent from business processes that can be used either individually or through compositions (Broens, 2004). For example, registering in an e-commerce site enables a consumer to buy assets. In some cases, the transaction processes are redirected to other servers that have the single purpose of verifying payment details and ensuring protection of such details. When the transaction is concluded, an amount is debited from the consumer’s bank account and an item is added to the consumer’s order log. After receiving the purchase, the consumer may evaluate the transaction, which can affect the seller’s popularity. These processes are transparent for the customer and may be reused whenever requested by any registered consumer. In this thesis, the processes structuring the services that contribute to the discovery of CS expertise depend on the information needs and on the responsibilities to provide such information. Hence, information sources for expertise discovery are further discussed. These processes are not constrained by any specific programming language.

Service-oriented architecture (SOA) is guided by the service-orientation paradigm and complies with design principles described in Table 2.9 (Erl, 2007).

Table 2.9: Design principles for services (Erl, 2007)

Design principle	**Explanation**
Standardised contract	"Services within the same service inventory are in compliance with the same contract design standards."
Loose coupling	"Service contracts impose low consumer coupling requirements and are themselves decoupled from their surrounding environment."
Abstraction	"Service contracts only contain essential information and information about services is limited to what is published in service contracts."
Reusability	"Services contain and express agnostic logic and can be positioned as reusable enterprise resources."
Autonomy	"Services exercise a high level of control over their underlying runtime execution environment."

Design principle	Explanation
Statelessness	"Services minimise resource consumption by deferring the management of state information when necessary."
Discoverability	"Services are supplemented with communicative meta data by which they can be effectively discovered and interpreted."
Composability	"Services are effective composition participants, regardless of the size and complexity of the composition."

The CS community relies on a rich network of connections and on the data they convey to improve the quality of expertise discovery. For example, a pilot questionnaire has shown that the top source for discovering expertise is networking in conferences in contrast to traditional organisational skill mapping. Agreeing with such an assumption, applications with SOA, which are focused on communication, are shaping trends with Big Data Analytics and Enterprise Architecture. In fact, organisations that rely on innovation have been using SOA to assist in reducing the gap between business and IT (Zimmermann et al., 2013), although such a topic is beyond the research context. The interoperability of distributed information sources (e.g. retrieving information from a social media profile), the encapsulation of processes and the use of an application programming interface (API) are examples of SOA applications suitable for CS expertise discovery. Service is also a key concept used and modelled with Enterprise Architecture. These applications are discussed in the following chapter.

2.4.4 Organisational Semiotics

Organisational semiotics (OS) regards the CS community (i.e. a system) as a social domain comprised of agents (e.g. people and organisations) performing specific behaviours (e.g. selecting experts). These behaviours follow certain patterns, which can be described using OS. Semiotics is the study of signs and meaningful communication. Liu (2000: p. 13) argues that semiotics “covers the whole cycle of a sign from its creation, through its processing, to its use, with more emphasis on the effect of signs”. Charles Sander Peirce (1857–1913), the founder of modern semiotics, suggested a process by which an interpretant is able to associate objects in the world with its representations (i.e. sign) and called it semiosis. This process is composed of three elements as presented in Figure 2.10. The dotted line means that the relationship between signifier and what it represents is subjective because it depends on the interpretant’s view. In other words, the interpretant is the one who establishes meaning.

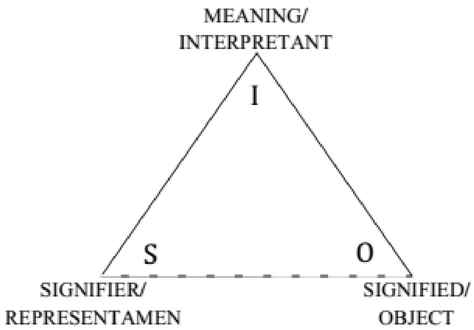


Figure 2.10: A version of the semiosis triangle from Peirce

Peirce also created three categories of signs (i.e. signifiers): icons, indexes and symbols. Icons have the purpose of resembling the object by presenting clear similarities (e.g. map, photos and drawings). Indexes are meanings derived from the original object (e.g. the smell of food in kitchen as lunchtime). Symbols are arbitrary signs that relate to convention or culture (e.g. traffic signs) (Liu, 2000). One of the key functions of signs is to extend behaviours beyond a specific place and moment. Thus, semiotics can deal with constraints within space and time dimensions (Stamper et al., 2000).

Semiotics is comprised of several branches because of its use in combination with other fields and applications. One of these branches, called OS, is focused in business and organisation, whether in the public or private sector (Liu, 2000). OS basically relies on signs and norms that describe behaviours (Stamper et al., 2000). A semiotic framework was created by Stamper to study signs in its six levels of manifestation. Stamper added three new dimensions (i.e. physics, empirics and social world) to the traditional dimensions of semiotics (cf. Figure 2.11).

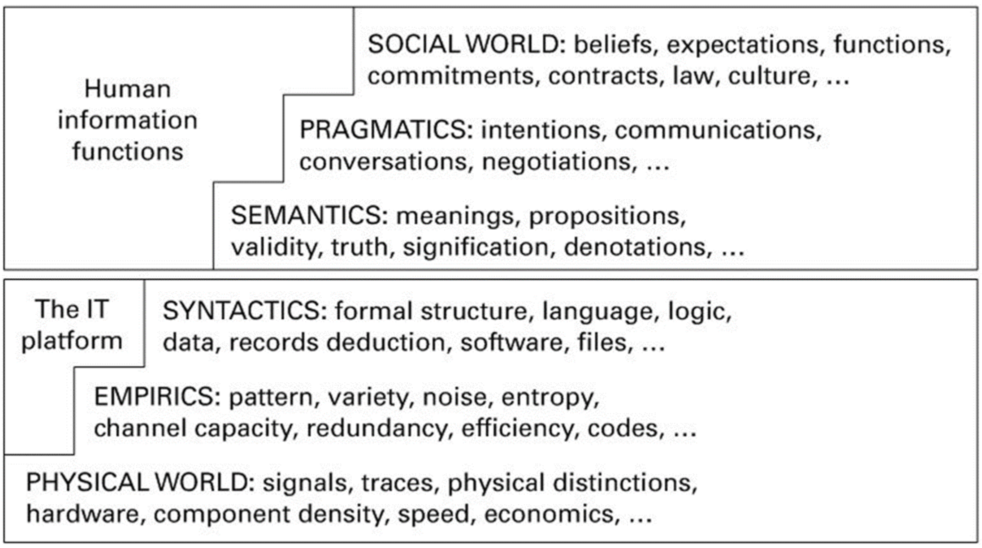


Figure 2.11: The semiotic framework (Stamper and Liu, 1994)

Semiotics have established the foundations on which actions lead to the creation of knowledge. Moreover, such a theory offers the tools to represent and dynamically capture the information flow in complex environments (Liu, 2000; Stamper et al., 2000). Therefore, semiotic-based ontology emerges as a suitable conceptualising approach capable of analysing the CS community as a socially constructed and complex IS. Moreover, the social world dimension relates to how interactions affect and shape the dynamic CS context over time. The MEASUR (Methods for Eliciting, Analysing and Specifying Users' Requirements) research programme was created in the 1970s by Ronald Stamper and covers semiotic methods for developing IS. MEASUR provides five major methods, two of which are Semantic Analysis Method (SAM) and Norm Analysis Method (NAM). The methods concern the three upper levels in the semiotic framework.

Aligned with the purpose of semantics in the semiotic framework, SAM is focused on information analysis, which conveys principles and concepts described by knowledge representation. The concept types in SAM are agents and affordances, which are mutually dependent. On one hand, agents perform behaviours when playing a specific role and eventually learn from such actions. On the other hand, behaviours are the expression of agents' knowledge and consequently depend on them (Liu, 2000; Stamper et al., 2000). These concepts are particularly relevant for the research context because expertise means knowledge based on experience (Davenport and Prusak, 2000). Such arguments imply that the discovery of expertise requires structuring the CS community in terms of its stakeholders (i.e. agents) and their responsibilities (i.e. behaviours). Moving on to a business ecosystem approach, these responsibilities represent different perspectives to achieve a single goal.

When using signs to represent the behaviours, it is possible to structure such behaviours from beginning to ending (i.e. a semiotic behaviour), capturing the full spectrum of underpinning actions. The concept of affordance was coined by James J. Gibson to represent a pattern of behaviour, which is formally described by norms. Each of these concept types may have properties, which are called determiners, and are connected via an ontological dependency. Each affordance is connected to only one or two antecedents (i.e. an agent or another affordance). Such a concept arrangement is called a semantic unit and it is the smallest piece of knowledge in an IS according to the OS approach. The principles to represent those concept types are explained in the following chapter.

Information analysis does not need to be linearly performed and corresponds to requirements engineering processes (cf. Figure 2.9). SAM helps to elicit and represent requirements in a formal way and comprises four steps (Liu, 2000):

- Step 1 (problem definition): description of the social system based on documents and interviews with system users, which encompasses the problem definition. After gathering sufficient evidence on the system, a problem statement summarises the context. In this thesis, an investigation was conducted on governmental documents, strategic initiatives and academic literature to describe the current CS expertise discovery context. Such investigation was triangulated by a pilot questionnaire and interviews with members of a CS community.
- Step 2 (candidate affordance generation): collect relevant keywords identified in Step 1, suitable to become affordances. It is worth noting that an agent may also be considered to be an affordance. The meaning of each collected keyword must be well conceptualised in order to avoid ambiguity and misunderstanding. During the articulation of the problem in this research, such keywords encompass business elements, which were determined via an application of stakeholder analysis and requirements engineering.
- Step 3 (candidate grouping): categorising keywords (i.e. semantic candidates) as agents, affordances, role names or determiners and then grouping these semantic candidates in semantic units (the notation is presented in the following chapter). In this thesis, these semantic units derive from a conceptualisation of the CS community in terms of the business elements, identified in the previous step.
- Step 4: drawing ontology chart (described in the following chapter).

NAM is performed after SAM, focuses on functional analysis of affordances and is closely related to pragmatics. The main purpose of NAM is to describe behaviour in affordances analytically within a social context by using norms. A norm describes which agent is responsible to perform what type of actions during a specific period. It specifies patterns of behaviour in the system agents by using constraints and conditions. Using norms in the semantic model enables the defining of “roles, functions, responsibilities and authorities of agents” (Liu, 2000: p. 47). Whether written in natural language or in machine-readable format, norms can formally detail behaviour patterns’ constraints and rules by using the construct in Figure 2.12 (Liu, 2000; Stamper et al., 2000).

Whenever <context>
If <condition>
Then <agent>
Is <deontic operator>
To do <action>

Deontic operators:
- Obliged
- Permitted
- Prohibited

Figure 2.12: Norm construct (Stamper and Liu, 1994)

Within the norm construct, the context defines the situation in which the behaviour takes place (e.g. an organisation as an employer can employ an individual). However, this context entails certain conditions detailing the requirements to apply the norm. If a condition is met, the agent is the entity in charge of performing an action according to the selected deontic operator (Sun et al., 2010). While SAM enables the representation of conceptual meanings and its relations, NAM is able to pragmatically contextualise their behaviour. Therefore, using SAM in combination with NAM provides "completeness, consistency, adaptability and interoperability" (Ousmanou, 2007: p. 116) of specifications.

Ultimately, the social layer in the semiotic framework enables the modelling of organisations as IS, since it takes into consideration the social context, which is beyond the modelling of data, processes and behaviour. Such an assumption means that either technical IS (e.g. machine-based or programmed), formal IS (e.g. regulations and manuals) and informal IS (e.g. social and cultural behaviours) can be systematically described through conditions and constraints (Liu, 2000). Thus, the CS community, as a complex organisation containing the aforementioned three types of IS, is capable of being modelled with OS. The combination of both SAM and NAM enables the performance of requirements analysis and concentrates on the upper levels of the semantic framework, where the human information functions occur. Indeed, those levels are not constrained by the lower layers in the semiotic framework when modelling IS (Liu and Li, 2015). Hence, OS is adopted as the approach to semantically model the CS community, taking into consideration agents' behavioural patterns, which affect expertise discovery. The representation of both methods is further discussed in Chapter 3.

2.4.5 Building Trust on People and Knowledge

Discovering CS expertise still requires a trusted environment (cf. Figure 2.2). Such an assumption is also supported by empirical research, where networking in conferences (i.e. face-to-face interactions) and peer recommendations are considered the most relevant sources for talent selection (as further discussed in Section 6.1). Trust is a "complex, multifaceted and context-dependent concept" (Koutrouli and Tsalgatidou, 2006: p. 153). Lewicki et al. (1998) define trust as "confident positive expectations regarding another's conduct" (p. 439). However, the challenge to define trust in a comprehensive way has led some researchers to focus on explaining trust typologies and defining its measurable components rather than creating a single and broad concept (McKnight and Chervany, 2001).

Trust has become a valuable asset for e-services (Vavilis et al., 2014) in partnerships between organisations (Harris, 2011) and in Talent Management (TM) (Altınöz et al., 2013; Boselie et al., 2005). In addition, trust and reputation systems are an increasing trend for decision-making support (Josang et al., 2007), which is a required feature for integrated TM (cf. Figure 2.7).

Reputation is a measurement of a given entity's trustworthiness (Swamynathan et al., 2010; Vavilis et al., 2014). Thirunarayan et al. (2014) argue that while "trust is local and subjective; reputation is global and objective" (p. 185). Reputation systems are discussed in the following chapter.

2.4.6 Viewpoints on Organisational Modelling for CS Expertise Discovery

This section explores the selected Knowledge Management and organisational modelling aspects (cf. solid rectangles in Figure 2.13) to address the viewpoints on managing talent for CS (cf. Figure 2.7 or cloud shapes in Figure 2.13). These aspects lead to the introduction of suitable approaches to enable proper information representation to identify the information sources and techniques to process such information (cf. dashed rectangles on top of Figure 2.13). Further discussion on these approaches is consequently performed in the following chapter.

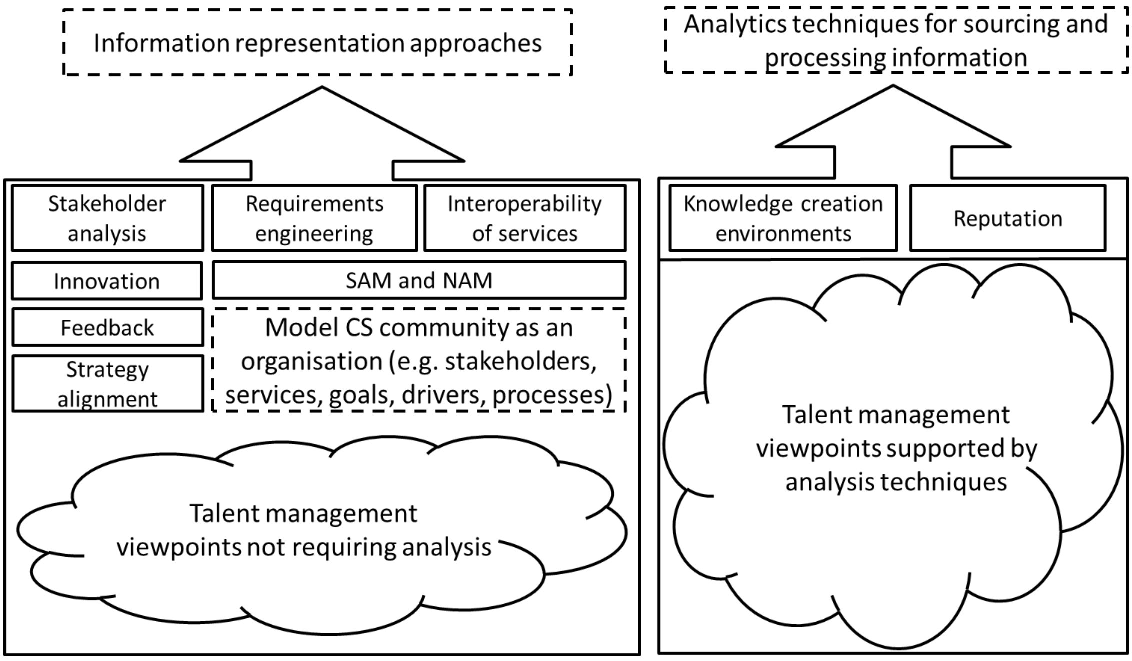


Figure 2.13: Knowledge and organisational theories supporting cyber security and talent management viewpoints

In line with current partnerships and initiatives, this thesis depicts the CS environment as a community of innovation where each actor contributes with specific behaviours to promote the overall expertise discovery in national CS (e.g. Figure 2.2). Organisational Semiotics is used to structure those behaviours, since the other organisational modelling theories discussed do not have such capability. The general term community is used to refer to the CS environment, considering the aspects of both the business ecosystem and the community of innovation theories.

Inspired by business ecosystem challenges and in order to optimise innovation within the CS community, organisational modelling theories provide valuable indicators to scope CS expertise discovery "as is" and draw how it is supposed "to be". Hence, analysing CS from an organisational perspective enables a better understanding of the motivations involved in the Talent Management viewpoints and delivers a theoretical underpinning to model the expertise discovery. From this perspective, national CS can be seen as an overarching organisation built upon public-private partnership, which delivers services to its stakeholders. This approach enables the identification of the roles and responsibilities of CS stakeholders, while managing conflicting concerns between partner organisations and single individuals. Hence, knowledge representation approaches aligned with previously discussed organisational modelling theories are discussed in the following chapter.

Aligned with the concepts of the community of innovation, an application of a pluralistic approach towards Talent Management is adopted in order to comply with the needs of a broader set of stakeholders within the dynamic and interactive CS community. Moreover, it seems that the adoption of a balanced approach towards talent in this matter is the most fair. While the high potentials and high performers can be more easily identified, opportunities should be offered to all those who are interested in joining the talent pool.

Modelling the CS community and benchmarking interactions amongst its members (either participants or organisations) enables innovation, which is a necessary commodity for the dynamic CS environment. Innovation to improve expertise discovery can be achieved in different scenarios (i.e. "ba" cf. Table 2.8), where indicators for information analytics approaches are suggested. For example, expertise discovery can be improved through externalisation and socialisation, which calls for a discussion on profiling techniques and social media. A combination of explicit knowledge in profiles can similarly generate new knowledge, which leads to a debate on approaches such as Big Data Analytics, data warehouse and performance measures in the following chapter. Special attention must be given with explicit feedback in order to promote creation of new tacit knowledge.

Moreover, locating the person with the right tacit knowledge for CS projects (i.e. mapping the expertise) and learning from such task leads to an investigation on Multi-Criteria Decision-making Methods (MCDM) and machine learning approaches as analytics techniques.

Despite a strong IT orientation, service-oriented theory enables the understanding of the CS community as a set of interoperable services, which are provided and requested by its own stakeholders. An important application for SOA concerns the ability to map expertise on different information sources in order to improve the global expertise discovery process. OS theory enables conceptualising the CS scenario based on the informational and functional analysis of its agents. SOA provides scalability and stateless features for enterprises, while OS has its own approach to modelling IS (Liu, 2000) and supports skills development (Ousmanou, 2007). These applications are required when modelling an expertise discovery system and deploying talent practices. In addition, both theories along with enterprise architecture (discussed in the following chapter) and social-technical systems are considered to be complementary approaches that lay the path towards compliance of business and IT systems (Liu and Li, 2015). Such a feature would also contribute to change management, which is a key issue in the CS knowledge domain. However, the IT platform (cf. Figure 2.11) is beyond the scope of this work. Knowledge creation environments (cf. Table 2.8) provide valuable cues regarding how to source information for discovering CS expertise. Applying reputation systems to these knowledge creation environments benchmarks the required trust (cf. Figure 2.7) in the CS expertise. Some related approaches are presented in the next chapter.

2.5 Summary

This chapter describes the interdisciplinary nature of CS communities and the major issue of discovering the expertise. Current initiatives concentrate on increasing awareness, providing education and regulations. However, coordinating efforts on such a scale requires proper requirements elicitation and analysis on information about capable experts. Therefore, an investigation on Talent Management and Knowledge Management supported identification of the main talent practices to discover the CS expertise and the theories to model such practices within the CS community. The best practices in those theories were grouped, in order to become more powerful, and adopted as the “viewpoints” in this work. Those viewpoints combine the strengths of these theories, which led to a robust research solution.

With regard to the second research question, this chapter identified competencies, certifications and academic areas as relevant CS capabilities, which are henceforth categorised as types of criteria for expertise discovery. Such criteria should be properly structured and measured. Hence, a discussion on profiling techniques and performance metrics is conducted in the next chapter, while the nature and types of criterion for CS expertise discovery are specified in Chapter 5.

Strategic Talent Management offers the basic set of practices to manage talents (e.g. identification, selection, development, retaining and analysis of talent gap) which are adapted as categories of processes to perform expertise discovery for CS. Those selected practices are comprised in an expertise discovery lifecycle, which is further introduced in Section 4.2.4.

Regarding the first research question, Talent Management (TM) is just part of the solution to find the most suitable talents for CS. Indeed, strategic decision-making is crucial to selecting the best people for positions and TM requires IS. Decision-making processes, however, have not yet received proper importance on such platforms and there is scarce literature proposing methods to support decision-making in TM. Hence, the context of this research offers an opportunity to address such a theoretical gap in the fields of TM and CS. Several TM viewpoints (cf. Figure 2.7) can address the research question; however, they require proper representation, measurement and information processing capabilities (cf. Figure 2.13). An investigation into the selected approaches is performed on the following chapter.

Developing a knowledge management system to address expertise discovery has manifold implications (cf. Table 2.8), which include codifying and abstracting the CS expertise (further discussion is conducted on approaches for sourcing such expertise); gathering and processing information to support finding people with a desired set of capabilities (e.g. competencies and other relevant criteria); enabling access to useful information; and linking expertise sources to analytical processes in order to promote the knowledge flow. Aligned with Table 2.7, this research work understands capabilities (e.g. competencies and skills) to be the type of knowledge and driving force required in projects. Moreover, expertise can be understood as a particular set of capabilities, which can be developed and discovered by linking knowledge sources and optimising the knowledge flow. These approaches comply with our definition of expertise discovery presented with this research background (cf. second paragraph in Section 1.1).

The performance in projects over time can build up one’s reputation in order to ensure the trusted relationships that are required in the CS community. While viewing the CS context as a single community of innovation, Knowledge Management plays a crucial role in understanding the knowledge flow in such an environment and organisational modelling theories provide the ways to structure the search for capable experts. Such an assumption addresses in part the first research question and still requires suitable techniques to process the search for experts.

The following chapter introduces knowledge representation tools and some analytical techniques to support required talent practices and to model other viewpoints, such as decision-making and reputational measures within the CS community.

Chapter 3

Information Analytics for Knowledge Discovery

The aim of this chapter is to introduce methods for modelling, sourcing and processing information to support expertise discovery in the multi-organisational CS scenario. The discussion includes approaches capable of supporting analytics so that expertise discovery can improve over time. Aligned with such a purpose, some well-established approaches and techniques are discussed. First, a concept of information analytics is briefly introduced. Afterwards, information representation is discussed through enterprise architecture, ontologies and profiling. Then, Big Data and related techniques are presented as valuable mechanisms for sourcing and processing information, which leads to the discovery of knowledge patterns in the resulting data. Finally, the main ideas applicable to the research problem are summarised.

3.1 Conducting Information Analytics

In this thesis, information analytics comprises a set of knowledge representation methods and techniques which enable modelling and discovery of insights from knowledge patterns. Since knowledge representation (KR) is focused on representing information about a given context, capabilities measurement and discovery of relevant patterns in data are delivered with proper analytics techniques.

Analytics can be classified as descriptive, predictive or prescriptive. The first type aggregates data to provide insight into past events and describe what has happened. The second type of analytics uses statistical models to forecast results, to understand the future and to identify trends. The third type relates to optimisation and decision-support functions in order to advise on what sort of action has to be taken. The scale and complexity of the CS communities require analytical capabilities to assist knowledge discovery in terms of CS expertise. In fact, analytics play a major role in decision support, which is one of the core viewpoints to address the research problem. This chapter investigates methods and techniques for knowledge representation, information sourcing and information processing, which serve the purpose of devising the research solution.

3.2 Knowledge Representation

According to Davis et al. (1993), KR plays five main roles. First, it is a substitute for something for reasoning purposes. The main idea is to reason on a surrogate before acting on the thing itself. Second, it is a set of ontological commitments, i.e. the rules to represent the real world. Third, it is a fragmentary theory of intelligent reasoning. In other words, each KR solution provides a specific perspective on inferring new knowledge. Fourth, it is a medium for efficient computation; therefore, it can use computational processing capabilities. Finally, KR is a medium of human expression or the language used to describe the world and to communicate with machines (Brewster and O'Hara, 2007). It requires proper information representation to understand the role of talents in CS and to take action for developing knowledge in such environment. The KR approaches in this thesis were selected in compliance with the organisational theories introduced in the previous chapter and deemed suitable for the context of CS expertise discovery. Approaches include enterprise architecture (EA), ontology modelling and profiling, which relate to the theories of SOA, OS and Talent Management. Then, some applications of information representation, which reflect the expertise discovery viewpoints, are discussed.

3.2.1 Enterprise Architecture

The main purpose of EA is to "improve the management and functioning of complex enterprises and their information systems" (Lapalme et al., 2016: p. 103). In this thesis, EA is presented as an alternative to deploy the service-oriented organisational approach in the context of the CS communities and to identify its organisational elements affecting expertise discovery. A graphical notation of EA is employed to facilitate discussions with stakeholders during articulation of the research problem. The importance of understanding such a method is because the research solution promotes revisiting the problem articulation for iterative improvement. In compliance with the overarching organisational approach (cf. introduced in Section 2.4), EA is examined as a means to represent the CS community's business elements along with its guiding motivational concepts.

EA offers a three-level framework comprising the business, application and technical layers. The business layer enables an integrated view of the organisation. The application layer comprises the applications and data structure based on business requirements. Finally, the technology layer is the technological infrastructure supporting all functioning systems (Mushi, 2012). EA aims to assure compliance between strategy planning and strategy execution, whether regarding external factors (e.g. government regulations and stakeholders needs) or internal factors (e.g. business and IT alignment) (Land et al., 2008). Services in either three layers are designed to fulfil organisational goals, which are motivated by internal or external drivers (The Open Group, 2013). EA depicts the services, functions and processes across multiple layers of an organisation and treats the different layers of organisations as independent units. The individual description of architectural components enables a clear distinction between business and IT layers.

EA extends the traditional service notation (Figure 3.1) from service-oriented theory by enabling the notation to represent other relevant elements for organisations, such as goals, stakeholders and processes. Such a feature helps to determine the value of services and make complex service compositions easier to understand.

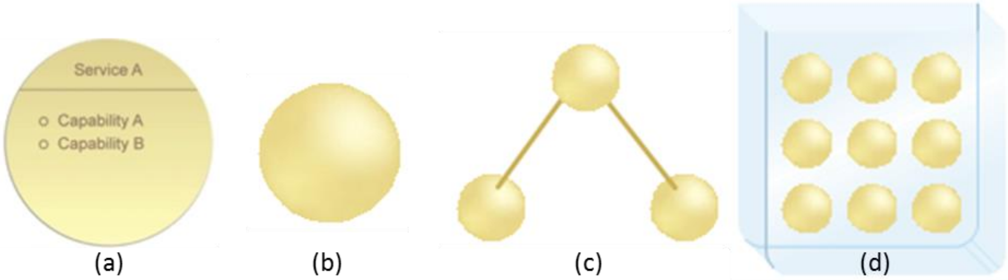


Figure 3.1: (a) Service chorded symbol, (b) service, (c) service composition and (d) service inventory (Erl, 2007)

The Open Group Architecture Framework (TOGAF) is a well-known EA framework comprising an IS layer which guides the alignment between business and IT layers. The Architecture Development Method (ADM) is the framework core. Archimate is a lightweight modelling language for representing an EA, which complements TOGAF with a graphical representation (The Open Group, 2013). The Archimate core language is comprised of three main types of elements (i.e. information, behaviour and structure) distributed in three layers (i.e. business, application and technology) and its features are closely related to ADM (Figure 3.2).

The information type of elements (i.e. the passive structure) comprises elements such as the business object. Behavioural elements include services, processes, functions and events. The structure type of elements (i.e. the active structure) similarly comprises actors and roles. Finally, motivational concepts used in this thesis include stakeholders, drivers, goals and principles. Although part of the application layer, components are modular parts of a software system that encapsulates behaviour, which can represent some of the techniques used in this thesis. Services are perceived as a functionality from an external point of view. However, they are actually described as a set of functions, processes and interfaces.

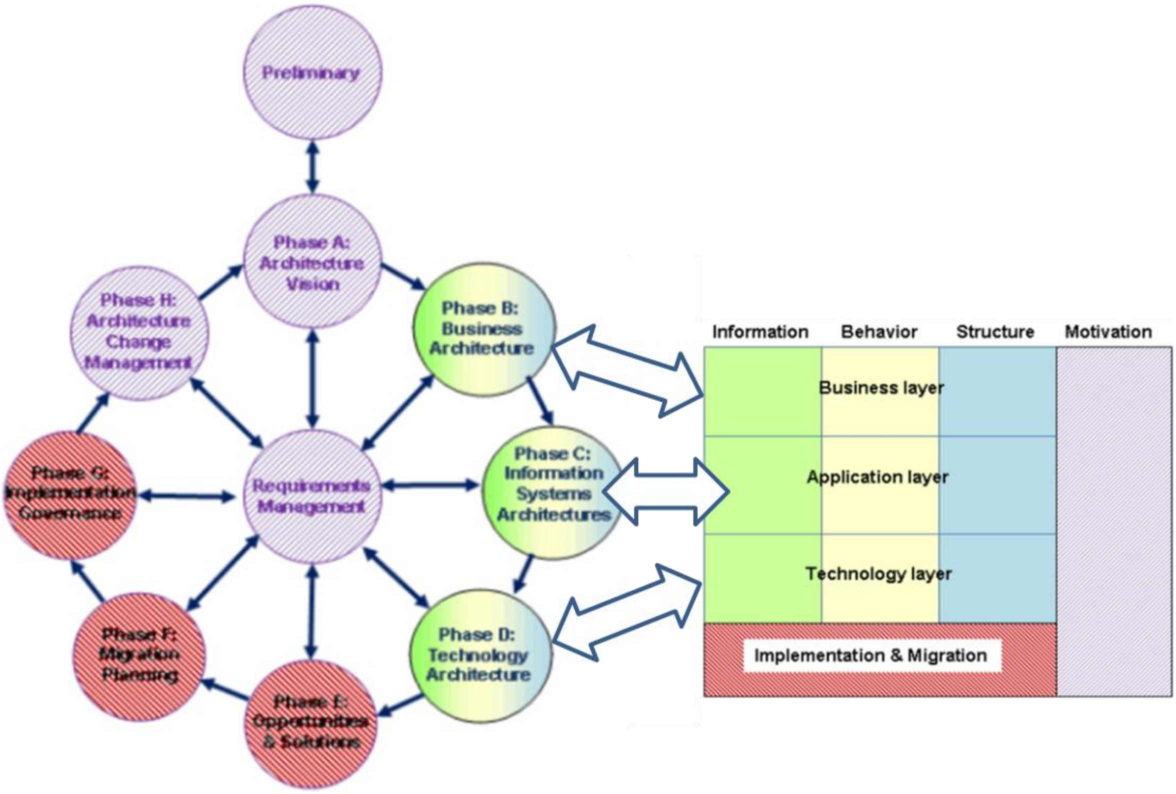


Figure 3.2: TOGAF ADM and Archimate with extensions (The Open Group, 2013)

In order to understand the aforementioned Archimate concepts, Figure 3.3 depicts an example of a function for selecting expertise with related business and motivational elements used in this thesis with Archimate.

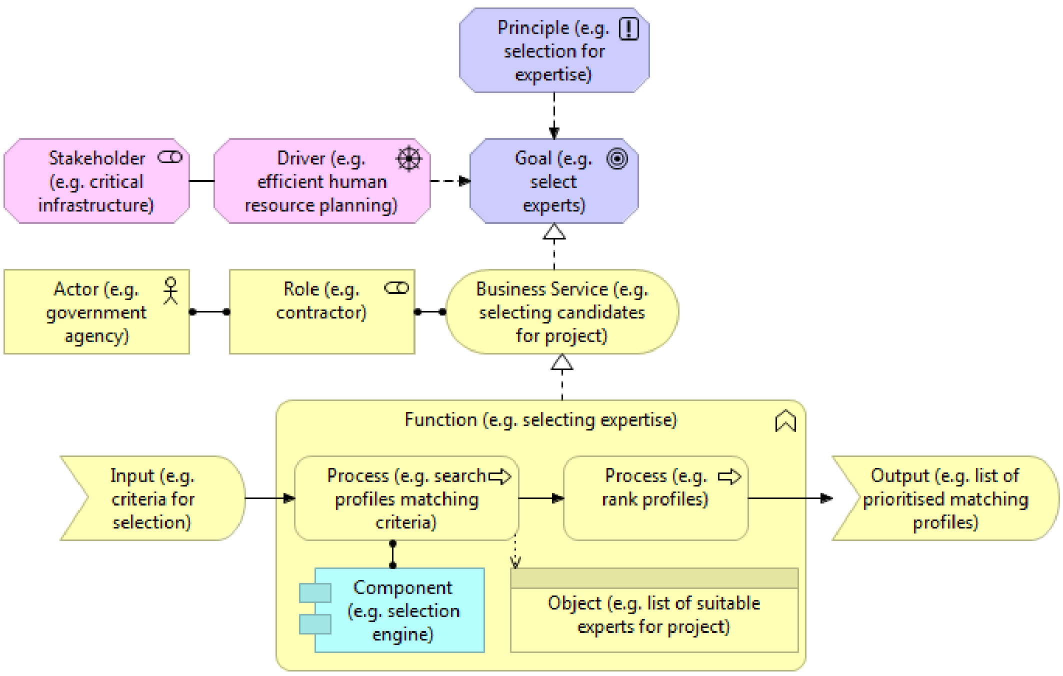


Figure 3.3: Example of selecting candidates for a project

The example depicts a service for selecting candidates for a project, which fulfils a business need from a contractor. Such a role is played by a government agency. The service is realised by a function comprised by a couple of processes. The first process is triggered when receiving criteria for selection, which is processed by a selection engine containing an algorithm. The result is a list of suitable experts for the project (i.e. a business object), which triggers a process for ranking the list of experts. Therefore, the service achieves its goal (i.e. to select experts), which is the end state that a critical infrastructure intends to reach. Such an end state is actually influenced by efficient human resource planning (i.e. a driver), which motivates change in the CS community (i.e. the organisation). The selection of expertise principle is a normative property in the context of CS expertise discovery. Principles should be specified through requirements and describe properties which affect goals. In this thesis, there are seven principles determining the way expertise discovery should be performed, which are further introduced during the articulation of the research problem. These expertise discovery principles encompass Talent Management practices and business ecosystems challenges, guiding the development of the research solution. A formal description of typical EA concepts can be found in The Open Group (2013).

EA is concerned with business and IT alignment. However, business and IT alignment is beyond the scope of this work. Moreover, this thesis adopts OS, which does not require the IT platform (cf. Section 2.4.4), to model the CS environment. Hence, only the motivational and business layer concepts in Archimate are further used in order to articulate the research problem. Kang et al. (2010) argue that current EAs still do not deliver proper semantics to enable a common understanding between people and systems, since relationships (i.e. behavioural elements) lack a logical formulation. Such a drawback leads to a discussion on ontology modelling.

3.2.2 Ontology Modelling

Ontology is defined as “a formal specification of a shared conceptualisation” (Borst, 1997: p. 12). It comprises a set of concepts well established and defined in a specific context, the connections among such concepts and properties that explain the concepts (Gruber, 1993; Sun et al., 2010). Ontology has become the favourite option for knowledge representation in the past few years and relies on contextualisation (Brewster and O’Hara, 2007), automation and semantics (Sani, 2011). Ontologies are widely used in developing Knowledge Management and in the emerging field of the Semantic Web (Staab and

Studer, 2004). This section introduces lightweight and heavyweight ontologies and a semiotic approach to ontology.

3.2.2.1 Lightweight ontology

Lightweight ontology has the ability to represent components of knowledge formally (i.e. classes, relations, attributes and instances of classes). Lightweight ontologies can be represented as a set of interconnected nodes (Liu et al., 2008), as depicted in Figure 3.4. In such an example, classes begin with capital letters and relationships with lower-case letters.

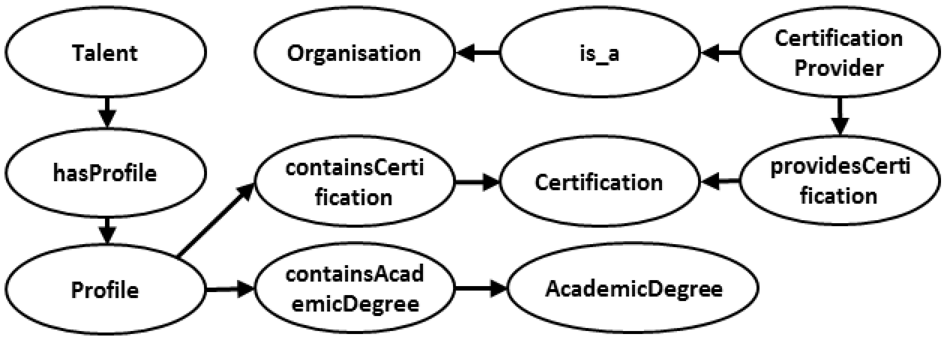


Figure 3.4: Example of lightweight ontology

The decision to build ontologies must follow some landmark assumptions regarding how to build them, the selection of the tool and the language selection (Corcho et al., 2003). Ontologies are generally coded in a particular language and syntax (e.g. Ontolingua, KIF and LOOM), which requires a specific tool (Liu et al., 2008). Terms in written language can have multiple meanings. Therefore, the concepts should be "ontologically committed". This means that there should be an effort to provide a shared and unique understanding of those values (e.g. using a vocabulary). Regarding the language, the Resource Description Framework (RDF) can describe resources (concepts) and model ontologies with metadata. RDF is actually a data model that can be expressed in a similar way to natural language. It is composed of a subject (a class, or a concept with a URI), a predicate (also called property, which defines attributes and relationships) and an object (a string, a number, other concept or even other RDF) (Gómez-Pérez et al., 2004). Thus, an RDF is also called a triple since it can be represented as three connected nodes. The subject is called the domain of a predicate, while the object is called the range of a predicate. Likewise, semantic units are comprised by concepts (and not data as in the RDF) that are connected to provide meaning. For example, in the semantic unit (Profile, containsAcademicDegree, AcademicDegree) shown in Figure 3.4, the domain of the relationship *containsAcademicDegree* is *Profile* and its range is *AcademicDegree*. Datatypes (e.g. age) are not allowed to be subjects in a semantic unit.

3.2.2.2 Heavyweight ontologies

Heavyweight ontologies enable significant richness to lightweight ontologies with the use of constraints (i.e. axioms) within concepts (Corcho et al., 2003; Sun et al., 2010).

Modelling heavyweight ontologies requires the use of first order logic or description logics to describe the axioms. However, a major drawback for a heavyweight ontology is that its axioms cannot be formally represented due to the lack of reasoning mechanisms (Ousmanou, 2007; Gómez-Pérez et al., 2004). In addition, some authors argue that ontologies require too much effort to be developed and may lack flexibility in complex real-world changing environments (Brewster and O'Hara, 2007).

3.2.2.3 Semiotic-based ontology

Semiotic-based ontology embraces the theory of Organisational Semiotics (OS), which relates with the semantic and norm analysis methods (SAM and NAM). Such methods cover a major requirement for knowledge representation which regards representation of axioms and temporality of constraints. Therefore, semiotic-based ontology can help overcome some of the heavyweight ontology drawbacks. Moreover, reasoning on temporal aspects as well as independence from specific ontology languages and flexible methods enables the combination of semiotic-based ontology with other approaches (Liu, 2000; Ousmanou, 2007). Conceptualisation of the context can be performed with axiomatic richness through representation of norms (Liu et al., 2008).

According to the semiotic approach, the existence of each concept depends on the existence of an antecedent concept. In contrast to the terms, classes and relationships used in ontology, OS uses agents and their behaviours (also called affordances), which are governed by norms. Agents may perform roles, and each role entails specific behaviours (Stamper et al., 2000).

Semiotic-based ontologies can be represented as schema containing descriptions of each semantic unit and other reference data, e.g. name, author and creation date, as depicted in Figure 3.5 (Liu et al., 2008). Each ontology chart comprises at least one semantic unit, which may have attributes (i.e. determiners) presented in different levels. For example, determiners may have their own attributes. Each semantic unit can be defined by a name, a type and may have numerous properties or no property at all. The affordances have a lifespan defined by its own norms (Stamper et al., 2000), i.e. a start norm and a finish norm, which are respectively triggered by a startAuthorityAgent and a finishAuthorityAgent. Such lifespans can also be defined in terms of time, i.e. startTime and finishTime. The ontology schema depicts a cardinality of one to up to two antecedents in each semantic unit. Research has been adopting for each affordance or business process a start norm, a finish norm to set its period of existence and one or more operational norms guiding conditions and actions (e.g. Sun et al., 2010; Sun and Mushi, 2010; Sun et al., 2014). However, operational norms are not necessarily required for a semantic unit.

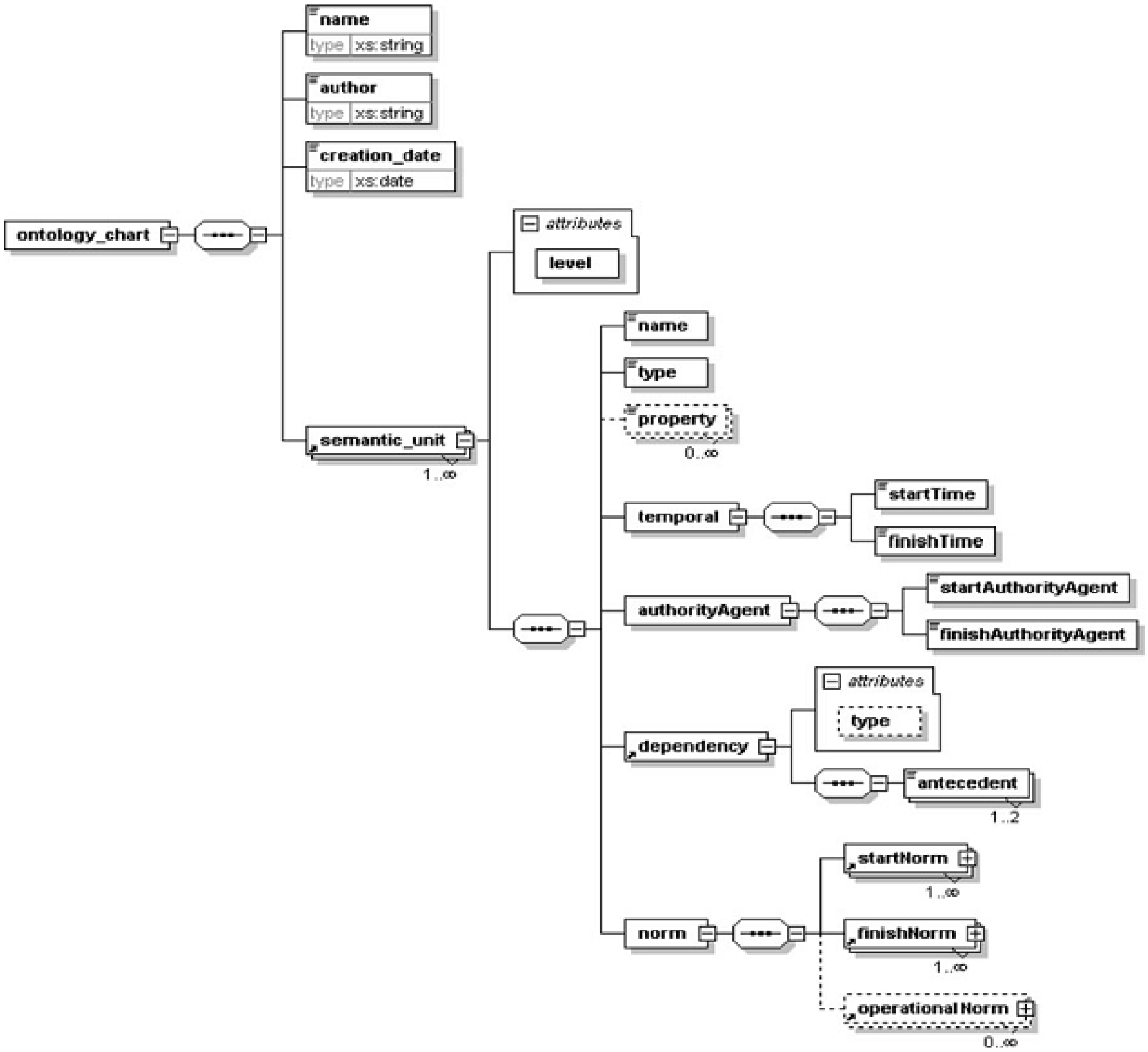


Figure 3.5: Ontology schema (Stamper and Liu, 1994; Sun et al., 2010)

There are a few design principles for building ontology charts (Ousmanou, 2007; Sani, 2011; Sun et al., 2010) which correspond to the last step in SAM (presented in the previous chapter):

- An ontology chart consists of a number of semantic units which are created and grouped during the first three steps of SAM.
- Semantic units:

    - Antecedent: the concept on the left is known as the antecedent. There may be one or two antecedents per semantic unit (either an agent or an affordance) connected to an affordance.
    - Affordance: the concept on the right side of the semantic unit is the dependent. Affordances are generally depicted as a rectangle and described by norms, which represent patterns of behaviour. However, an agent can also be afforded by another agent. For example, the concept of society is afforded by the concept of nation in Figure 3.6.
    - Ontological dependency: represented by a line, it connects concepts, which are part of a semantic unit (i.e. antecedents to their dependents).
- Agents and roles: agents are designed as ovals. When these actors perform behaviours (i.e. connected to affordances via an ontological dependency), they may play roles, which are represented by a semi-circle on the respective ontological dependency.
- Determiners: the features of the concepts are preceded by a hash symbol.

In addition, semantic modelling can add temporal information to the ontological dependencies, which can be stored in a semantic temporal database (Liu, 2000; Stamper and Liu, 1994). The underlying norms that govern each affordance can be represented and further implemented with a workflow engine (e.g. an activity diagram) to better understand interactions (e.g. Ousmanou, 2007). In this thesis, an ontology chart is used to model expertise discovery in the CS community.

In order to illustrate an example of ontology chart, let us suppose that an organisation (i.e. an agent) contracts (i.e. an affordance) a talent (i.e. another agent), as presented in Figure 3.6. Both agents afford “contracts”, which exist only during a period of valid conditions. In addition, if either one of the agents (i.e. an antecedent) ceases to exist, so does the affordance “contracts”. Here are some basic kinds of concepts used in an ontology chart (cf. Figure 3.6), which were introduced in Chapter 2: agent (the stakeholders or actors), role (also considered as an agent with specific responsibilities) and affordance (a concept that represents interactions between two roles or agents) (Stamper et al., 2000; Sun et al., 2010).

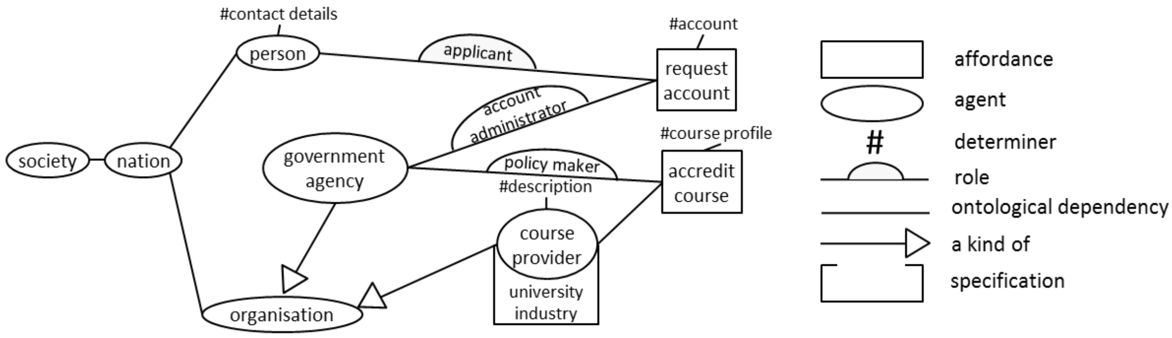


Figure 3.6: Ontology chart

3.2.3 Profiling and Profile Retrieval

Profiles enable the structuring and encapsulation of relevant information about a given subject or entity. Profiling is widespread in many disciplines and refers to store the user context and to configure applications to meet user's interests (Golemati et al., 2007). Profiling information about experts plays a major role in the context of this thesis, since the research problem regards the methodologies and information for discovering CS expertise as currently unstructured. Structuring relevant information about the available expertise promotes internalisation of knowledge (cf. Table 2.8) and is a necessary step towards knowledge pooling, which is one of the viewpoints to be addressed in this work. This section provides a brief explanation concerning user profiling and its applications.

A user profile represents a set of user's information needs, interests and preferences. There is a conceptual difference between user profiling and a user profile. User profiling involves the act of collecting and managing personal data and representing them in a user model or user profile. Such data can be explicitly provided by the user himself or implicitly inferred through behaviour patterns. Implicit information is normally obtained through data mining and other approaches that capture data from external information sources. The success of a search however relies on a sound definition of the user profile, which can be formally structured using an ontology (Calegari and Pasi, 2013). Profiled data may either be transactional which may be updated over time or summarised (i.e. data is persistent and changes are associated to a timestamp). Profiling practices include supervised and unsupervised learning; individual profiling, which focuses on gathering data about a specific person, and group profiles, when a large set of data is used to discover knowledge patterns about a group and assumed behaviours on members of such group (Hildebrandt and Gutwirth, 2008). This thesis adopts two types of profiles with different features. The first one relies on explicitly provided and transactional data for individual profiling to support the identification of experts. An empirical research (i.e. via an online questionnaire) determined the information needed to discover CS expertise. Such information supported structuring the profiles for the research solution, which are used to store profiled data. The research solution also adopts a group profile when analysing the overall expertise of the CS community (as further presented in Section 5.2.6). In this case, the profiled information to measure the expertise was drawn from existing individual profiles and stored in a list of variables using the summarised data so that these variables can be further used for analytics.

In contrast to data structures, a profile is not restricted to any particular method to organise data (e.g. graphs, tables, arrays or lists) or any particular technology. Profiles similarly do not require relations or constraints among its elements, although ontology can bring reasoning capabilities when used with profiles, as seen in the study by Golemati et al. (2007). Profiles cluster a variety of knowledge objects about a specific topic (e.g. individuals with a relevant set of competencies, documents and tasks) and are used to provide knowledge outputs. User's profiles must address some construction requirements in order to support knowledge creation and recognise skills, competencies, experience and prior knowledge (Owen, 1999; Sun and Mushi, 2010). Profiles also can be used to enrich services by defining a data structure composed of attributes and respective types (Table 3.1), which in this thesis is explicitly provided and refers to individual profiling (The Open Group, 2013).

Table 3.1: Example of data attributes and types for expert selection service

Attribute	**Type**	**Description**
Skill	String	Qualitative criteria that defines required capabilities
Number of skills	Integer	Quantitative criteria indicating the amount of capabilities
Postgraduate	Boolean	Quantitative criteria indicating if a PhD or MSc degree is required

There are some techniques for building user models, which include identifying user's concepts and vocabulary, measuring the responses that satisfy the user and the use of stereotypes to infer many facts from a sample (Rich, 1983). In concord with the former technique, this thesis adopted questionnaires and interviews to determine the information needed to discover experts and investigated additional issues with regard to user's features. Some of these issues involve securing private information, identifying what sort of information the user profile should contain, and determining how to acquire it and its purpose. Among other applications, user profiles are used as learning standards, e.g. IEEE LTSC Personal and Private Information (PAPI) standard and the Universal Learning

Format (Ousmanou, 2007) and for social networking purposes (Kaplan and Haenlein, 2010). However, those profile standards are focused on learning activities rather than criteria for other domain areas as seen in Mushi (2012). In addition, applying Big Data Analytics to a set of profiles can create valuable knowledge with regard to analysing expertise supply and demands for CS capabilities. Such new knowledge can be used to improve decision-making and expertise discovery over time.

3.2.4 Applications for Knowledge Representation on Expertise Discovery

The literature shows a growing number of ontology applications to address talent practices, such as recruitment (e.g. García-Sánchez et al., 2006), information provision (e.g. Ousmanou, 2007; Sun et al., 2010) and decision-making (e.g. Sun et al., 2014; Šaša Bastinos and Krisper, 2013). This section discusses strengths of and drawbacks to some KR approaches in order to select an appropriate solution for the research problem.

The use of ontology to manage competencies has also been the object of several applications. Staab and Studer (2004) describe the development of an ontology application for skills management in a Swiss company. Among other valuable lessons learned, skills ontology should be developed and maintained by experts. In addition, self-assessment and suggestions of new skills should be made available to participants. Both aspects are adopted in this thesis via recommended capabilities and feedback. However, the focus of that work was in matching profiles with required skills rather than providing a robust decision support for talent selection (e.g. considering ranking and the overall value of experts' profiles). Sure et al. (2000) developed IT skill matching capabilities using compensatory profile matching and weighting (M_C). Each skill is weighted (from unimportant to very important) and is graded (beginner to expert) within the job profile (p_j) and the individual profile (p_i) according to Eq. (1). The method enables the finding of suitable candidates that compensate missing skills with overdeveloped required skills and calculate the match result. Moreover, ontology is used to infer additional skills not present in the profile. However, the method is not able to rank the matching candidates according to additional skills in their profiles or based on additional criteria (e.g. certifications and academic degrees). Such an issue is addressed in this thesis since all types of capabilities are weighted following a recognised technique for ranking purposes (i.e. Analytic Hierarchy Process) and candidates are also able to compensate with overdeveloped capabilities (i.e. not just skills) requested in projects.

$$M_C(pi, pj) = \frac{p_i^T \; x \; (W * p_j)}{p_j^T \; x \; (W * p_j)};\; p_i, p_j \in P_{DB} \tag{1}$$

García-Sánchez et al. (2006) proposed an ontology-based system for recruitment, which describes a set of relevant profile attributes. In such a system, employment requests are recorded in curriculum vitae by the applicants themselves. The user profile is described in detail (i.e. concepts, relations and constraints). The method reported by García-Sánchez et al. (2006) requires developing taxonomies to categorise values in order to facilitate retrieval of attributes and the selection process. However, the system does not address criteria that support trust and is restricted to small regions where the applicants know the advertisers. In addition, it does not consider feedback from applicants supporting ontology evolution. The matching between profiles and job positions is semantically performed. However, it does not consider additional criteria based on applicants' performance, nor the subjectiveness of the decision maker. Although profile constraints is beyond the scope of this thesis and ontological modelling of CS capabilities is proposed as future work, this thesis addresses the drawbacks in the work by García-Sánchez et al. (2006) with regard to CS expertise discovery.

Colucci et al. (2011) propose an integrated competence management system using ontology reasoning with Description Logics. Some of the system features address core competence and knowledge gaps in the individual and organisational level. Such knowledge management system provides management in three levels (i.e. HR allocation choices, strategic choices and training programmes). HR allocation choices enable proper assignment of tasks according to semantic similarity with skills in user profiles with automatic team composition. Strategic choices treat agencies as a "competence warehouse" whereas the sum of the knowledge of each employee reflects the knowledge of the agency as a whole. Considering that all competencies in an organisation are mapped, either individual or collective training programmes can be developed. Such efficiency however would rely on the level of the ontology's formal description. A major drawback for this approach is that reasoning engines do not support evaluation in some information structures and do not consider flexible decision-making using weighted factors, which resembles human thinking (Colucci et al., 2011). Moreover, the nature of CS seems to require a "trans-organisational capability warehouse", due to the dynamic knowledge environment and to the collaborative approach of the initiatives involved. Hence, the research solution adopts the three levels presented in the work by Colucci et al. (2011) through selection for expertise, expertise analysis and provision of feedback principles and addresses its drawbacks for the context by adapting and combining flexible analytical techniques.

Assuming that OS can provide significant support to conceptualise Talent Management practices, it seems that semiotic-based ontology is a suited approach to represent meanings (semantics) and actions (pragmatics) within the multi-organisational CS context (social world). In fact, recent works successfully applied a combination of user profiles with a semiotic-based ontology (e.g. Ousmanou, 2007; Sun et al., 2010; Sun and Mushi, 2010). A semiotic-based ontology also enables the articulation of information embedded in profiles. In these works, profiles are used to organise information about users and other agents involved in tailored information provision, which is tightly related to talent practices (i.e. skills development). Such features are required for dynamic scenarios that involve personal development and training to support lifelong learning (Sun et al., 2010). Despite contemplating development of competencies, those approaches do not address talent selection practice, which is addressed by this research solution. Numerous examples of profiles, such as on social networking sites (SNS) and curriculum databases[7], constitute valuable information sources for profiling CS experts.

3.3 Information Sources and Processing Techniques

Establishing metrics is a viewpoint for CS expertise discovery. The absence of metrics has been a major concern for managing experts. Without the usage of metrics, decision-making and performance levels cannot be effectively appraised (Lewis and Heckman, 2006). Expertise selection should rely on efficient decision support techniques to allocate the right person to the right position (Lin, 2010). The increasing variety of information sources along with the intensive use of cyberspace is providing large amounts of data. Since the speed of data creation has surpassed its processing capability, a new paradigm is rising to solve the so called Big Data problems (Chen and Zhang, 2014). The advent of the Web 2.0 has also shifted the role of individuals from passive information consumers to active content creators, generating a significant leap in available personal information (Kaplan and Haenlein, 2010). In this section, data warehousing, Big Data and some of its related techniques are introduced as a means to enhance user profiles with a wider set of data sources, to improve data management and to support decision-making, performance assessment and other features related to managing the expertise (cf. Figure 2.13).

[7] e.g. Lattes platform available in <http://lattes.cnpq.br/>.

Moreover, a discussion on hybrid applications of information analytics methods is performed to identify suitability for CS expertise discovery.

3.3.1 The Role of Information Analytics in Expertise Discovery

There are two major approaches to manage the massive amount of data in the context of national CS expertise: data warehouse (DW) and Big Data. The former relies on extracting, transforming and loading data from different sources in order to supplement business intelligence. DW creates knowledge through combination of explicit knowledge sources (cf. Table 2.8) through data-driven (focused on structure of data sources), requirement-driven (focused on goals and user needs) or hybrid approaches. When adopting a DW, the metrics should be carefully defined to reflect the goals and indicators (Di Tria et al., 2017). Big Data focuses on volume (massive amount of data), variety (different sources and formats), velocity and value (generated from large groups of data) (Ferguson, 2014; Hashem et al., 2015; McAfee and Brynjolfsson, 2012). Application domains include business, social computing (e.g. reputation systems and social media), interdisciplinary research complex systems, and the public sector (Chen and Zhang, 2014). New kinds of data have emerged resulting in increasing complexity for data analysis (Ferguson, 2014). Among Big Data sources such as internet of things, machine and transactions, this thesis particularly relies on social media as discussed in the following section. One of the major advantages of Big Data is dealing with unstructured or semi-structured data (Chen and Zhang, 2014), as relational databases or DW cannot deal with these types of data. Gathering structured data (or information) is called retrieval, while gathering semi-structured or unstructured data is called extraction (Abdulrahman, 2012). In agreement with the work by Colucci et al. (2011), this thesis adopts a hybrid DW approach to measure the CS community as a capability warehouse and to identify the knowledge trends that promote continuous improvement. In order to promote scalability and performance, we assume that profiles deal with transactional data, which requires timely retrieval for the purpose of analysis. Parts of profiled information (i.e. capabilities) are retrieved and transformed into capability metrics. These metrics support the creation of indicators (as discussed in Sections 3.3.3 and 5.2.6) for the purpose of raising general awareness on relevant and emerging CS capabilities. Although the research solution is not restricted to any data structure (which is a topic beyond the scope of this thesis), the analytics is deliberately aligned with Big Data techniques for a twofold purpose. First, those techniques tend to be scalable in terms of data. Second, some techniques are capable of processing unstructured data, which opens opportunities for creating knowledge out of the rich data environment of the CS expertise domain.

Big Data and analytics are closely related because both approaches seek to generate business intelligence from data (McAfee and Brynjolfsson, 2012). Hence, the use of the term Big Data Analytics relates to data processing, while Big Data is concerned with data management (which is beyond this research scope). Big Data Analytics encompass techniques from different areas of study (Figure 3.7). Data mining, for example, serves the purpose to discover valid patterns in data and present knowledge in an understandable way to users. Indeed, data mining supporting HRM practices is increasing in the research community. The main purpose of data mining in HRM is to create knowledge and support decision-making (Kum et al., 2009; Strohmeier and Piazza, 2013). In this thesis, mining profiled information serves to determine performance metrics. Machine learning is a branch of artificial intelligence that aims to develop and improve behaviours in computers with the use of empirical data, and ultimately enable automatic decisions (Chen and Zhang, 2014). Exemplar-based reasoning (EBR), which is closely related to machine learning, is further discussed as a method for learning (i.e. acquiring tacit knowledge) from past projects. Optimisation methods (e.g. Multi-Criteria Decision-making Methods) relate to the selection of the most adequate element (or elements) within a set taking into consideration a set of criteria. Decision-making involves a series of steps such as knowing the problem, knowing the purpose, identifying people affected by the decision (groups and stakeholders), criteria, sub criteria, and alternative actions (Saaty, 2008). However, it is useful to begin decision-making by reducing the alternatives. This procedure is called screening (Chen et al., 2008). Multi-Criteria Decision-making Methods include Data Envelopment Analysis (DEA) and Analytic Hierarchy Process (AHP), which are adopted in the research solution for assisting selection for expertise. Data visualisation has the purpose of using graphs to present analytical results (Hashem et al., 2015), from which knowledge patterns can emerge to facilitate decision-making. Reputation and recommender systems are Big Data techniques, which play an important role in benchmarking trust in the CS community (cf. Figure 2.13). Visualisation methods (e.g. dashboards from key result indicators) are also considered as a Big Data Analytics technique and are commonly used in DW architectures. Those techniques can be combined or used to fulfil different purposes. For example, social network analysis (SNA) is often used as a visualisation method[8] and can be used to draw capability trends in communities of practice (e.g. Fontenele et al., 2014). Hashem et al. (2015) add SNA, neural networks, signal processing, pattern recognition and statistics to the list of Big Data Analytics techniques, some of which may be used in future work.

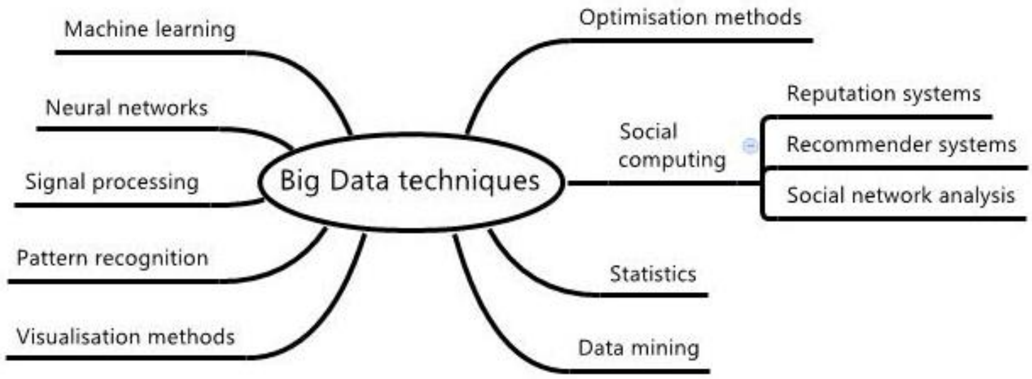


Figure 3.7: Some of the techniques used for Big Data (adapted from Hashem et al., 2015)

Some of the major challenges to use Big Data involve the use of inconsistent data (originated from different sources) and reliability issues (due to trustworthiness of the sources). Because Big Data Analytics have a profound influence for decision makers (Chen and Zhang, 2014), some techniques relevant for sourcing and processing data for expertise discovery are further discussed. The selection of these methods relies on their flexibility, because of their solid foundation in research and their alignment with requirements in this work.

3.3.2 Social Media as Useful Information Sources

Social media platforms are widely used in CoP to promote knowledge exchange and to support trust, although it seems that CS communities still do not properly use its benefits. Social media is "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0 and that allow the creation and exchange of user generated content" (Kaplan and Haenlein, 2010: p. 61). In other words, social media is considered to be a technology based social network. Social media is also an important Big Data source (Hashem et al., 2015) and a growing trend in Knowledge Management applications (Giuffrida and Dittrich, 2013; Jeners and Prinz, 2012). Social media has become a resource to improve profile reliability, since a digital identity is built through a process of co-creation, as a result of social interactions (El Ouirdi et al., 2015). In this thesis, social media is adopted as a means to perform and to source peer evaluations based

[8] Using tools such as Pajek (available in http://mrvar.fdv.uni-lj.si/pajek/) or Gephi (available in https://gephi.org/)

on social interactions focused on knowledge processes. In addition, it is believed that such interactions may promote the creation of tacit knowledge through socialisation and internalisation (cf. Table 2.8), which may also contribute to CS expertise discovery. This section presents how to use some social media platforms in order to evaluate shared knowledge, thus promoting trust in expertise discovery. However, it is worth noting that the implementation of a social media platform is beyond the scope of the research solution.

Social media applications enable the creation of knowledge networks where people collaboratively generate and share knowledge (Yates and Paquette, 2011). Companies use blogs to update employees and customers and videos for recruiting future employees (Kaplan and Haenlein, 2010). On the one hand, social media supports all processes of Knowledge Management (KM), knowledge reuse and faster decision cycles (Yates and Paquette, 2011). On the other hand, leadership shifting may occur with unwanted results, such as support of extreme communities or hate issues (Avolio et al., 2014). In Table 3.2 some examples are presented on the use of social media in knowledge management systems (KMS).

Table 3.2: Social media applied in KM (adapted from Kaplan and Haenlein, 2010, Giuffrida and Dittrich, 2013 and Yates and Paquette, 2011)

Classification	Example	Using in KMS and characteristics
Blogs	Weblog	Author-centred; informal knowledge sharing; externalisation of knowledge; improve reputation; document ideas; engage people inside and outside the organisation. Drawbacks: requires motivation and feedback.
	Microblog	Author-centred; informal broadcast medium; workspace awareness; transparency; can be used for publishing news about groups' findings. Drawbacks: data protection and privacy issues.
Collaborative projects	Wiki	Asynchronous and collaborative nature; knowledge construction and sharing; knowledge repository; raises awareness; provides history feature. Challenges: modify other's content properly; vandalism (although it is rare); information sensitivity; explicit ownership of collaboration; requires guidelines for usage.
		Good to share lessons learned. Drawbacks: shared ownership complicates KM; needs descriptive metadata; avoid information overload and check information accuracy.
	Google Docs	Manage documents in cloud computing environment; simultaneous and ubiquitous document editing.
	Microsoft SharePoint	Can be used for sharing knowledge within staff.
Content communities	Social bookmarking	Manage and tag online resources; knowledge distribution; expert finding; tagging is a collaborative activity; lightweight and customised tag vocabularies. Date tags help distinguish old from new knowledge
	Video and photo sharing	Create interaction for users; proved to be very useful in disaster response; creates situation-specific ontologies when tagged. Examples: YouTube and Flickr.

Classification	Example	Using in KMS and characteristics
Social networking services	Organisation al social networking site (SNS)	Greater importance to people search and context awareness; socialising; project planning; forge new connections with experts and contact them; improves reputation. Challenges: provide feedback on content visualisation; identity and multiple profile management.
	Instant messaging	Facilitate knowledge sharing; probe availability for meetings; synchronous and spontaneous communication; faster and more informal than email; suited for distributed teamwork

People tend to self-disclose more in computer communications (Avolio et al., 2014) and are willing to share experiential knowledge (i.e. expertise). However, most of the time, they do not have proper procedures or mechanisms to do so (Butler et al., 2008). An individual may self-disclose to manipulate others' impressions of himself (Kaplan and Haenlein, 2010) and because social media satisfies the need to belong and for self-presentation (Avolio et al., 2014).

Currently, SNS (e.g. Facebook and LinkedIn) is the most successful type of social media (Kaplan and Haenlein, 2010). External social media serve different purposes (Table 3.3) and provide a Big Data environment to build trust (Avolio et al., 2014).

Table 3.3: Examples of SNS

Main purpose	Social media	Website
Social	Facebook	https://www.facebook.com
	Google+	https://plus.google.com/
Academics and research	Academia.edu	http://www.academia.edu/
	Epernicus	https://www.epernicus.com/
	Figshare	http://figshare.com/
	Harvard Catalyst Profiles	http://profiles.catalyst.harvard.edu/
	Mendeley	http://www.mendeley.com/
	ResearchGate	http://www.researchgate.net/
	Social Science Research Network	http://www.ssrn.com/en/
	VIVO	https://wiki.duraspace.org/display/VIVO/VIVO+FAQs
		http://www.vivoweb.org/download
Professional	LinkedIn	http://www.linkedin.com/
	Xing	https://www.xing.com/

Some SNS (e.g. Academia.edu and ResearchGate) enable uploading articles and creating discussion forums, thus facilitating the creation of CoP. Instead of developing an application from the ground up, there are open-source solutions such as VIVO and Harvard Catalyst Profiles developed for the scientific community (Gewin, 2010). Both solutions are ontology-based, and therefore support extended functionalities.

Professional and academic social media also can provide valuable information for expertise discovery. The lack of proper searching mechanisms, the unreliable data, copyright issues and access restrictions, however, hampers such practice. Some initiatives in CS and government databases already support creating profiles where users can register their competencies. These initiatives, however, disregard decision support features for selecting talents as well as reliability aspects about informed expertise. For example, in terms of integration, automated information sourcing from corporate social networking sites is already possible through APIs. In fact, some social media provide these interfaces to enable data extraction (e.g. Facebook with its Graph API and LinkedIn) to support information retrieval on user profiles (van Dam and van de Velden, 2015). However, access to interesting features such as work experience, interests and contact information depends on the service provider and may not be available (Russell, 2011; van Dam and van de Velden, 2015). Even public academic databases, such as the Brazilian Lattes database, do not allow a customised search by given features despite curricula being saved in XML format (Castaño, 2008). In addition, acquiring data to support Talent Management is recommended in order to shift from a method-driven to a domain-driven approach. The domain-driven approach requires identification of domain-specific criteria and features (Strohmeier and Piazza, 2013). Therefore, it seems that a domain-driven social medium is more suitable for the CS community. Such an approach allows profiles to be manually populated by users themselves and eventually additional information can be retrieved from external social media via API to improve trust.

3.3.3 Information Processing with Performance Metrics

Performance benchmarking, measurement and evaluation are some of the required aspects for discovering CS expertise (cf. Figure 2.7) since measuring correlates highly with knowing. These aspects fall under the definition of information analytics and support visualisation methods. In this thesis, such a topic is adopted to develop a mechanism to measure the “CS capability warehouse” in the CS community and disseminate those measures among its actors. The lack of such a mechanism and awareness were identified as part of the research problem. Raising general awareness with clear performance metrics supports a systemic optimisation of CS expertise; i.e., actors are expected to prioritise the relevant capabilities with regard to their roles and responsibilities. This section provides a brief discussion on the role of the key result indicator.

Performance measures may benchmark past events (e.g. number of projects with missing expertise last month), current events (i.e. updated continuously) or future events (e.g. number of initiatives to develop the missing expertise in the next month). There are four types of performance measures: Key Result Indicator (KRI), result indicator (RI), performance indicator (PI) and key performance indicator (KPI). Such indicators are employed as metrics to support routine business reporting and the use of dashboards. This thesis adopts KRI for benchmarking the expertise. The basic difference between KRI and KPI is that the former measures (i.e. relates to descriptive analytics), supports information governance for a broader set of users and enables the creation of RI, PI and KPI. The latter, measured with higher frequency, tells us what should be done to obtain specific results (i.e. relates to prescriptive analytics), affects critical success factors and concerns the information management staff. Generally, a governance report should consist of up to ten KRI, among other indicators. A quarterly measure for the number of employment candidates and a number of potential recruits from contractor referrals is suggested. Likewise, a monthly measure for the number of users in a system or employees in an organisation is advised (Parmenter, 2010).

A KRI summarises activity in a critical success factor (Parmenter, 2010), which in this thesis refers to the selection of the right expertise for the right project within the CS community. Hence, the types of capability encompassed by CS expertise (i.e. competence, academic area and certification) were identified in a pilot empirical research. In order to benchmark the performance of these capabilities in the CS community and be/remain aligned with DW principles, a capability metric was devised as a tuple, consisting of the supply and demand frequency of each capability in a given time. These values are extracted from the pool of expert and project profiles and a timestamp is added. This metric is the basic data unit used to determine KRI in the context of the CS community (as further discussed in Section 5.2.6). Hence, the research solution extends the concept of competence warehouse in the report by Colucci et al. (2011) by determining the relevance of each capability in terms of supply and demand.

This thesis adopts KRI as a means to benchmark and raise awareness on the available CS expertise so that actors in the CS community can take timely action to address the knowledge shortage according to their own responsibilities. Actors may further create their own KPI to promote and benchmark improvement of the KRI over time. Developing KPI for the CS community, however, is beyond the scope of this thesis. The use of dashboards with KRI as a visualisation method is adopted as a means to enhance awareness on CS expertise.

3.3.4 Exemplar-based Reasoning

Exemplar-based Reasoning (EBR) is used to facilitate decision-making based on previous experience. The changing knowledge requirements in CS are part of the research problem. This thesis assumes that learning from successful CS projects can improve the requirement definition of future projects and ultimately the expertise discovery itself. EBR is one of the manifold approaches of case-based reasoning (CBR), which is described as a four-stage cycle: 1) retrieve similar past cases with some type of similarity threshold; 2) reuse the solution (in part or as a whole) to those cases; 3) revise (i.e. test) the adopted solution and 4) retain successful solutions for future problem-solving either by uploading the new case or by updating existing cases (Aamodt and Plaza, 1994; Voskoglou, 2011). The method can be combined with rule-based reasoning, fuzzy methods and other techniques. Many approaches can be used to design the information retrieval mechanism.

EBR focuses on determining the classification of the problem instead of solving past problems. This means that EBR reuses past problem features to structure the definition of new problems (Aamodt and Plaza, 1994; Voskoglou, 2011). There are many similarity measures that can be used with CBR or EBR, such as Euclidean Distance, Cosine Similarity and Jaccard Coefficient (*J*). The latter is suited for binary data that does not require numerical attributes and can be described using Eq. (2), where *A* and *B* are two sets being analysed in terms of similarity.

$$J\,(A,B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} = \frac{|A \cap B|}{|A \cup B|} \qquad (2)$$

The research solution adopts EBR as a means to suggest additional capabilities during the definition of CS projects based upon successful similar projects. Hence, the research solution is supposed to not only improve the search for the “right expertise”, but also to define the “right project” to address a specific problem. By applying Eq. (2) in the context of this thesis, *A* represents the set of capability requirements, which are being defined in a project, while *B* represents those requirements in successfully concluded projects. *J* calculates the similarity between projects *A* and *B*.

3.3.5 Data Envelopment Analysis

This thesis adopts Data Envelopment Analysis for selecting the most suitable candidates for a CS project. Data Envelopment Analysis makes it possible to work with a scalable and dynamic pool of experts and provides a manageable subset of efficient candidates based on a set of quantitative criteria. Data Envelopment Analysis is a method "whereby, within a set of comparable Decision-Making Units (DMU), those exhibiting best practice could be identified and would form an efficient frontier" (Cook and Seiford, 2009: p. 1). In addition, Cook and Seiford (2009: p.2) state that Data Envelopment Analysis "enables one to measure the level of efficiency of non-frontier units and to identify benchmarks against which such inefficient units can be compared". Efficiency is obtained comparing benefit (output) / cost (input) ratios among a given Decision-Making Units and those that belong to an efficient frontier (Charnes et al., 1978). Efficiency can be increased either by reducing input while maintaining a constant output value (input-oriented model) or by increasing output while maintaining a constant input value (output-oriented model) (Cook and Seiford, 2009). A major characteristic of Data Envelopment Analysis is that it relies on changing the weights of inputs and outputs of each Decision-Making Unit in order to maximise relative efficiency (Lin, 2010), as presented in Eq. (3).

$$
\begin{aligned}
e_0 = \max\ & \textstyle\sum_r u_r y_{r0} \,/\, \sum_i v_i x_{i0} \\
\text{s.t.}\ & \textstyle\sum_r u_r y_{rj} - \sum_i v_i x_{ij} \leq 0, \quad \forall j \\
& u_r, v_i \geq \varepsilon, \quad \forall r, i
\end{aligned} \tag{3}
$$

In order to convert the original Data Envelopment Analysis formula to linear programming, the problem in Eq. (3) can be converted to the model in Eq. (4).

$$
\begin{aligned}
& \mu_r = {}^{u_r}\!/\!{}_{\sum_i v_i x_{i0}} \\
& v_i = {}^{v_i}\!/\!{}_{\sum_i v_i x_{i0}} \\
& e_0 = max \textstyle\sum_r \mu_r y_{r0} \\
& \text{s.t. } \textstyle\sum_i v_i x_{i0} = 1 \\
& \quad \textstyle\sum_r \mu_r y_{rj} - \sum_i v_i x_{ij} \leq 0, \quad \forall j \\
& \quad \mu_r, v_i \geq \varepsilon, \qquad \forall r, i
\end{aligned} \tag{4}
$$

Where
n is the total number of DMUs
j is a given DMU (j=1,…,n)
m is the total number of inputs
i is a given input (i=1,…,m)
s is the total number of outputs
r is a given output (r=1,…,s)
ε is a non-archimedian number to enforce positive values
u, *v* are the weights of outputs and inputs, respectively

When measuring employee performance, for example, the input can be the payment and the time spent in the company, while the output can be his production. Data Envelopment Analysis can also be used for benchmarking instead of calculating performance when the evaluated features do not relate with a productivity frontier (Cook et al., 2014). The original Data Envelopment Analysis does not support ranking because Decision-Making Units are just classified as ‘efficient’ or ‘not efficient’ (Sinuany-Stern et al., 2000). Khodabakhshi and Aryavash (2012) claim to rank Decision-Making Units using a single and differentiated Data Envelopment Analysis approach. Notwithstanding this, input and output must have quantitative values, which is not always the case when dealing with user profile information. Some qualitative data can be transformed into quantitative data by using discrete scales, such as in employee satisfaction (e.g. Cook et al., 2014) and patient feedback (e.g. Sun et al., 2014).

In order to rank Decision-Making Units and to properly use Data Envelopment Analysis with quantitative and qualitative values, the literature provides examples of combining Data Envelopment Analysis with other methods (Hatami-Marbini et al., 2011; Lin, 2010), which are further discussed in Section 3.3.8. Table 3.4 summarises the strengths and weaknesses of Data Envelopment Analysis features.

Table 3.4: Data Envelopment Analysis features (adapted from Chebat et al., 1994; Cook and Seiford, 2009; Cook et al., 2014)

Advantages	Disadvantages
Allows to work in small samples	Only presents relative efficiency in specific set of criteria and to a given set of Decision-Making Units
Uses the same scale for different types of inputs	Simplicity of the model can hide mistakes in inputs and outputs
Weight can regulate proportion of efficient Decision-Making Units	Loses discriminatory power when number of inputs/outputs increase in relation to Decision-Making Units
Several models to choose from	Not suitable to use with percentiles and ratios in inputs and outputs simultaneously

The basic difference between Data Envelopment Analysis and regression analysis is that the former relates units to an efficient frontier while the latter relates the set of units to average performance (Chebat et al., 1994). Cook and Seiford (2009) present a review of the most relevant Data Envelopment Analysis models, either single level (e.g. constant returns to scale, variable returns to scale, additive model, Russel measure and alternative views) or multilevel. In this thesis, Data Envelopment Analysis is used to process the quantitative values with regard to criteria for discovering expertise. These values do not represent ratios as recommended in Table 3.4. The simplest Data Envelopment Analysis model (i.e. the constant returns to scale) is adopted and its flexibility to work in small or large samples enables a scalable research solution. The loss of discriminatory power due to an increasing number of inputs and outputs (cf. Table 3.4) is expected to be addressed with an optimisation process to rank the Decision-Making Units.

3.3.6 Analytic Hierarchy Process

A number of suitable candidates for CS projects can be selected. Analytic Hierarchy Process hence plays an important role for decision support by ranking those candidates. Ranking considers the relevance of each project criterion based on their weight. Analytic Hierarchy Process is a widely used Multi-Criteria Decision-making Methods created by Saaty (1977). Other examples of Multi-Criteria Decision-making Methods may be found in several studies (Belton and Stewart, 2002; Figueira et al., 2005).

Analytic Hierarchy Process can clarify the core of complex problems to decision makers whereas identifying and categorising criteria (Hor et al., 2010), even if the criteria is intangible and have no measurements as a basis (Saaty, 2008). The method comprises four steps as depicted in Figure 3.8.

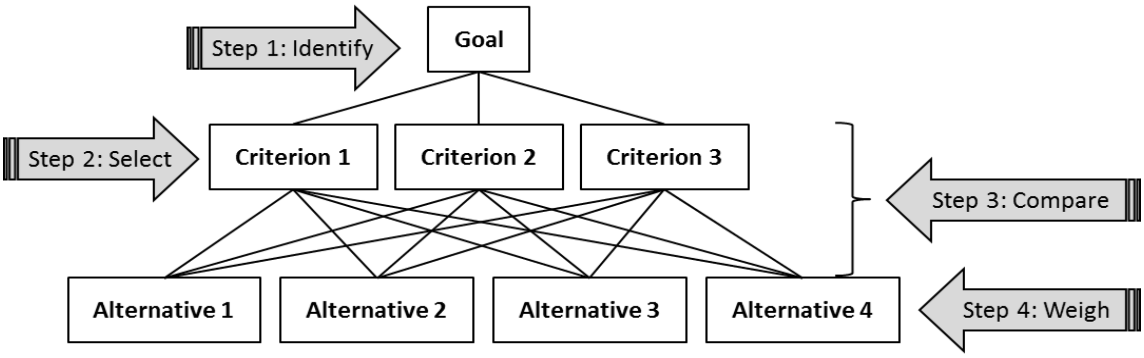


Figure 3.8: A simple Analytic Hierarchy Process hierarchy

When using the Analytic Hierarchy Process, the first thing to do is to define the problem and identify the goal. The second step is to select criteria (and sub-criteria if required) to be evaluated and arrange them hierarchically down to the lowest level, which usually corresponds to the set of alternatives. It is advisable to cluster sub-criteria in order not to increase distortion on the weights, because various authors argued that the more detailed in sub-criteria a given criterion is, the more weight it tends to receive. In the third step, a pairwise comparison is conducted between n criteria in the same level (Figure 3.9a) in order to define weights (w) and using a scale from 1 to 9 (Saaty, 1977). The example in Figure 3.9b shows three criteria and corresponding weights for recruiting an individual. In such an example, academic background is considered slightly more important than experience (three times), whereas it is strongly favoured over certification (seven times).

$$\begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{bmatrix}$$

(a)

$$\begin{array}{rc} & \begin{matrix} \mathit{Academic} & \mathit{Certification} & \mathit{Experience} \end{matrix} \\ \begin{matrix} \mathit{Academic} \\ \mathit{Certification} \\ \mathit{Experience} \end{matrix} & \begin{bmatrix} 1 & 7 & 3 \\ \frac{1}{7} & 1 & \frac{1}{5} \\ \frac{1}{3} & 5 & 1 \end{bmatrix} \end{array}$$

(b)

Figure 3.9: (a) Pairwise comparison matrix, (b) Example of comparing three criteria

After defining the pairwise weights (cf. Figure 3.9b), the geometric mean of each row must be obtained and each value is normalised in order to obtain the proper weight of each criterion, as shown in the third column of Table 3.5.

Table 3.5: Example of calculating criterion weights and lambda max

Criteria and sum	**Geometric mean**	**Normalised value**	**Sum of columns x normalised value**
Academic	2.759	0.649	(1+1/7+1/3) x 0.649 = 0.958
Certification	0.306	0.072	(7+1+5) x 0.072 = 0.935
Experience	1.185	0.279	(3+0.2+1) x 0.279 = 1.172
∑	4.250	1	3.065 (λ_{max})

For each (sub) criteria pairwise comparison, a consistency check is advised in order to fix incoherent judgements. The consistency ratio (*CR*) results from dividing the consistency index (*CI*) by a random index. The *CR* must be below 10% to be an acceptable judgement, although other authors are not in consensus as to the requirement for random indices (Ishizaka and Labib, 2011). A random index depends on the number of criteria, as shown in Table 3.6.

Table 3.6: Random indices according to number of criteria (Saaty, 1977)

N	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Random index	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

CI can be obtained according to Eq. (5). λ_{max} is obtained by initially totalling the weights in each column of the pairwise matrix and multiplying by its corresponding normalised value. Then, the sum of all resulting products constitutes λ_{max}. (cf. last column in Table 3.5).

$$CI = \frac{(\lambda_{max} - n)}{(n-1)} \tag{5}$$

In the given example, *CR* equals to 0.056, which is acceptable. The same procedure is repeated for every set of (sub) criteria. Finally, the priorities obtained on each level weigh the priorities on the following level. This procedure enables calculating the weights of the alternatives (Saaty, 1977).

Due to the number of possible combinations, authors argue that criteria should be reduced by a significant amount. The last step is to determine global priority. However, an additional sensitivity analysis may be performed to see how slight modifications of weights, local priorities and comparisons can impact the results (Ishizaka and Labib, 2011). The Analytic Hierarchy Process provides more benefits than conventional techniques such as Delphi (Vaidya and Kumar, 2006) and is being successfully used in a broad range of applications, including HRM. There are also several software solutions to enable implementation with the Analytic Hierarchy Process. However, a simple spreadsheet template can be used instead (Ishizaka and Labib, 2011; Saaty, 2008). Fuzzy methods also help reduce the subjective impact of the criteria and can be applied with the Analytic Hierarchy Process. Notwithstanding, decision makers still need to perform the pairwise comparison (Güngör et al., 2009; Vaidya and Kumar, 2006).

Authors argue that compared to other Multi-Criteria Decision-making Methods, the Analytic Hierarchy Process incorporates both quantitative and qualitative values in the process. The Analytic Network Process (ANP) is preferred over the Analytic Hierarchy Process when there are intricate connections between criteria in HRM and different levels of decision-making. The former method is an advanced, however complex, version of the latter. Both methods are considered to be the best approaches out of the Multi-Criteria Decision-making Methods (Hor et al., 2010; Ishizaka and Labib (2011).

The Analytic Hierarchy Process received some criticism, such as the order of comparisons among criteria, which supposedly would influence in judging the weights (Webber et al., 1997), the absence of zero in the preference scale (Dodd and Donegan, 1995), the rank reversal problem and even the 1–9 scale. The problem of rank reversal means that the ranking of the result of a given set of alternatives may change by adding or removing one of the alternatives. Despite criticism, the method has been receiving attention in scientific literature through several different usages (Ishizaka and Labib, 2011). This thesis regards the Analytic Hierarchy Process as a suitable method to process qualitative criteria for CS expertise discovery and to rank the suitable experts.

3.3.7 Reputation Systems

Reputation is a viewpoint adopted in this thesis (cf. Figure 2.13) to enable profile co-creation, to measure experience in given capabilities and to encourage interactions based on trust and knowledge exchange within the CS community. Reputation measures an entity's trustworthiness with the aggregation of all the ratings that such an entity receives. Each rating represents an evaluation that a user attributes to another when both users share an event (Swamynathan et al., 2010; Vavilis et al., 2014). The evaluated entity can be an organisation, an individual or a department (Ferris et al., 2007). Trust and reputation can support decisions, although they are not proper decision-making methods (Josang et al., 2007). The purpose of Reputation Systems (RS) in the context of this work is twofold: to ensure trust in profiled information and as a mechanism to evaluate performance and experience.

A RS structure can be either centralised or decentralised (Vavilis et al., 2014). In centralised models, a reputation centre collects all the ratings and publishes an updated reputation score for each participant. This score can help decide whether to negotiate with a given participant (Josang et al., 2007). The main processes and respective elements of a reputation-based trust system are represented in Table 3.7.

Table 3.7: Processes of a reputation-based trust system

Processes	Design considerations	References
Collection	(1) Generation of ratings; (2) types of ratings	(Swamynathan et al., 2010)
Information gathering	(1) Trust information storage, dissemination and search mechanisms; (2) Local control over trust information stored locally on a peer; (3) Credibility of the recommender; (4) Type of behaviour considered; (5) Context dependency	(Koutrouli and Tsalgatidou, 2006)
Storage	(1) the various forms; (2) choice of data; (3) storage methodology	(Swamynathan et al., 2010)
Aggregation	Building a reputation profile from individual user ratings	(Swamynathan et al., 2010)
Reputation estimation	(1) Initialisation of trust information; (2) Scope of trust information (global vs. localised information); (3) Trustworthiness estimation method; (4) Transitivity extent; (5) Recency dependency	(Koutrouli and Tsalgatidou, 2006)
Reputation representation	(1) Range of trustworthiness values; (2) Rank or threshold based; (3) Distrust representation	(Koutrouli and Tsalgatidou, 2006)
Communication	Exchange protocols	(Swamynathan et al., 2010)

Vavilis et al. (2014) presented a framework with requirements (as the desirable characteristics presented on Table 3.8) to evaluate Reputation Systems.

Table 3.8: Requirements for reputation systems (Vavilis et al., 2014)

Group	Requirements
Formulation (information and aggregation method)	Involves the measure and the mathematical model. Ratings should discriminate user behaviour. Reputation should discriminate user behaviour. The Reputation Systems should be able to discriminate "incorrect" ratings. An entity should not be able to provide a rating for itself. Aggregation of ratings should be meaningful. Reputation should be assessed using a sufficient amount of

Group	Requirements
	information. The Reputation Systems should differentiate reputation information by the interaction it represents. Reputation should capture the evolution of user behaviour.
Fair treatment of new users	Users should not gain advantage of their new status. New users should not be penalised for their status.
Calculation and dissemination	Involves the algorithm to assess reputation and ratings among users.
	Users should not be able to modify ratings directly. Users should not be able to modify reputation values directly. Users should not be responsible for calculating their own reputation directly.

The formulation group can be categorised in three types of rating scales. Firstly, a frequency scale determines the occurrence of behaviour (e.g. “always”, “seldom” and “never”). Secondly, an evaluation concept scale delivers subjective ratings (e.g. “not satisfactory” and “outstanding”). Thirdly, a standard type of scale measures behaviour according to a standard (e.g. “below”, “meet” and “above” a standard) (Kanij et al., 2014). The latter scale is adopted in this thesis to rate experience in CS projects, due to its higher objectivity. Moreover, Table 3.9 suggests some features that should support each requirement (cf. Table 3.8). It is worth noting that the importance of requirements, features and technical solutions vary between different application domains (Vavilis et al., 2014).

Table 3.9: Features and technical solutions (adapted from Vavilis et al., 2014: p. 149)

Feature	Description
Trust / distrust	Reputation metrics should consider the whole spectrum of both concepts
Absolute reputation values	Use of absolute values instead of relative ones (such as ranking users)
Origin / target	Professionals must have the same expertise in order to be evaluated accordingly
(un) Certainty	“Confidence on trust information”
Interaction context	Attribute weights to evaluations according to costs
Timestamp	Indicates the time in which an interaction occurred. Recent transactions should have greater weight than older interactions (Koutrouli and Tsalgatidou, 2006)
Transaction proofs	Technical solution that assures that the interaction actually happened
Privacy	“(…) ability to be anonymous and not to let other peers monitor its transactions and recommendations. However, for a reputation system to work, information about a peer’s identity, its transactions and provided recommendations needs to be monitored” (Koutrouli and Tsalgatidou, 2012: p. 64). “For security, privacy and management reasons, we assume that every peer maintains the attributes of the users associated to his domain.” (Swamynathan et al., 2010: p. 243)

Reputation Systems are considered a Talent Management practice (Tarique and Schuler, 2010); however, there is still limited research on trust-base choice mechanisms upon people selection (Hu and Wang, 2014) and on how to develop and measure HR reputation and integration with social networks (Ferris et al., 2007). Reputation is being used in social media to measure knowledge creation based on contributions and interactions (e.g. ResearchGate, 2014); notwithstanding, it is not clear if their algorithms comply with the aforementioned Reputation Systems requirements. Collaborative filtering systems (often called recommender systems) are similar to Reputation Systems in terms of collecting ratings. However, ratings of recommender systems are subject to taste, thus providing subjective results. Hence, collaborative filtering can also be used as an alternative to Reputation Systems when using subjective inputs (Josang et al., 2007). In this thesis, a recommender system is adopted to rate knowledge exchange interactions between peers with subjective ratings. Reputation Systems have some drawbacks which should be considered before adoption. First, authors suggest that personal reputation should only be considered when there is a vacuum of required information in a user’s profile (Ferris et al., 2007). Second, there is a wide variety of attacks to which Reputation Systems are highly susceptible. Some of them are common to most IS, such as denial of service (DoS), whereas others specifically attack the system’s trustworthiness (e.g. providing biased information), which cannot be detected by traditional centralised security solutions. Third, works focusing on the credibility of Reputation Systems are still at an early stage (Josang et al., 2007; Koutrouli and Tsalgatidou, 2012). Fourth, reputation is compared according to the same expertise (cf. Table 3.9) in contrast with comparing multiple criteria in user profiles. Moreover, this aspect seems to hinder the use of reputation scores retrieved from external social media, since it may input bias in information analytics, given that population and reputation criteria are not the same. Finally, some SNS use confirmed data (e.g. certification number fields and digital object identifier) instead of relying on Reputation Systems. Such solutions build profiles by retrieving information from verifiable sources (Gewin, 2009), relying on ontology reasoning to enhance semantics and assuring trust without using subjective evaluations.

Since reputation plays an important role in the CS community, an algorithm compliant with the aforementioned features and requirements is adopted in order to serve the research solution. A Reputation System for reciprocal evaluation was developed to incorporate incentives for honest bids, where each user holds a buyer and a seller reputation, each ranging from 0 to 1. A rating is attributed after each transaction, when one user plays the role of buyer and the other one is the seller. The algorithm considers positive and negative ratings and weighs bids according to its recency and the reputation of the evaluator. Negative bids have a deeper impact than positive bids in order to incentivise honesty. Ratings are used to calculate scores, which determine a user’s reputation. For example, *TSS* in Eq. (6) refers to the total weighted reviews that a user *A* would have in case of receiving only positive ratings when performing the role of seller. *SS* represents the actual weighted evaluations that *A* received as a seller in Eq. (7). These two variables are used to compute the current *SR* for *A* in Eq. (8). The buying reputation (*BR*) is similarly computed according to a total buying score (*TBS*) and to a buying score (*BS*) (Lin et al., 2015). Despite being focused on e-commerce, such a Reputation System is adopted in this thesis aligned with the CS context. The algorithms are thus adapted to calculate evaluations from peers and employers (instead of buyers and sellers) and *M* assorts a rating gradation instead of the amount of trading.

$$TSS_A^{t+1} = TSS_A^t + BR_B^t \times M^t \tag{6}$$

where
TSS: the total selling score, which initial value is 1
A: the selling party
B: the buying party
BR: the buying reputation, scored between 0 and 1
t: the time before rating occurs
M: the amount of trading

$$SS_A^{t+1} = \begin{cases} SS_A^t + BR_B^t \times M^t, if\ F^t \geq 0,\ \ else \\ SS_A^t + (-1/100) \times TSS_A^t \times BR_B^t \times M^t \end{cases} \tag{7}$$

where
SS: the selling score, which initial value is 1
F: the feedback rating

$$SR_A^t = \alpha_A^t \times SPR_A^t + (1 - \alpha_A^t) \times SSL_A^t \tag{8}$$

where
SR: the selling reputation, scored between 0 and 1
$\alpha_A^t = 1/(1 + exp^{-(\log TSS_A^t - k)})$,
$SPR_A^t = SS_A^t / TSS_A^t$,
$SSL_A^t = 1/(1 + exp^{-(\log SS_A^t - k)})$
k: an exponential smoothing variable

Due to its compliance with the Reputation System principles, this thesis adapts the algorithm presented by Lin et al. (2015) in such a way to enable experts to receive evaluations from peers and contractors.

3.3.8 Hybrid applications of Information Analytics

This section discusses the applications of some analytical techniques with the purpose of comparing and contrasting different approaches suitable for Talent Management viewpoints that are required to promote CS expertise discovery. Moreover, strengths and weaknesses of such techniques and their combinations are discussed. Combining Multi-Criteria Decision-making Methods may reduce some of their drawbacks. For example, Analytic Hierarchy Process/Data Envelopment Analysis combinations can provide mathematical evaluation of the weights for Data Envelopment Analysis and enable ranking the Decision-Making Units (Sinuany-Stern et al., 2000). Data Envelopment Analysis with an assurance region is suggested as an alternative using Analytic Hierarchy Process to generate weights for Data Envelopment Analysis for recruiting personnel (Wang et al., 2008). However, all qualitative and quantitative data are treated as the same because the processes focus on drawing the weights and ranking. Most applications of Analytic Hierarchy Process/Data Envelopment Analysis in HRM focus on ranking Decision-Making Units and calculating organisational performance (Feng et al., 2004; Tseng and Lee, 2009). Although there is not much research using Data Envelopment Analysis for talent selection, Data Envelopment Analysis is an adequate method to obtain applicants' relative performance. Even when input or output data from certain Decision-Making Units may be not exactly known, it still is a powerful technique to calculate relative effectiveness (Lin, 2010).

In contrast to the aforementioned approaches, Analytic Hierarchy Process and Data Envelopment Analysis can be used as complementary methods dealing with qualitative and quantitative data separately. This approach focuses on optimising decision-making based on the nature of data rather than combining methods. Such an approach offers flexibility for dealing with different types of data stored in profiles. Mushi (2012) and Sun et al. (2014) used CBR, Analytic Hierarchy Process and Data Envelopment Analysis to support decision-making. Sun et al. (2014) have studied the participation of patients in the decision process of healthcare service provision. A request for healthcare service provision triggers a new case C^{new} comprising a set of criteria {*diagnosis, personalised feature, weight, pf value*}. A case-based engine retrieves previous cases containing the same set of criteria with values within a given threshold using a weighted similarity algorithm. If more than one case is retrieved, the requester is able to choose one from the matching list. If just one case is retrieved, it is automatically selected. If no case is retrieved within the threshold, a new case is created and saved on the requester's profile. The Data Envelopment Analysis is used to select the efficient options using objective parameters, whereas the Analytic Hierarchy Process considers subjectiveness from patients' preferences to rank the efficient Decision-Making Units using qualitative data. In such research, Data Envelopment Analysis is used before the Analytic Hierarchy Process because there is a larger available amount of quantitative data compared with qualitative data (Sun et al., 2014). Fontenele and Sun (2016) adopted a similar approach with Analytic Hierarchy Process and Data Envelopment Analysis, which is tailored to select CS experts, and includes a Reputation

System for evaluating expertise. These evaluations eventually update quantitative values for selecting experts. The successful combination of CBR, Data Envelopment Analysis and Analytic Hierarchy Process in the report by Sun et al. (2014) and Data Envelopment Analysis, the Analytic Hierarchy Process and Reputation Systems in Fontenele and Sun (2016) inspired the adoption of those methods for the research solution, although Exemplar-Based Reasoning is adopted instead of Case-Based Reasoning. In this thesis, both quantitative and qualitative criteria are therefore capable of being used for selecting experts to work in CS projects, which are refined with criteria from successful projects.

Finally, the literature provided different combinations of using Multi-Criteria Decision-making Methods and ontology reasoning, as introduced in the report by Colucci et al. (2011) to support decision-making for expertise discovery. Table 3.10 summarises the approaches discussed in the report by Šaša Bastinos and Krisper (2013) and draws on Multi-Criteria Decision-making Methods discussed on previous sections. Although requiring additional consistency checks, balancing Multi-Criteria Decision-making Methods with ontology seems more suitable when the subjectiveness of the decision maker should be taken into consideration.

Table 3.10: Decision-making approaches (extended from Šaša Bastinos and Krisper, 2013)

Features	Plain MCDM	MCDM with ontology	Ontology reasoning
Automation	Several methods and combination of methods available in the literature	Partial automation	Full automation
	Manual information retrieval; however, there are tools for some methods	Retrieves information from ontology	Ontology is used for information retrieval and reasoning
Complexity	Complexity depends on the use of the adopted method	Complexity depends on the use of the adopted method	Adds complexity to the system (restricted number of criteria and alternatives)
Objectivity vs. subjectivity	May use both quantitative or qualitative data	May use both quantitative or qualitative data	Objective and quantitative results
	Decision has a level of subjectivity	Decision has a level of subjectivity	Decisions are objective and do not rely on decision maker
Flexibility	Multi-Criteria Decision-making Methods are flexible to solve a wide variety of decision problems	Multi-Criteria Decision-making Methods are flexible to solve a wide variety of decision problems	Reasoning engines do not allow evaluation in some information structures
	Should avoid more than three subcriteria and large qualitative values per criterion	Should avoid more than three subcriteria and large qualitative values per criterion	
Consistency	Requires additional	Requires additional	Provides consistent results

Features	Plain MCDM	MCDM with ontology	Ontology reasoning
	consistency checks	consistency checks	

3.4 Summary

This chapter explored methods to represent, to source and to process the knowledge required to support expertise discovery. The investigation concentrated on best practices and state-of-the-art approaches matching the viewpoints in order to address the first research question.

EA and semiotic-based ontology were presented as alternatives to model the complex, unstructured and dynamic scenario where expertise discovery takes place. EA is capable of portraying high-level organisation features, motivational concepts, business-oriented requirements and services related to inter-organisational strategic Talent Management. Archimate, in particular, delivers a user-friendly interface to perform requirements elicitation and analysis, which is particularly useful when dealing with the complex CS scenario. On the other hand, semiotic-based ontology can help overcome limited semantic expressiveness in EA languages by describing norms that govern Archimate behaviour concepts (e.g. processes). By enabling heavyweight ontology, OS enables the formalising of axioms in order to conceptualise behaviours and workflows in CS expertise discovery. OS, however, lacks the representation of underpinning motivational concepts. Thus, Archimate emerges as a useful complementary knowledge representation method to approach stakeholders when performing requirements analysis. Such an approach facilitates identifying and structuring critical services for the CS community, which is regarded as part of the research problem.

The CS expertise can contain both qualitative (e.g. competencies and certifications) and quantitative data (e.g. number of concluded projects and years of experience in a given subject). Currently, mechanisms capable of processing such type of criteria are rather unstructured, manual or empiric. Therefore, Analytic Hierarchy Process and Data Envelopment Analysis were introduced as versatile Multi-Criteria Decision-making Methods to deal with finding the proper expertise in an effective fashion, which is crucial to the research problem.

The lack of awareness and review mechanisms for expertise, along with the fast changing knowledge domain, were identified as part of the research problem. Hence, an application of performance metrics with visualisation methods seems appropriate to enable the necessary feedback to the CS community. The research problem also states that currently there is no learning from the rich and dynamic CS knowledge domain. In addition to the learning experience provided by visualisation methods, performance metrics and feedback, this thesis adopts information retrieval and Exemplar-Based Reasoning. Both techniques may prove to be valuable tools to suggest relevant capabilities based on successful expertise discovery cases. CS projects can therefore be stored as a case for further reuse. An application of information retrieval on DW addresses part of the second research question by creating a metric based on the frequency of capabilities in both expert and project profiles. Such a metric is used to generate KRI, which enables raising of awareness by identifying changes and measuring shortages and gaps in the dynamic CS expertise environment. These functionalities address the third research question. In addition, these indicators may be useful in supporting KPI development for expertise discovery, although such a feature is beyond this research scope.

After a critical discussion on Reputation Systems, using a combination of confirmed data with reputation measures can improve information reliability in co-created expert profiles over time. In addition to addressing trust issues in the current CS scenario, Reputation Systems are adopted as a mechanism to promote knowledge exchange in order to raise awareness, which is part of the research problem.

Among the latest applications on information analytics, no work properly addressing the research problem was found. Instead, the literature provided the building blocks to pave the foundations to develop a method that efficiently supports expertise discovery for the CS community.

Chapter 4

Research Methodology

This research seeks to devise a method to discover expertise for CS. Research should be grounded in an appropriate set of paradigms, methods and techniques for scientific validity. This chapter conducts a review of the research methodology in order to provide awareness of the methodological options. It first presents an investigation about research paradigms, designs, methods and techniques and their impact on the IS field. Then, it discusses and selects the proper approaches according to the research theme and context.

4.1 A Review of Core Content in Research Methodology

This section settles theoretical differences among research paradigms, designs, methods and techniques through a brief review of such concepts drawn from discussions of meaningful works.

4.1.1 Research Paradigms and Philosophical Groundings

All research is based on the underlying assumptions for suitability and validity (Myers, 1997), which are also called scientific paradigms. Scientific paradigms relate to universally accepted models derived from examples of scientific practice (Burrell and Morgan, 1979; Kuhn, 2012). According to (Guba, 1990), a paradigm is “a basic set of beliefs that guides action”. The transformation of paradigms via scientific revolutions is the typical pattern of science evolution. Studying paradigms enables individuals to share the same standards, views and practices within a scientific community (Kuhn, 2012). This section introduces the philosophical foundation on which the research paradigms are based.

4.1.1.1 Philosophical Groundings on Research Paradigms

The scientific paradigm that is chosen by a researcher guides the method of thinking and acting throughout the research as well as the nature of research itself (Ousmanou, 2007). Guba and Lincoln (1994) state, however, that researchers should clearly know their guiding paradigm before starting their research. According to Lincoln and Guba (1985), Guba and Lincoln (1994) and Vaishnavi and Jr (2007), any research should have underlying philosophical assumptions about the following topics:

- Ontology (the nature of reality, i.e. "what is real and what is not?", "what is fundamental and what is derivative?");
- Epistemology (exploring the nature of knowledge, i.e. "on what does knowledge depend?", "how do we know things?");
- Methodology (in what way can one obtain such knowledge) and
- Axiology (the study of values, i.e. "what values does an individual or group hold, and why?").

Ontological beliefs relate to the essence of the phenomena under investigation. Such beliefs might be either objective or subjective. Objectivists believe that the empirical world exists independently of any human observation (Harris, 2011; Liu, 2000; Orlikowski and Baroudi, 1991). Subjectivists believe that the existence of reality depends on the actions of humans (Orlikowski and Baroudi, 1991), and is a knowledge product of individual consciousness, which requires a "knower" entity (Liu, 2000). Furthermore, objectivist beliefs tend to provide a better support for natural sciences rather than to social sciences, and conversely subjectivist beliefs tend to provide a better support for social sciences due to the nature of studied phenomena (Harris, 2011; Liu, 2000). Epistemological beliefs concern the issues that contribute to create and validate knowledge about a phenomenon. These criteria relate to a given research paradigm's notion of valid knowledge. The methodological beliefs relate to the adequate set of research methods and techniques for data collection (Orlikowski and Baroudi, 1991).

A research framework for conducting research studies is described in Figure 4.1. The figure shows that the options on each level influence the choices on other levels; however, there are no strict correlations among them.

Guba and Lincoln (1994) suggested four qualitative-related research paradigms: positivism, post-positivism, critical theory and constructivism. Vaishnavi and Jr (2007) added the design science as an approach loosely based on interpretivism. Previously, Orlikowski and Baroudi (1991) suggested three categories relating to research epistemology: positivist, critical (or postpositivist) and interpretive (also known as antipositivism).

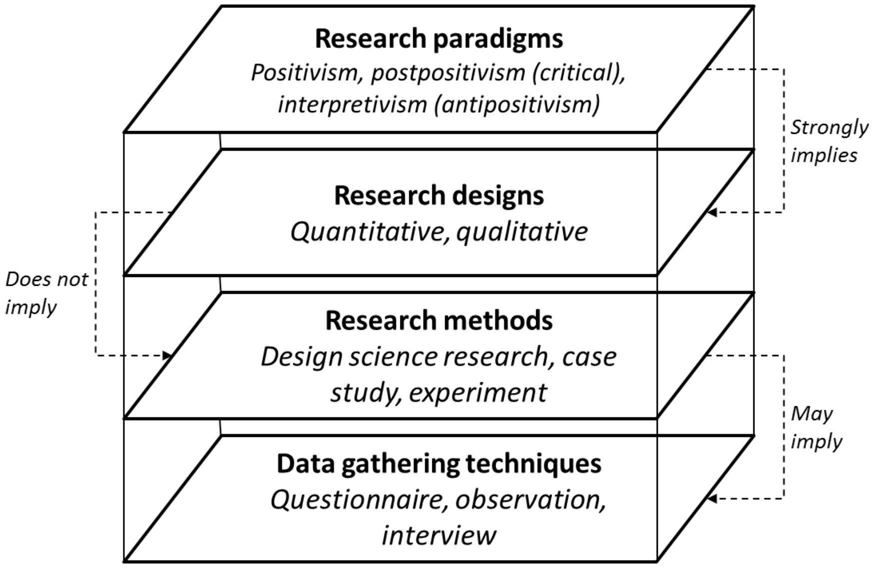


Figure 4.1: Research framework for conducting Research Studies (adapted from Pickard, 2002: p. 10)

Positivism and interpretivism seem to adopt opposite stances. While the former is usually quantitative and rely on natural science methods, the latter is usually qualitative and urges the necessity for methods other than those of natural science (Gable, 1994). Indeed, there is, most of the time, no clear distinction between “interpretive” and “qualitative” research. However, Klein and Myers (1999) remark that qualitative research can still be carried out with either philosophical stances.

Each research paradigm has an underlying philosophical stance. Table 4.1 comprises ideas from Lincoln and Guba (1985), Orlikowski and Baroudi (1991), Gable (1994), Vaishnavi and Jr (2007) and Harris (2011). Although Design Science Research is not fully accepted as a research paradigm (Weber, 2010), it also can be analysed under such stances.

Table 4.1: Characteristics of major research paradigms and design science (adapted from Ousmanou, 2007: p. 19)

Stance	**Positivism**	**Postpositivism (Critical)**	**Interpretivism (Antipositivism)**	**Design Science Research**
Ontological	Realism (single reality, knowable, probabilistic)	Critical realism	Relativism, multiple realities, socially constructed	Multiple realism, socio-technologically enabled
Epistemological	Objectivist, dualist, empirical testability of theories	Objectivist, modified dualist	Subjectivist, transactional (knowledge emerges from interaction)	Knowing through making, iteration reveals meaning, improving the world through intervention
Methodological	Quantitative (statistical)	Quantitative and qualitative	Qualitative (participation, hermeneutical)	Qualitative exploration and quantitative confirmations
Purpose	Prediction/ control, explanation / verification	Generalisation, falsification	Transfer of findings	Developmental (measure artefactual impacts on the system)

4.1.1.2 Positivist Research

Positivism is the mainstream research paradigm (Harris, 2011; Orlikowski and Baroudi, 1991) and assumes "the existence of an objective, independent and stable reality, which is available for discovery and analysis" (Ousmanou, 2007: p. 19). This paradigm only considers observable phenomena; therefore, even social facts are considered as being independent from human interactions. The observer can maintain an isolated stance from what is being observed and draw annotations upon the object of research (Liu, 2000; Ousmanou, 2007). Dualism underlies positivism through its epistemological stance (as presented in Table 4.1) and relates to two entities that are completely independent from each other (e.g. an observer and the target of observation) (Harris, 2011). Because of such independence, the substitution of the observer should not affect the subject, which means that the research process is replicable. Positivist research normally begins with hypothesis identification. Afterwards, the hypotheses are empirically tested using structured experimentation and statistics. Positivist research usually applies the quantitative approach (Ousmanou, 2007).

4.1.1.3 Interpretivist (or Antipositivist) Research

Interpretivism comprises two streams: empirical interpretivism and critical theory (Ousmanou, 2007). The former aims to show aspects of the interpretive structure (such as empirical conditions and nature of relationships), how experience and knowledge can be derived from such a structure (Lavine et al., 2011) and includes constructivism, semiotics and phenomenology. The latter debates about ideology in social structures (Ousmanou, 2007). Phenomenology emphasises direct observation of phenomena (Bernard, 2000).

Interpretivists are also known as antipositivists because such a philosophical stance emerged in contradiction with the long established positivism (Harris, 2011). Interpretivists adopt a holistic view by acknowledging multiple realities (Ousmanou, 2007) and the access to these realities is only achieved through social constructions (i.e. language, consciousness and shared meanings) (Myers, 1997); thus, the term "constructivism" also refers to interpretivism (Harris, 2011). Relativist ontology (interpretivism) is opposed to dualistic ontology (positivism), because the former believes that there cannot exist a neutral observer, whereas the latter adopts the independent observer and observed subject stance (Harris, 2011).

Interpretive research aims to discover what things are rather than how they function (Ousmanou, 2007). It does not predefine variables, rather focusing on the complex human nature as the situation evolves (Myers, 1997). Because of the highly complex relationships between phenomena and events (phenomenological position), interpretivists seek to identify patterns within such relationships, rather than analysing separate parts of the world, as positivists would do (Harris, 2011; Liu, 2000).

Context is defined as “a kind of container in which the phenomenon resides” (Dervin et al., 2003: p. 112). Dervin et al. (2003) argued that without context, it is not possible to understand any sort of human behaviour. By considering the context and multiple points of view, interpretivism can explain meanings underneath actions of individuals. Epistemologically, there is a transactional relation between the “knower” and the “known” because they influence each other. This relationship results in the creation of time and context bound knowledge (Lincoln and Guba, 1985). Additionally, Orlikowski and Baroudi (1991) argued that interpretive studies seek to understand how members of a social group connect their particular realities and draw meaning and social action from this enactment.

4.1.2 Research in Information Systems

This section presents Information Systems (IS) under the perspective of traditional research paradigms and methods. In addition, some development paradigms specific to IS field are introduced.

4.1.2.1 Purpose of Information Systems Research

IS development is a multi-disciplinary subject that draws primarily from systems theory and from other disciplines relevant to application domains, such as management science, organisation theory, sociology and computer science (Harris, 2011; UKAIS, 2014). Hirschheim et al. (1995) defines IS under functional and structural perspectives. The structural perspective views IS as a set of organisational components (such as people, processes, and technology) that serves a given purpose, while the functional perspective views IS as a technological-based platform to support “recording, storing and disseminating linguistic expressions” (Hirschheim et al., 1995: p. 11).

IS uses applied research to contribute along with building knowledge to the support of the use of ICT in organisations (Liu, 2000) to understand the possibilities, choices and consequences of ICT usage. Subjects of analysis can be individuals, groups of people and organisational components (Harris, 2011). IS research has the theoretical objective to expand the discipline’s knowledge frontiers and the practical objective of providing useful solutions to organisations and society (Ousmanou, 2007). The discipline of IS can be associated with either positivist or interpretivist paradigm (Klein and Myers, 1999).

4.1.2.2 Information Systems Development Paradigms

Hirschheim et al. (1995) proposed another approach to research paradigms focusing in IS development. Such a proposal is based in Burrell and Morgan (1979) and draws four perspectives from assumptions about reality (ontology) and knowledge (epistemology). An order-conflict dimension defines the former and a subjective-objective dimension depicts the latter. Both dimensions are represented as orthogonal axes that generate the new paradigms in the four resulting quadrants illustrated in Figure 4.2.

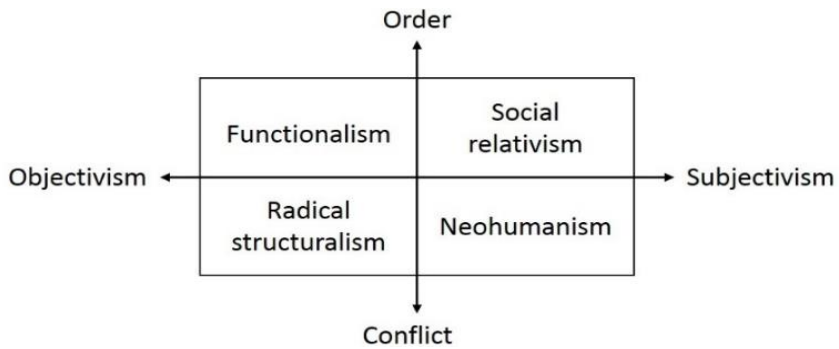


Figure 4.2: Information Systems development paradigms (Hirschheim et al., 1995: p. 48)

The functionalist paradigm aims to provide explanations of the status quo of integration and of reality. It analyses the components of the system individually to acknowledge how the whole system works. The social relativist paradigm seeks to study the system as a whole and adopts elements of the interpretive stance. In such a paradigm, the researcher is an actor of the system, rather than being a passive observer. The radical structuralist paradigm is concerned with a constantly evolving scenario and analysis of economic power relationships. Finally, the neohumanist paradigm looks for radical changes in the system as a whole, focusing on all forms of obstacles to emancipation and evolution (Hirschheim et al., 1995).

4.1.2.3 Other Approaches to Research Paradigms

There are three forms of research: problem-solving, exploratory and testing-out. Problem-solving starts with a problem in the real world and the research gathers all the available resources in order to develop a method to solve it. While exploratory research deals with new problems about which little is known, testing-out research tries to "find the limits of previously proposed generalisations" (Phillips and Pugh, 2000: p. 51).

Rule-based approaches for data modelling relate to the social relativist paradigm. Such approaches are tightly connected to Constructivism, in which reality is created according to interactions of agents in a social environment (Hirschheim et al., 1995). Aligned with rule-based data modelling, the radical subjectivist paradigm sees reality as a construct of agents' behaviour. When people interact with the world, this leads to the emergence of new beliefs and behaviours; therefore, reality is constantly shifting. These behaviours are represented by a system of norms; thus, knowledge and reality are dependent on agents. This paradigm is governed by two axioms: "there is no knowledge without a knower" and "there is no knowing without action" (Liu, 2000: p. 26).

4.1.3 Research Design

Most authors use the term "research methodology" instead of "research design". However, to avoid ambiguity with the multiplicity of methodology meanings (such as the philosophical stance or the set of research paradigm, research methods and techniques adopted in a given research), this work adopts the latter term.

Research design is the theoretical perspective of the research from which derives the general nature of research activities (Ousmanou, 2007). There are two basic types of research design: the quantitative and the qualitative approaches.

Quantitative design, also known as the "scientific approach", relies on the broader principles of positivism, explains phenomena using metrics and was originally developed to study natural phenomena (Harris, 2011; Myers, 1997).

The qualitative design seeks to comprehend subjective perceptions of reality by focusing on insights instead of statistical results; thus, its methods tend to dominate in social science research (Harris, 2011). This approach helps researchers to understand social and cultural contexts while much of such context should be lost if one was to use quantitative data (Myers, 1997).

Most critics of qualitative design come from the physical sciences. The incapacity to manage independent variables, incorrect conclusions and low randomisation possibility constitute the major drawbacks of the qualitative approach (Gable, 1994). Although reflecting (in part) distinct research paradigms, qualitative and quantitative designs may complement each other and can be used together (Myers, 1997).

4.1.4 Research Methods and Techniques

A research method is a strategy to perform empirical investigation (Ousmanou, 2007). It influences the way to collect data and reveals the researcher's suppositions, abilities and practices. There are several types of research methods, such as action research, case study, survey, grounded theory, ethnography, experiment (Myers, 1997; Yin, 2009), formal methods, numerical (Myers, 1997), history and analysis of archival records (Yin, 2009). Such methods can be combined in a so-called mixed method research, which enables researchers to address complicated research questions and to acquire a stronger array of data by dealing with both quantitative and qualitative data. On the other hand, mixed methods are normally more difficult to conduct than single methods (Yin, 2014). Methods are not defined by the employed techniques for data gathering, but by the main purpose of the research. The following subsections introduce techniques for data collection and some research methods.

4.1.4.1 Data Gathering Techniques

Myers (1997) argued that some research methods were developed within either natural sciences (e.g. survey, experiment, formal method and numerical method) or social sciences (e.g. action research, case study and ethnography). However, Yin (2009) suggested that each research method can be used for either exploratory, descriptive and explanatory purposes. Bernard (2000) added that the research methods are not "owned" by any sort of discipline as long as it is coherent with the adopted paradigm.

Each method uses one or more techniques to collect empirical data for the research, although techniques are not strictly bonded to a specific method. During the problem-solving stage, the researcher has to answer questions such as "what kind of data do I need to address the research question?", "where can I collect the data that I need?", "which instruments and procedures are available?" and "how can I collect such data?" (Ousmanou, 2007: p. 31). This work discusses interviews, questionnaires and observations.

Conducting interviews is a very common and established qualitative data gathering technique that can assume many stances depending on researcher's experience, research question and participants' availability. It can be tailored to participants, such as when questions are too complex to formulate, and verbal responses are easier for participants rather than written answers. Interviews can also be used to validate data or information collected with other techniques (Ousmanou, 2007). Depending on their rigour, interviews are classified as highly structured/standardised, semi structured or unstructured/informal. When directed to a group of people experts, interviews are called focus groups. In this case, data is primarily obtained from the interactions of interviewees. Data quality is substantially enriched by constructivism. However, specific skills are required to conduct group discussions (Merriam, 2014). For example, the interviewer should drive discussions within the group and should be able to capture knowledge from these discussions.

Distributing questionnaires is another popular technique with some advantages (such as low cost, larger samples, anonymity for participants and replicability) and drawbacks (such as low response rate, bias, and simplicity of questions and lack of opportunity to clarify questions). Questions should be easily understood in order to avoid misinterpretation (Foddy, 1993).

Observation is a technique that requires watching and recording notes of the behaviour of the studied subject. There are two major remarks for a researcher who intends to use observation as a technique. Firstly, one might be concerned to get so involved to eventually lose objectivity. However, this is not a problem when objectivity is not desired in the research, especially in that which is socially driven. Secondly, one might think that it is impossible to be an observer completely detached from the observed context. In this case, the researcher should try to minimise the impact on participants by adopting a more subjective approach (Ousmanou, 2007).

4.1.4.2 Design Science Research

Design science (or "science of the artificial") research (DSR) is fit to ICT research because of the artificial and complex nature of the questions in ICT, which may arise from a non-existent background and from a variety of sources. Design Science Research emerges from interpretivism and aligns with constructivism. In Design Science Research, knowledge emerges from an iterative process of problem awareness, suggestion, development, evaluation and conclusion. A suggestion of a first approach to address the problem can be conducted using preliminary quantitative and qualitative data collection and analysis. Then, iterative development of an artefact is conducted. Finally, empirical methods (e.g. experiments) can be used for evaluation and the research concludes as an opportunity for further improvement using insights or suggestions. Such an iterative process is known as the general design cycle. Thus, the outcome of Design Science Research is a functional artefact as a solution to a problem in the real world, which still can be refined after being delivered (Vaishnavi and Jr, 2007).

Due to its developmentalist nature, Design Science Research often delivers meaningful contribution to practice. On the other hand, the approach has received critiques due to the challenge to contribute to theory. Authors have argued that a pluralistic approach of Design

Science Research with other paradigms and methods that address such shortcomings and ensures a theoretical contribution to the research domain (Weber, 2010). Design Science Research is similar to action research; however, the time frame of the former is greatly shortened in terms of group interactions, which are typical in the latter (Vaishnavi and Jr, 2007).

Design Science Research relies upon both qualitative and quantitative designs. The range of possible artefacts comprise "technological rules, technical capabilities, constructs, conceptual designs, models, methods and instantiations, such as prototypes or commercial products" (Harris, 2011: p. 21). An artefact developed under Design Science Research is guided by the use of patterns, usually defined as "a solution to a problem in a recurring context" (Vaishnavi and Jr, 2007: p. 58). This assertion supports that Design Science Research is rather goal-oriented than strictly algorithmic and delivers "a feasible" solution rather than "the only possible" solution. Those patterns (e.g. complex system analysis and others for evaluation and validation) drive the development of each stage of the iterative processes in Design Science Research.

4.1.4.3 Case Study

Case studies are built around the research questions (Harris, 2011) and can be defined as "an empirical enquiry that investigates a contemporary phenomenon ("the case") in depth and within its real-world context, especially when the boundaries between phenomenon and context may not be clearly evident" (Yin, 2014: p. 16).

Case studies enable the researcher to combine different methods for harvesting data and develop a qualitative or quantitative analysis (Ousmanou, 2007). Yin (2014) suggests that a case study is the preferred method to (1) answer "how" and "why" questions; (2) when the researcher has little control over events; and (3) focus on contemporary events in real-life context. In contrast with other qualitative methods (e.g. ethnography and grounded theory), a case study requires building a preliminary theory. Although it is possible to be adopted in either type of research design, the case study emphasises the use of qualitative analysis and implies collecting data through direct observation, well-defined interviews and integrating multiple fields while trying to understand the research problem. On the one hand, it captures the richness of the scenario, whereas on the other hand, results might be too specific to an organisation, thus not generalisable. In addition, lack of controllability, repeatability and generalisability are weaknesses in this method that can be surpassed and that in any event may appear in other research methods (Gable, 1994).

There is also a main difference between surveys and case studies. In surveys, questioning must be structured and consistent in order to permit response grouping and to produce generalisation from a sample population, whereas case studies are concerned with a given perspective (Ousmanou, 2007)

There are many advantages to using the case study method in IS. First, the researcher can learn the state of the art in a natural environment and create theories from practice. Second, it enables the researcher to comprehend the complexity of the processes. Finally, new subjects in IS discipline provide valuable insights (Gable, 1994). Because IS deals with "social, organisation and cultural issues", Ousmanou (2007) argues that the case study method is adequate when the researcher seeks holistic and profound study of phenomena from the involved stakeholder's point of view. It is worth noting that the case study method can be also adopted under a Design Science Research approach (Vaishnavi and Jr, 2007). This empirical research draws elements from case study, although it does not comply with case study in the traditional sense.

4.1.4.4 Simulation and Experimentation

Both simulation and experimentation are regarded as methods suited for the validation and evaluation of research solutions for complex problems not susceptible to mathematical proof.

Simulation is appropriate for time consuming, costly or unfeasible real-life settings. Such a method involves three steps. Firstly, a conceptual model representing the problem should be developed. Secondly, an initial dataset should be created to test the model. The dataset must relate to the goals of the artefact and to the environment where the artefact should function. Thirdly, the model should be exercised and simulated by using some software or with some amount of programming. The performance data should be analysed in light of the solution goals. Considering that the simulation represents real-life situations and that data analysis complies with the solution's goals, the solution is validated. Experimentation can be employed as a solution development (e.g. a model, prototype or system) for hypothesis testing. Data generated from the solution is used to validate or reject the hypotheses, either by developing or testing the system under varying environments (Vaishnavi and Jr, 2007). Experiments can also be used to answer "how" and "why" kinds of research questions. The main difference between experiments and case studies is that the former method is recommended when control of behavioural events is required (Yin, 2014).

4.2 Research Methods Adapted by this Study

Aligned with the social relativist paradigm, the radical subjectivist approach enables a viable perspective to understand the views and actions of different stakeholders in the complex CS scenario. Knowledge is derived from actions in the social environment (Hirschheim et al., 1995). In compliance with those approaches, CS experts improve their capabilities according to interactions with peers and work experience gained in projects. Considering the reviewed research paradigms and methods, experiment methods guided by the empirical interpretive paradigm and Design Science Research seems to be the most appropriate choice because of the strong sociological context in this work. Such mixed methods approach strengthens both theoretical and practical contributions. Although interpretivism naturally leads to the choice of the qualitative design, this research also uses elements of quantitative approach. This is due to the use of socio-technical fields for information analytics and the use of statistics and the use of algorithms. This is a problem-solving and empirical research, because it begins with the awareness of a problem in the real world and gathers the resources to solve it. The complexity of the research problem claims for an iterative development in order to assure compliance with user requirements. Thus, research draws from Design Science Research approach, since an artefact (i.e. a method) is developed through knowledge gained from interactions with a CS community. This section ultimately describes how the thesis addresses the research objectives.

4.2.1 Overview of Approach for Research Solution

A Method for Discovering Expertise in Cyber Security communities (DECYSE) is devised according to Figure 4.3. The solution design is iteratively developed throughout the whole research process.

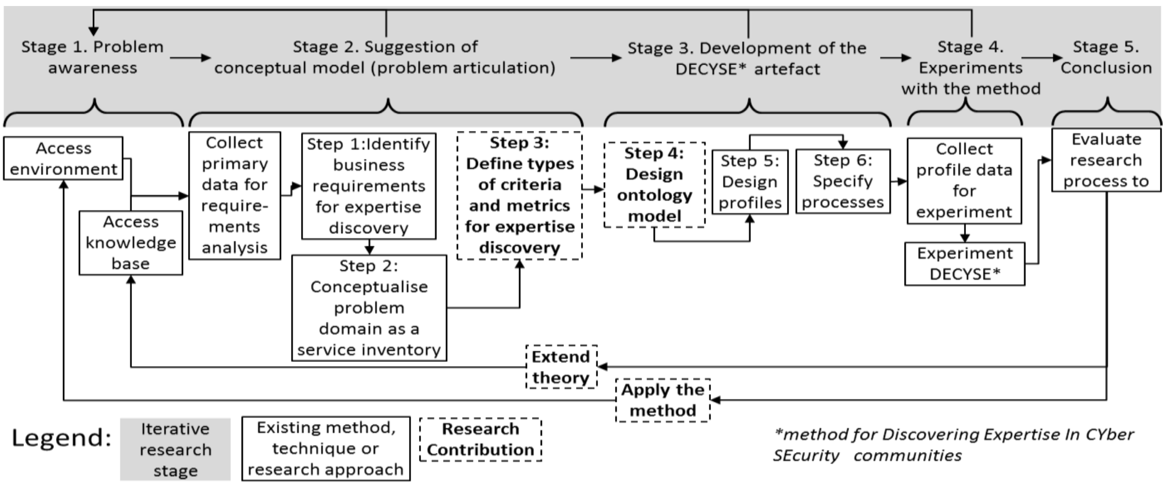


Figure 4.3: Solution design to devise research contribution

In compliance with Design Science Research, five major stages drive the whole research project. First, the nature of CS communities was investigated providing evidence that the problem of expertise discovery is regarded as having high relevance and has not yet been properly solved. Such investigation addressed the first research objective. Based on problem awareness and following the second and third research objectives, a thorough literature review on related theories supported by an empirical research was conducted. The outcomes showed that neither Talent Management nor Knowledge Management approaches alone were capable of solving the issue of CS expertise discovery. Afterwards, a conceptual model is suggested based on preliminary findings. Then, using elements from such conceptual model, the DECYSE method is developed aligned with the fourth and fifth research objectives. After a series of iterations with a test dataset, DECYSE is tested in order to validate the main contribution of this research project (cf. sixth research objective). Finally, DECYSE is evaluated according to its experimentation and conclusions lead to theory improvement and implications of the method for practice (cf. seventh research objective). The following sections describe in further detail how the solution design was performed.

4.2.2 Data Sources

The manifold challenges for sourcing data in the scale of a CS community encompassed capturing the requirements for articulating the research solution and the dataset for experimenting DECYSE. The questionnaires identified the participants who were either experienced individuals or those expecting to work with CS projects. In order to ensure that the sample was the most representative of the stakeholders involved, participants were also asked about their working area (e.g. private sector, academia and government). Data sources for developing and experimenting the method involved a Brazilian CS community mainly comprised by RENASIC, the Brazilian Army and public servants enrolled in a national information security course. During data collection, there were around 750 experts in RENASIC, 100 military working with CS in the Brazilian Army and 114 students enrolled in the national information security course. Such a CS community was chosen in compliance with the motivations driving this research (cf. last paragraph in Section 1.1) and to demonstrate that the solution can even be applied to societies in the early stages of deploying strategic CS projects. Communication with experts was conducted via emails, online calls, participation in conferences and online discussion forums. The number of people who were approached varied during the research stages, as presented in the following section.

RENASIC is the Information Security and Cryptography National Network supported by the Brazilian Government. RENASIC includes a community of innovation called COMSIC, which gathers experts from universities, research centres, government agencies and the private sector using a discussion forum (Brazilian Government, 2013b). The Brazilian Army is the organisation in charge of coordinating the CS effort in Brazil. Additionally, the Brazilian Government (2008a) intends to promote recruitment, selection, development and retaining of personnel to support CS. Time is of the essence for using the Brazilian CS setting. Currently, the Brazilian Government (2014) is in the early stages of projects involving the creation of a National School for Cyber Defence to coordinate CS skills development and accreditation of CS products and services.

4.2.3 Data Collection

Data collection for this research fulfils a twofold purpose. Firstly, primary data are used to triangulate findings in a literature review and develop the artefact, since some subjects are based on government documents and practices than in academic research. Secondly, they are used to empirically validate the method.

In this work, data is collected using a mixed method approach by way of observation, questionnaires (which include open-ended questions) and semi-structured interviews. Design Science Research supports the iterative creation of the method in parallel to the development of the aforementioned projects from the Brazilian Government (2014). Compliant with the solution design (cf. Figure 4.3), the process of data collection during research phases is described in Table 4.2.

Table 4.2: Research stages, activities, respective data sources and research techniques

Research stage	Activity	Data sources and research methods	Research technique
Problem awareness (preparation)	Identify the literature and practice gap in expertise discovery for cyber security	Consulting practice Theories and methodologies Official documents (secondary data)	Document analysis
Suggestion of conceptual model (problem articulation)	To identify key stakeholders and requirements for developing a conceptual model	Empirical research with primary data	Participatory observation and questionnaire (Appendix G)
Development of the artefact	To elicit additional requirements, design and test the method	Consulting practice and empirical research (both primary and secondary data)	Questionnaire (Appendix H), interview (Appendix I) and

Research stage	Activity	Data sources and research methods	Research technique
			document analysis
Experiments with the method	To experiment the method and to check validity according to elicited requirements	Acquire primary data (combination of real profile data with test data) to perform experimentation	Questionnaire (Appendix J) and literature review
Conclusion (solution evaluation)	To evaluate the method according to acceptance by experts and compliance with research objectives	Consulting practice (presentation and discussions)	Questionnaire (Appendix K)

A pilot investigation was conducted using observation and a first online questionnaire (Appendix G). The purpose of the pilot investigation is to provide a better understanding of CS, ensure relevance of the research problem, triangulate findings from the literature and elicit preliminary requirements for developing a conceptual view of the solution. Notes based on documents and definitions are structured on a mind map (i.e. a brainstorming technique) in order to provide an overview of the CS scenario. In fact, mind maps were widely used during the literature review in order to structure and balance the relevant topics and identify theoretical gaps. The first questionnaire was devised to provide both quantitative and qualitative results and is partly based on (Hiltrop, 1999). It (Appendix G) was conducted among 238 people involved with CS in different sectors of Brazilian society and with different knowledge background. Participants were mainly those working in the Brazilian Army or connected to RENASIC.

The empirical research used a second questionnaire and two interviews. The second questionnaire (Appendix H) was devised to acquire additional data and understand some findings on the first questionnaire. In addition to the former participants, federal government agency employees enrolled in a national information security course were addressed resulting in 66 valid answers. Moreover, the research aimed to determine the criteria to perform expertise discovery. A couple of semi-structured interviews (guided by questions in Appendix I) were devised to appraise the criteria for expertise discovery under the lens of CS experts. Both interviewees (I1 and I2) are retired senior military officers actively working in Brazilian CS projects (e.g. the Brazilian National School for Cyber Defence). Questions aimed to triangulate findings from the pilot investigation, improving accuracy on selected criteria and on the conceptual model. Interviews were followed by a brief presentation providing a holistic view of the artefact in order to capture feedback and improve the method.

4.2.4 Articulating the Research Problem

More than simply providing a solution for the research problem, knowledge management systems require maintenance through a particular set of processes in order to ensure its viability over time (Staab and Studer, 2004). This section provides an overview of the iterative development of DECYSE and how selected theories and subjects in information analytics are applied to such a method to enable its ongoing improvement.

The current CS community context was depicted during the problem articulation stage (cf. second row in Table 4.2). According to a questionnaire and findings in the literature, policy makers publish regulations, critical knowledge advisors publish frameworks with critical skills and CS strategies should provide incentive for research bodies. Currently, the discovery and selection of talents is conducted with contractors relying on different ways to hire workers. Most of the information about experts is ill-structured or not accurately measured (e.g. personal networks, social media and recommendations). Social media (in its many forms) is used to exchange knowledge and provide recommendations about peers and services. However, such interactions are not completely fit for the CS community, nor made accessible for analysis for many reasons (e.g. copyright, profit and privacy protection). Evaluations on service providers in the CS community (when they occur) normally are not shared, which leaves a gap while tracking their performance. Performance evaluation in CS projects was not found in the literature or in practice. Some form of evaluation among peers was detected in general social media platforms, however they do not seem to be reliable enough or do not provide suitable metrics to support expertise discovery. The current CS knowledge environment thus appears to be highly unstructured.

The literature review along with pilot questionnaires guided the selection of the stakeholders, roles and viewpoints (cf. Figure 2.7) to develop a conceptual model for DECYSE. Some of these viewpoints are identified as top drivers for both the Talent Management and CS fields. Subjects discussed in previous chapters (as shown in the white boxes in Figure 4.4) provide the means to address such motivations (depicted in grey background in Figure 4.4). These and other drivers inspire the creation of goals for expertise discovery, which in turn are eventually realised by services.

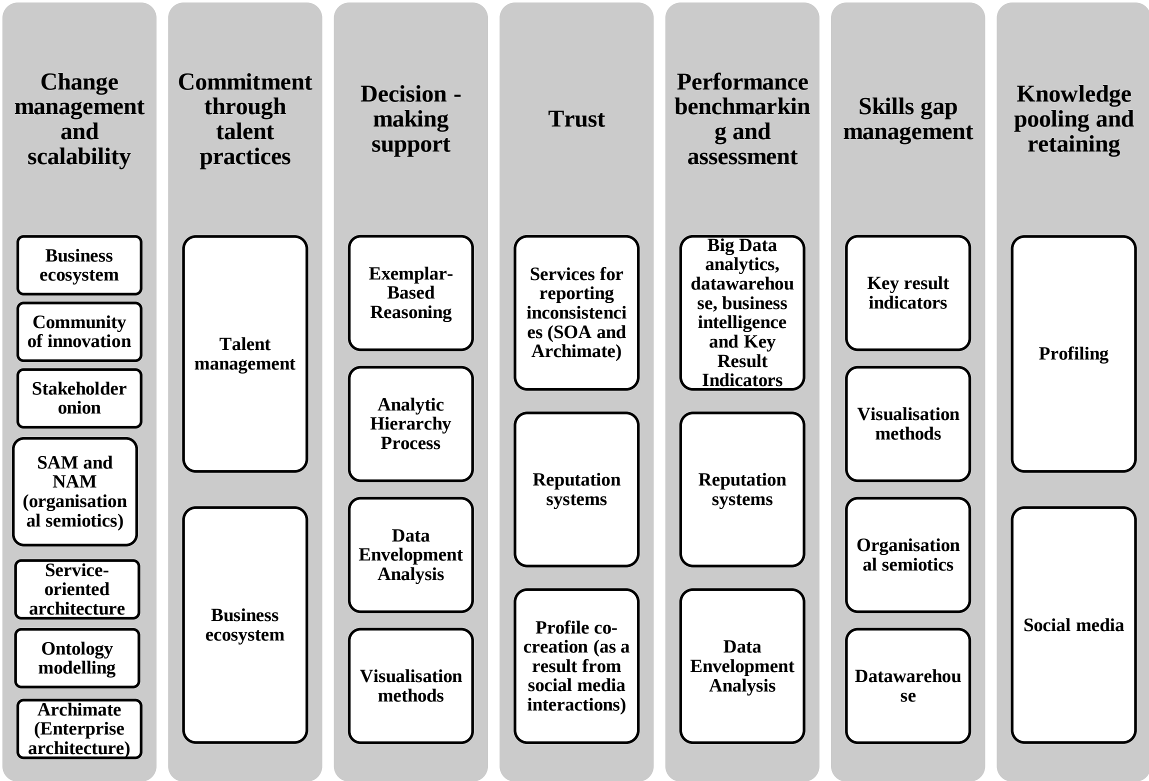


Figure 4.4: Top drivers and selected approaches for expertise discovery

DECYSE complies with the specification of expertise discovery principles (top layer in Figure 4.5) for the CS community, which are the normative properties to perform CS expertise discovery. These principles are derived from selected Talent Management practices and business ecosystem challenges to ensure that the research problem is articulated accordingly. In order to promote alignment with these expertise discovery principles, the requirements for the CS ecosystem are determined through a pairwise analysis between stakeholders, under each principle. The analysis determines the information services that each stakeholder requires and can provide to other actors, under the lens of these principles. Services that are relevant for discovering capabilities should be associated to expertise discovery principles and drivers elicited by stakeholders. These services are realised through expertise discovery functions, which comprise business concepts (e.g. roles and processes) in the CS community (middle layer in Figure 4.5). Selected business elements can be semantically integrated and transformed into a conceptual view using an ontology model to provide a solution for the problem (bottom layer in Figure 4.5). On the one hand, business elements facilitate requirements engineering and visualisation of motivational concepts by using Archimate. This ontology model combined with profiles and norms can assure consistency, completeness and rigorousness of representation and deliver automation support. On the other hand, the iterative nature of method development (illustrated with curved arrows in Figure 4.5) is a set of six steps (cf. Figure 4.3) that ensures compliance between the model and the domain context. The first three steps support articulation of the research problem, while the remaining three steps relate to the design of the DECYSE method. Understanding such steps as meta-processes allows those players with CS coordination roles to revisit requirements in order to improve the DECYSE method over time. The adopted layered view in Figure 4.5 intends to facilitate the understanding of DECYSE's development, since Archimate and semiotic-based ontology are simply different KR approaches, each one with their own strengths and weaknesses.

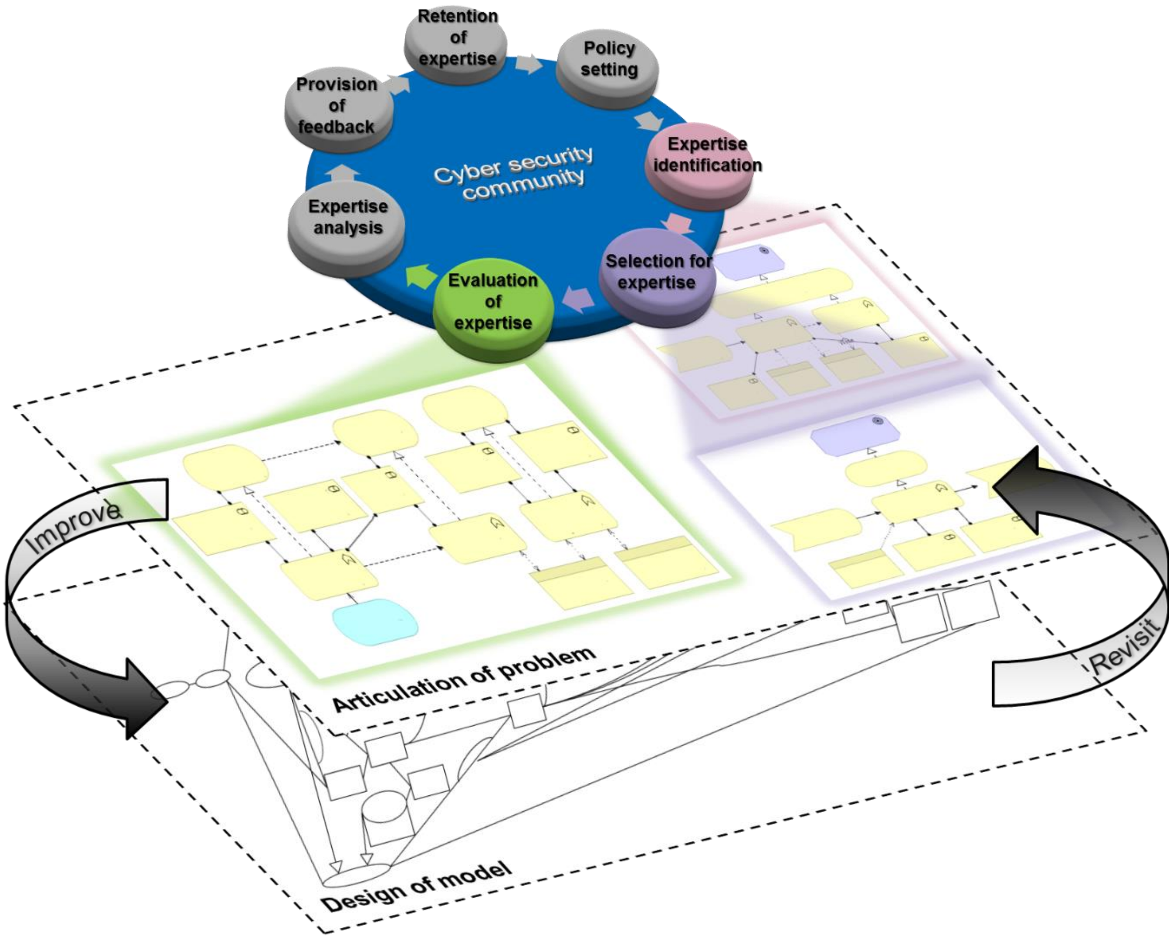


Figure 4.5: Expertise discovery lifecycle guiding DECYSE's iterative development

The expertise discovery lifecycle (top layer in Figure 4.5) obtains its structure from the talent lifecycle (cf. Schiemann, 2014) and seven principles from different Talent Management approaches (cf. Section 2.2.4), namely the following: policy setting, expertise identification, selection for expertise, evaluation of expertise, expertise analysis, provision of feedback and retention of expertise. The main contribution is that the expertise discovery lifecycle implies that services are applicable under the perspective of each actor within the CS community, and not simply focusing on the talent as a subject. Thus, the challenge in our approach relies on identifying the services that each actor can provide or benefit from a CS ecosystem under those selected principles in order to enable integration, joint stewardship, innovation and support co-evolution. While expertise discovery principles support identification of services for the CS community, drivers become the motivational aspects, which indicate the relevance of such services. These drivers were determined during the problem articulation stage (cf. Figure 4.3 and Figure 4.5), which was developed according to the following three steps:

- step 1: identifying business requirements for CS expertise discovery. These requirements include stakeholders, roles, drivers and contextualised content requirements for the problem domain. This step relates to the first stage in the Semantic Analysis Method and broadly with requirements elicitation (cf. Figure 2.9). In line with the business ecosystems approach, this work suggests that requirements should be drawn in a pairwise fashion between stakeholders. This means that each stakeholder should define their information needs and inform the information they can provide to benefit other stakeholders. The stakeholder onion is selected among other tools to analyse stakeholders mainly because it relates to the Organisational Semiotics approach adopted in this research. Selected Talent Management practices (presented on the top level of Figure 4.5) are the guiding principles used to discover the expertise in the CS context;
- step 2: conceptualising of the problem domain in the business layer as a service inventory described through functions and processes. The service-oriented approach offers flexibility through loose coupling of stateless services, which are offered according to roles in CS. The use of Archimate as a modelling language favours clear understanding when discussing with stakeholders during requirements negotiation (cf. Figure 2.9). It also enables the representation of motivational concepts and is aligned with the service-oriented approach;
- step 3: determining the types of criteria and metrics supporting CS expertise discovery. The purpose of this step is to determine the key concepts that should be profiled for information processing. These key concepts include both quantitative and qualitative data. For example, reputation systems and key result indicators are adopted in DECYSE to measure the experience and the expertise gap, respectively. Further data collection and analysis perform iterative refinement of DECYSE.

The purpose of the first two steps is to identify the concepts in the CS community business layer and serve as a reference for future improvement of DECYSE. After identifying major requirements and services based on two pilot questionnaires and the literature review, a conceptual view of the method is devised. The third step encompasses elements to answer the second research question. The aforementioned steps articulating the CS community business layer are presented in Section 5.1.

4.2.5 Design of DECYSE Method

This stage aims to document a subset of the CS business services and its embedded concepts into a single ontology model. The DECYSE method, which is one of the main contributions of this work, was designed after the problem articulation stage and comprises the following steps (cf. Figure 4.3):

- step 4: selecting and modelling the identified business elements through semantic analysis and documenting the process through an ontology schema (i.e. the DECYSE method). Organisational Semiotics is used to represent the dynamics of expertise discovery information space within the CS environment. Semiotic modelling enables the ontology evolution; therefore, user profiles and domain knowledge can be updated when necessary;
- step 5: structuring information requirements in profiles. Profiling techniques were already successfully combined with semiotic-based ontology in previous studies. In this thesis, profiles are used to capture both explicit (e.g. as a result from peer evaluations) and implicit data (e.g. metrics from expertise analysis) for processing. Social media provides numerous examples of relevant criteria for creating talent profiles and is a valuable data source for profiling individuals. In our approach, social media principles (e.g. co-creation) combined with the criteria for expertise discovery (determined in step 3) are used as the basis upon which interactions between actors build their reputation and where profiles are able to articulate with services provided by the CS community. Hence, the solution does not focus on developing a social media platform;
- step 6: specifying processes through norms. The capability of representing axioms is enabled by the richness of semiotic-based ontology. Thus, norm analysis is employed as a means to describe those analytical methods and techniques underpinning the affordances in DECYSE. Further implementation of these processes should require the use of databases (e.g. for profiling information) and programming. In order to support process specification for DECYSE, a CS

maturity model was devised (Figure 4.6). Such model complies with the iterative nature of the expertise discovery principles (Figure 4.5), because information flows (represented as arrows in the model) promote continuous improvement. However, it is worth noting that an initial set of CS capabilities should be defined before implementing DECYSE (represented with dashed lines in Figure 4.6). In addition, the CS maturity model presents how the selected analytical methods and techniques in literature (which relate to ICT) are layered for the specification of processes for DECYSE.

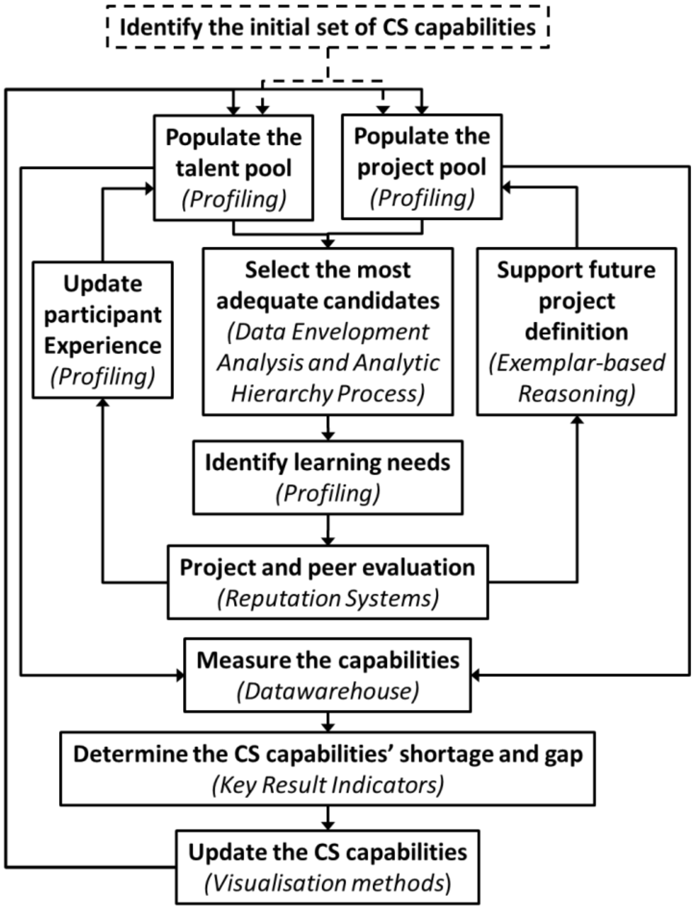


Figure 4.6: The CS maturity model for expertise discovery with its ICT layering

Profiling techniques are adopted to map the expertise in participants and projects by using an initial set of CS capabilities. Aligned with the adopted interpretive stance, decision-making techniques with ontology seem the proper choice. Despite lacking full automation (cf. Table 3.10), it considers a degree of subjectivity from the decision-maker, although requiring additional consistency checks. Numerical criteria extracted from profiles provide the required inputs and outputs to use with Data Envelopment Analysis, which is used to significantly reduce options in a talent pool to a select group of the fittest candidates. The adoption of the Analytic Hierarchy Process relies on its simplicity, flexibility to combine with other methods and due to its extensive use in the literature. In our approach, the Analytic Hierarchy Process enables the prioritisation of candidates based on qualitative criteria defined on CS project profiles. Both techniques support the selection for expertise principle (cf. Figure 4.5). Decisions can be supported by metrics derived from repeated performance evaluations from peers and former contractors. Thus, reputation algorithms are used to measure experts in terms of trustworthiness and experience, which eventually affect new expertise discovery processes. Profiling information contributes to the provision of feedback principle, whether identifying one's learning needs based on project requirements or updating a participant's experience as a result of interactions with peers and contractors. Exemplar-based reasoning is adopted to assist the creation of project profiles based on positive experience with similar projects. Finally, the principles of Datawarehouse, result indicators, visualisation methods and analytics are embarked on DECYSE in order to enable the use of business intelligence and global awareness on CS expertise shortage and gap.

Figure 4.7 summarises the aforementioned steps and their underpinning theoretical principles, based on state-of-the-art literature, as the theoretical framework of DECYSE.

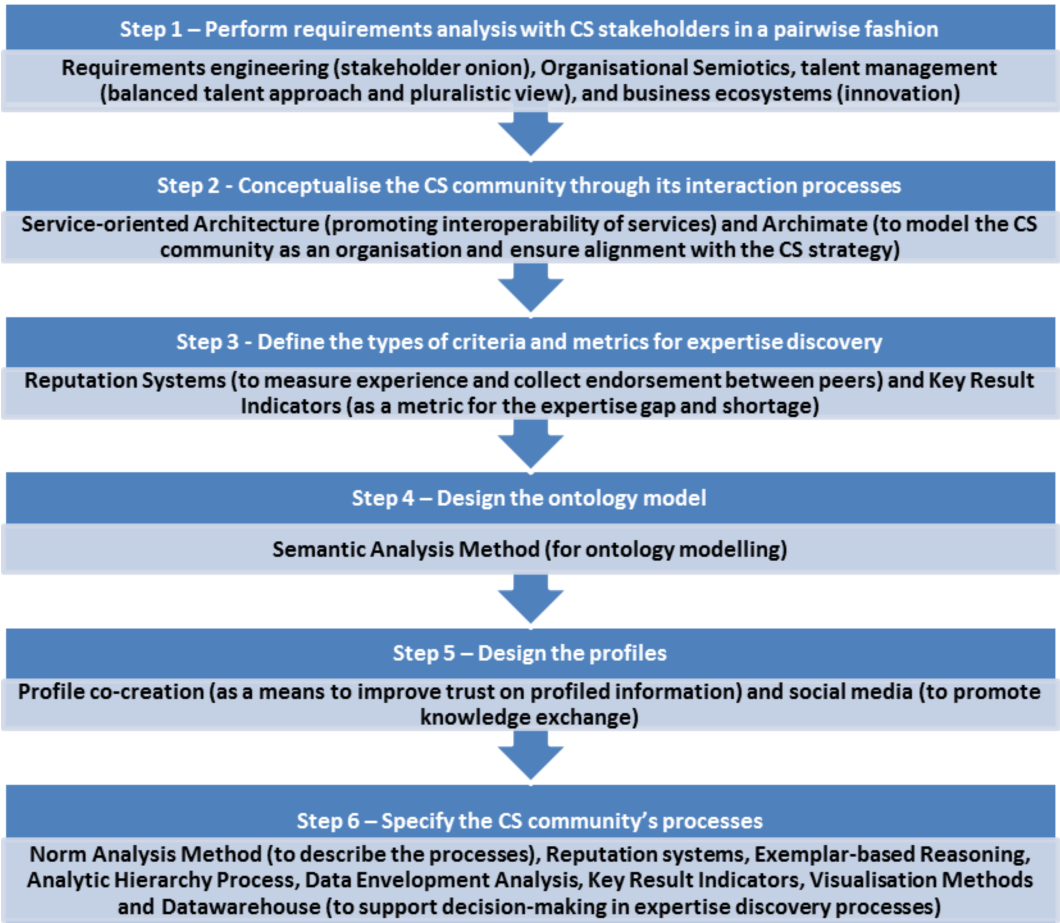


Figure 4.7: The theoretical framework of DECYSE

During iterative refinements, the DECYSE method was experimented with a test dataset, which was devised using real life values and enabled to exercise the method in different conditions. The DECYSE method with its profiles and analytical processes is described in Section 5.2.

4.2.6 Validation and Evaluation

There are numerous ways to validate research, although validation is affected by the adopted research methods. For example, Design Science Research has a particular set of validation patterns to evaluate an artefact (Vaishnavi and Jr, 2007). This research adopts mixed methods guided by Design Science Research, as well as construct, external, internal, content and benchmarking validity. Construct validity is the appropriateness and consistency of measurements because of testing scores. External validity relates to how much the research solution can be generalised to be applied to other settings. Internal validity refers to claiming that the results derive from expected combinations or relationships between adequate variables and approaches adopted for the research solution. Content validity is encompassed by construct validity and can use experts to assess the extent to which a testing item properly addresses a content or a problem (Taylor, 2013). In line with such a type of validation, synoptic validity refers to checking if reasonable outputs are achieved according to a set of inputs into the artefact (Finlay, 1989). Benchmarking is a Design Science Research validation pattern suitable for experiments where there are no metrics available beforehand. In such a case, a scenario should be created to evaluate the artefact. The merit of the measure relies on its independent validity, on not being biased towards the artefact and on meeting the criteria specified in the benchmark (Vaishnavi and Jr, 2007).

After a series of iterations with the test dataset in order to refine the solution in compliance with the research problem, a third questionnaire (Appendix J) was devised to profile a talent pool in order to experiment and validate the method. The questionnaire also identified expertise requirements used in past CS projects, which determined the remaining input to perform the experiment in more realistic conditions. Hence, the final experiments (presented in Chapter 6) used cleansed data acquired from real profiles. Generated data was used when it was not feasible to be collected from a real world scenario or when it resulted from interactions between participants (e.g. an individual’s reputation or the perceived degree of experience in a given competency). Such data was randomly created using a range of pre-defined values in order to support the validation of DECYSE. Results of such an experiment were compared to other test results (construct validity) or analysed in terms of its merit if no previous measure was available (benchmarking validity pattern). Validation also took into account DECYSE's generalisability (external validity) and accuracy according to the method's underpinning techniques (internal validity).

In order to ensure content validity, the method and results of its experimentation were presented to and evaluated by a board of experts (cf. Table 4.2) working in a Brazilian agency involved in the CS coordination effort and a potential user of the model. DECYSE was appraised in terms of applicability, usability and acceptance of the method (Appendix K). A detailed evaluation is conducted in Chapter 7, which includes the Design Science Research patterns under which the research project was conducted and discussions about the experiments. The evaluation also includes the contributions of DECYSE to both theory and practice and its limitations.

4.3 Summary

This chapter presented a thorough review in traditional and IS research paradigms according to their philosophical groundings. Moreover, research designs, methods and techniques were discussed along with other approaches, such as the radical subjectivist paradigm. This paradigm guides the development of methods able to work on dynamic environments where knowledge grows according to interactions among their actors. Information analytics approaches were carefully selected in order to ensure alignment between expertise discovery requirements and CS community motivations over time.

The understanding of the aforementioned concepts helped the author to select the most appropriate approach to address the research problem. The strong social context combined with a plurality of interacting stakeholders and subjective requirements pointed to the adoption of radical subjectivist paradigm. Mixed methods comprising Design Science Research, experiment, observation, questionnaire and interview were adopted as the most appropriate methods and data collection techniques in this scenario. The artefact is designed, developed and validated under the light of the Design Science Research approach due to the complexity of the CS context and use of numerous information analytics methods and techniques. The DECYSE method was devised in six steps within the stages of problem articulation and method design. The DECYSE method itself includes seven expertise discovery principles, which encompass categories of processes described in the following chapter.

Chapter 5

DECYSE: A Method for Discovering Expertise in Cyber Security Communities

This chapter describes the method for discovering expertise in cyber security communities (DECYSE) which is the main contribution made by this thesis. Prior to presenting the method, the DECYSE requirements are articulated and reveal the complexity of the method, which is further described through its affordances. Such articulation is part of a conceptual model (i.e. the research framework for developing the DECYSE method), which is another research contribution. DECYSE is capable of representing, sourcing and processing the information, which promotes the expertise discovery lifecycle. The method systematically suggests those most fit individuals to participate in CS projects given a set of criteria and measures the overall expertise shortage and gap in dynamic profiles.

5.1 Articulation of the DECYSE requirements

Expertise discovery in a CS community involves a complex process and multiple stakeholders. In line with the steps to develop the DECYSE method and their underpinning theory and viewpoints (cf. Figure 4.7), this section introduces and details the framework for developing DECYSE method. This study has started with a thorough articulation of its requirements prior to the DECYSE modelling. Firstly, the CS community is articulated in terms of its business elements in order to define the current context. Such a discussion is conducted under an organisational perspective to enable modelling of motivational concepts, to identify business concepts in the CS community and to facilitate further ontological representation of the services related to expertise discovery. The types of criteria for expertise discovery are introduced to support the profiles embarked in the DECYSE ontological model (i.e. the context “to be”). Secondly, the concepts and the categories of processes comprised by the DECYSE ontological model are introduced.

5.1.1 Identifying the Business Elements under the CS Community

The framework for developing the DECYSE method (Figure 5.1) guides the meta-processes to apply the DECYSE method in a given context and to improve it over time. It structures the viewpoints for CS expertise discovery (Figure 2.7 and Figure 2.13), as well as the steps, theory and analytical techniques presented in the theoretical framework (Figure 4.7). It is worth noting that the viewpoints were drawn from literature and triangulated with some primary data analysis. In the framework, the dashed arrow represents a new iteration to improve the DECYSE method by revisiting the articulation of the research problem (Figure 4.5). The current and following sections detail the framework for developing DECYSE.

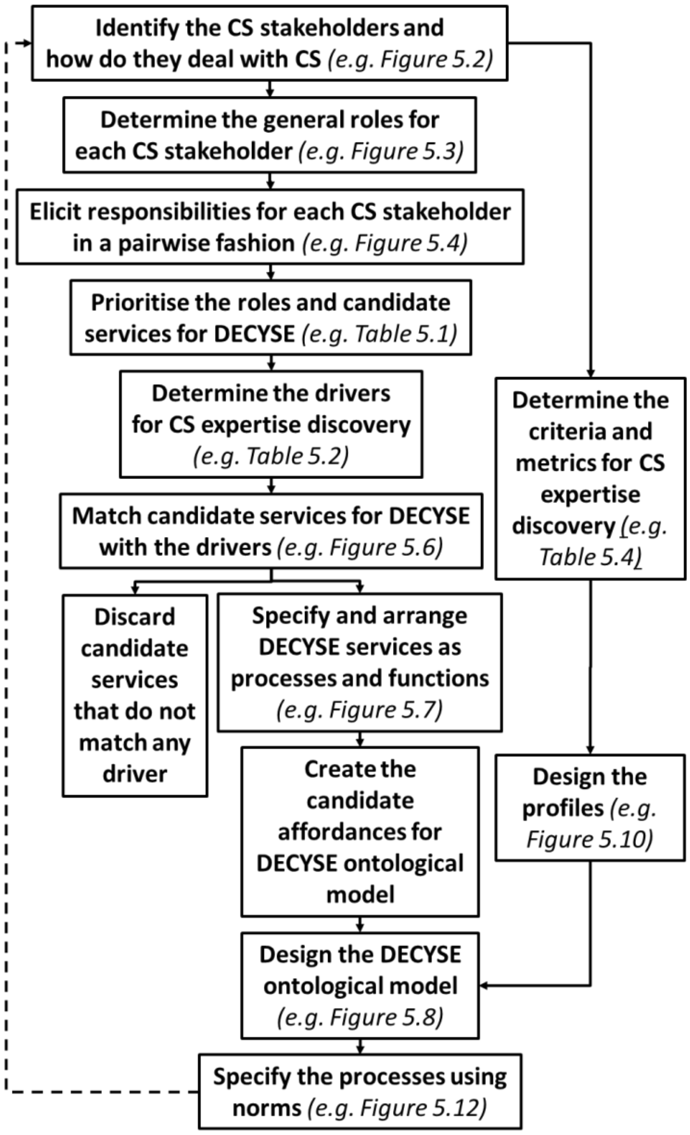


Figure 5.1: The research framework for developing the DECYSE method

The CS scenario can be understood according to the questions posed in Figure 5.2. Such mind map summarises the CS concepts in Appendix B and observations collected in an early stage of research work (cf. Section 4.2.4). The DECYSE architecture concentrates on "who are the stakeholders involved" and on "how to deal with CS". As one may notice, partnerships with clear roles and managing people, knowledge and trust constitute challenging requirements to increasing CS. The literature already provides a broad set of examples of regulations (e.g. policies, strategies, standards) and procedures (e.g. actions, operations, technologies). However, trust, people, knowledge and joint effort are required for CS (cf. Figure 5.2). Those requirements served as a starting point to determine the business layer concepts in the CS community, since they have not been systematically addressed. CS has a general set of stakeholders in charge of coordination, cooperation and collaboration activities (cf. Section 2.1.3). Since the context of this work relies on the concept of nation, those stakeholders related to the international community (i.e. performing a collaboration role) are beyond this research scope.

In our approach, the CS context comprehends a set of national stakeholders providing services to each other as a means to enable co-evolution. Such services are realised by workflows, which are influenced by a set of goals and drivers that guide semantic conceptualisation of national CS. The systematic integration of those services under the light of expertise discovery principles (cf. Figure 4.5) is assumed to improve innovation in the CS community by addressing business ecosystems' challenges.

The CS community was articulated in line with the research problem (cf. Section 4.2.4), while the following sections provide an overview on how to determine the business concepts (i.e. introduced in Figure 3.3) describing the "to be" context. Such concepts were identified when performing data collection with a CS community, in order to articulate the research problem. The key purpose of the following sections include presenting selected expertise discovery services for the CS community under an ecosystemic perspective and providing evidence to support and improve DECYSE in future iterations.

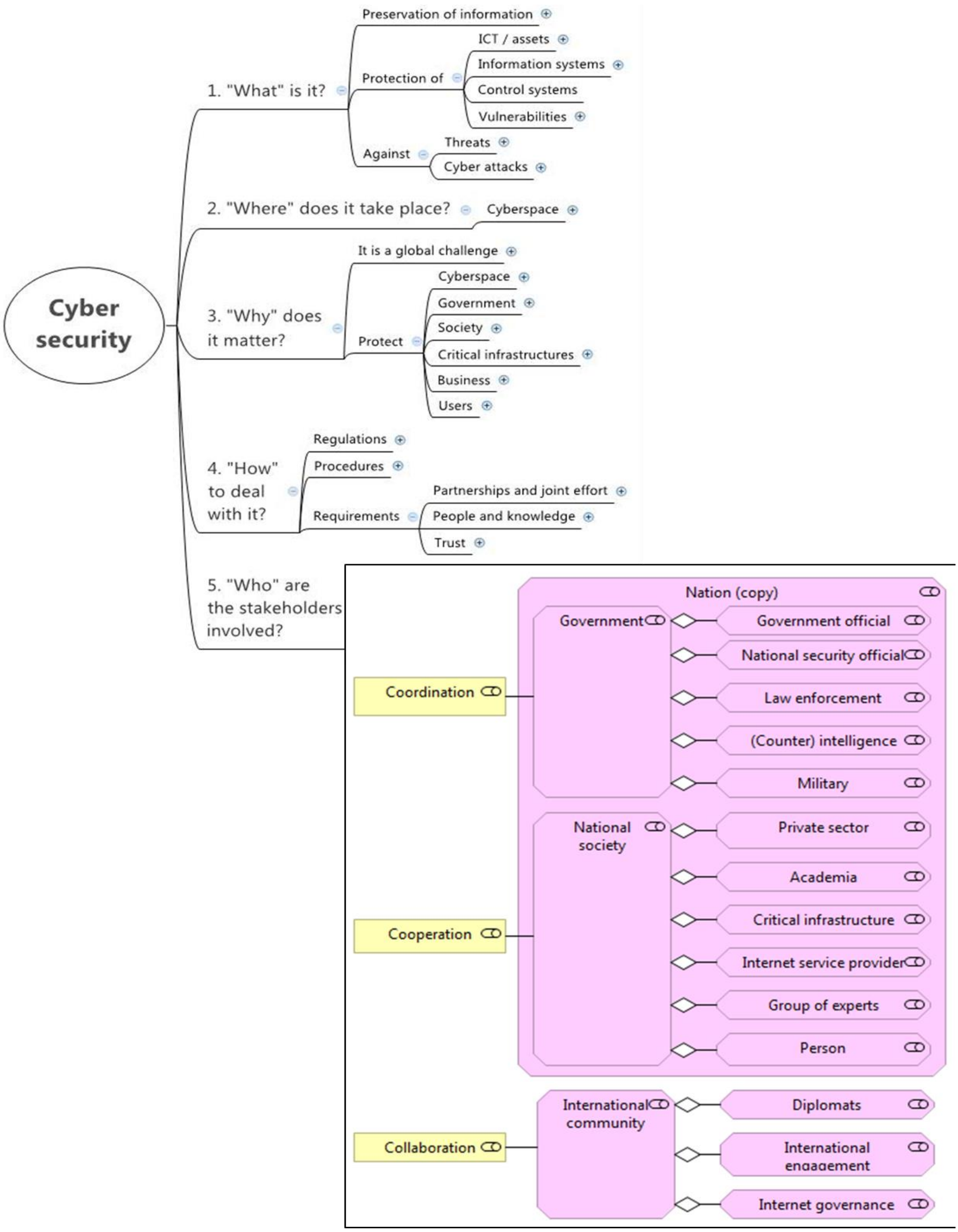


Figure 5.2: The main questions defining the cyber security context

5.1.2 Identification of Business Requirements for Expertise Discovery

A first step for articulating the research problem (cf. Step 1 in Figure 4.3) is to carry out the identification of business concepts (i.e. actors, roles, requirements, services and drivers) in the context of the CS community. The empirical research determined the requirement to integrate current CS expertise discovery principles and to fill in existing gaps, which would underpin the context "to be". Its results identified a set of responsibilities and services that should be offered to and provided by stakeholders in order to improve those principles in CS. Some of those services are further discussed. Stakeholders within the government and society can be classified using the stakeholder onion tool (Figure 5.3) and their responsibilities described according to the roles they perform within the CS community.

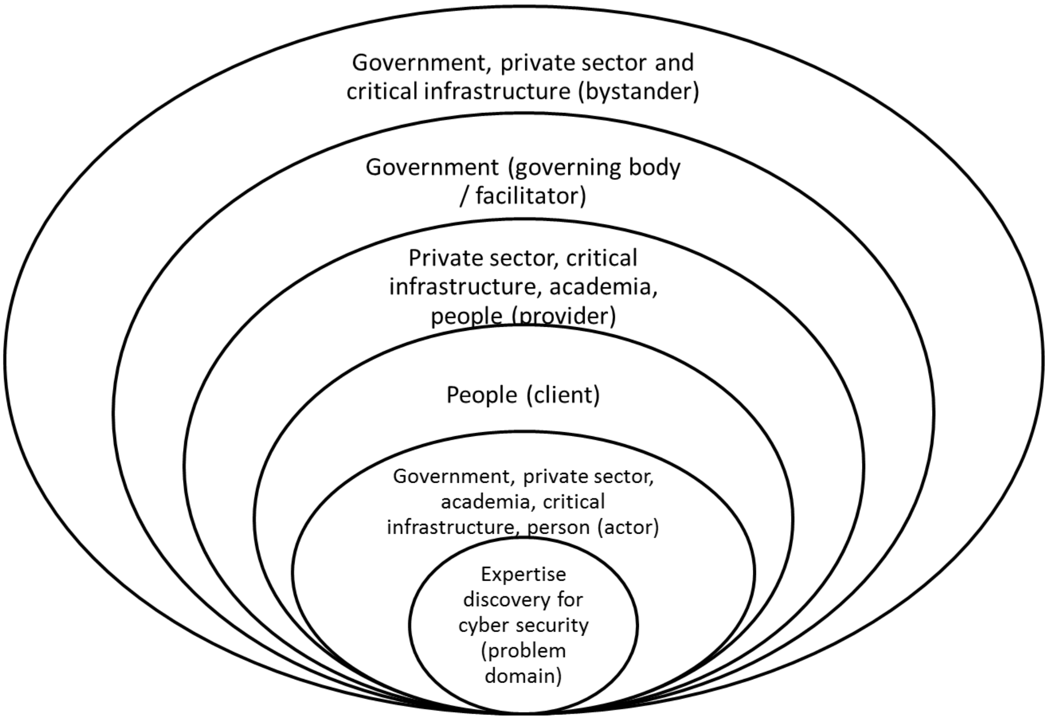


Figure 5.3: National cyber security stakeholders and their general roles

In this research work, each stakeholder may perform one or more general roles (due to the scale of our scope). The responsibilities that each stakeholder has, define DECYSE roles. At least one specific DECYSE role should be established for each stakeholder. Likewise, DECYSE roles involve at least one responsibility, which in turn becomes a service for the CS community. In order to support co-evolution and facilitate service elicitation, each stakeholder's perspective should be analysed according to the services he / she can provide to and benefit from the CS ecosystem under the light of each expertise discovery principle. Such an approach intends to promote innovation by establishing links between stakeholders and encouraging to share the stewardship of the community. Figure 5.4 illustrates an example of how to brainstorm stakeholder's responsibilities according to the adopted approach. The answers to the questions "what are my responsibilities to" and "what are my information needs from" each other stakeholder determines the business requirements, which specifies each expertise discovery principle.

Figure 5.4: Eliciting stakeholders' responsibilities for CS expertise discovery

The general roles provided by the stakeholder onion and respective responsibilities within the CS context served as a starting point to draw DECYSE roles and respective services, as shown in Table 5.1. The responsibilities (i.e. the requirements for expertise discovery) that stakeholders have elicited during the problem definition phase are contextualised through DECYSE roles and services. The general roles are eventually refined according to the relevance of DECYSE services, meaning that services associated to actor roles fall under the scope of the model, while other services are prioritised (compliant with the stakeholder onion hierarchy) for future iterations. The example in Table 5.1 aids the identification and prioritisation of the DECYSE roles and services, with fourteen services within the scope.

Table 5.1: Roles, responsibilities and services for CS community

General role	Stakeholder	Responsibility	DECYSE role	DECYSE candidate service
Actor	People	Keeps own profile up to date	Participant	"participant profile instance managing"
Actor	People	Evaluates peers according to their actions	Peer reviewer	"evaluating peer"
Actor	Government official, private sector and critical infrastructure	Evaluates contracted experts according to their performance in projects	Contractor	"evaluating work"
Actor	Government official	Suggests a list of suitable candidates to join a project	Data steward	"selecting expert"
Actor	Government official	Ranks a list of candidates suitable for a project	Data steward	"optimising selected expert"
Actor	Government official, private sector and critical infrastructure	Identifies required capabilities for candidates to join projects or key positions	Contractor	"generating criteria for expertise discovery"
Actor	Private sector and Academia	Offers training and certifications	Course provider	"offering course or certification"

General role	Stakeholder	Responsibility	DECYSE role	DECYSE candidate service
Actor	Government official	Offers training recommendations	Account administrator	“offering training recommendation”
Actor	National security official	Coordinates initiatives among stakeholders, devices and publishes CS strategy	Cyber security coordinator	“conducting CS strategy”
Actor	People	Applies for an account and a participant profile	Profile applicant	“creating participant profile instance”
Actor	Government official	Maintains profile schema	Account administrator	“maintaining profile schema”
Actor	Government official	Sets policies for profiles, privacy and capabilities according to CS strategy	Policy maker	“deploying CS strategy”
Actor	Government official	Archives long unused profiles	Quality controller	“archiving unused profile”
Actor	Government official	Performs data analytics on capabilities within talent and project pools	Quality controller	“performing analytics on capabilities”
Client	People	Assesses their own training needs and applies for course	Course applicant	“applying for course”
Client	People	Evaluates concluded courses	Course graduate	“evaluating course”
Provider	Private sector and Academia	Enables remote access to course results	Course provider	“validating certificate of conclusion”
Provider	Critical infrastructure	Identifies critical skills and flags critical projects	Critical knowledge advisor	“flagging critical skills and projects”
Provider	People	Exchanges knowledge with peers	Knowledge provider	“exchanging knowledge with peers”
Provider	People, private sector and critical infrastructure	Reports problems or advises improvements	Participant	“reporting a claim”
Facilitator	Government official	Accredits institutions providing courses and certifications	Quality controller	“accrediting course or certification”
Facilitator	Government official	Determines performance by accredited institutions	Quality controller	“accessing course performance”
Governing body	Government official	Disseminates regulations to nationals	Cyber security coordinator	“publishing regulation”
Bystander	Academia and Government official	Increases research and development in CS	Research body	“increasing research and development in CS”
Bystander	Government official, private sector and critical infrastructure	Provides lifelong training for employees by addressing their knowledge gap	Contractor	“providing lifelong training for participants”

It is worth noting that different stakeholders may play the same candidate role (e.g. contractor) and each stakeholder (e.g. government official) may play more than one candidate role, as shown in the example in Figure 5.5. However, candidate roles and responsibilities should preferably be normalised in a one to one relationship in order to provide a clear understanding and to facilitate further ontological modelling.

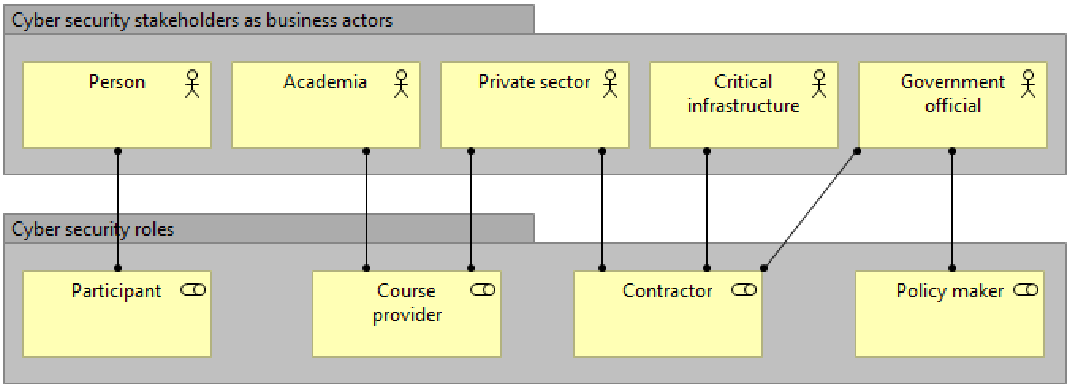


Figure 5.5: Conceptual view of stakeholders and roles in DECYSE

Determining the drivers for expertise discovery is crucial to ensuring alignment and relevance of services for the CS community. During the problem articulation stage, two hundred thirty-eight participants in the first questionnaire (Appendix G) were asked to value factors that contribute to attract and retain talents in CS using a Likert scale from 1 (not important) to 5 (extremely important). Considering value 5 as a valid result, Table 5.2 demonstrates the frequency of participants who deemed the respective factors as being extremely important. Results show that skills development, course opportunities and proper HR planning are the top drivers for attracting and retaining talents in CS. Such a finding is compliant with the ongoing initiatives drawn from the literature review and supports the relevance of this research work. In addition to the factors in Table 5.2, the empirical research identified other drivers, which were also used to develop a conceptual view of the service inventory for the CS community.

Table 5.2: Essential drivers for expertise discovery in cyber security

Frequency	Factor that contributes to attract and retain talents in cyber security
167	Skill development and courses
161	Efficient human resource planning
148	Recognition and rewards
145	Good payment
137	Career development and promotions
109	Equal opportunities
80	Transparency on goals, outcomes and intentions
75	Feedbacks
75	Leadership of the coordinating agency
71	Governmental agency coordinating cyber security effort
59	Autonomy and decentralisation of decision-making

Frequency	Factor that contributes to attract and retain talents in cyber security
56	Participation and teamwork
56	Possibilities of recruitment for specific tasks

The answers to open ended questions, which were related to additional drivers for CS expertise discovery, included five major topics. Firstly, there is a need for discussion forums to exchange knowledge and to remove barriers between stakeholders. Our method builds on such requirement to measure the collaboration between users (as further discussed). Secondly, incentives for both organisational and protean careers (with flexible work hours) are required. Hence, DECYSE deliberately considers the perspectives of participants and contractors. Thirdly, opportunities to work with state-of-the-art technology seem relevant to participants. In our approach, a description of the CS project, which might contain such details, is forwarded to its candidates. However, such a request is beyond the scope of DECYSE. Fourthly, special care should be taken to balance information privacy and visibility for opportunities. Thus, our model only discloses profiled information by which the expert wants to be discovered. In addition, DECYSE enables access to a range of up-to-date recommended capabilities for the CS community, which may create opportunities for self-improvement.

5.1.3 The Conceptualisation of the Problem Domain as a Business Service Inventory

The second step for articulating the research problem (cf. Step 2 in Figure 4.3) refers to conceptualising the processes and information flow in a service inventory, which comprises the DECYSE services listed in Table 5.1. CS services should be aligned with motivational concepts (e.g. drivers and goals) drawn from the literature, regulations and CS documents. Stakeholders should agree goals and drivers in order to improve synergy within the community. Services realise goals which are motivated by drivers. If a service does not have a goal within the context, it is not within the scope. If a goal is not realised by at least one service, the model is not efficient. Figure 5.6 illustrates the services "generating criteria for expertise" and "evaluating work" performed by a contractor, which is linked to/associated with the actor's government official, critical infrastructure and private sector (cf. Table 5.1). Those services realise and may share goals (e.g. "to support talent discovery"), which are influenced by drivers. Some drivers in Figure 5.6 (e.g. "good payment" and "work with state-of-the-art technologies") do not influence the identified goals, since they are not within the scope of the solution. However, they serve to guide the development of future services, which can be used to improve or extend the DECYSE method in further iterations. There are similar general roles (i.e. those other than “actors”) and respective responsibilities in Table 5.1 that are beyond the scope of this research.

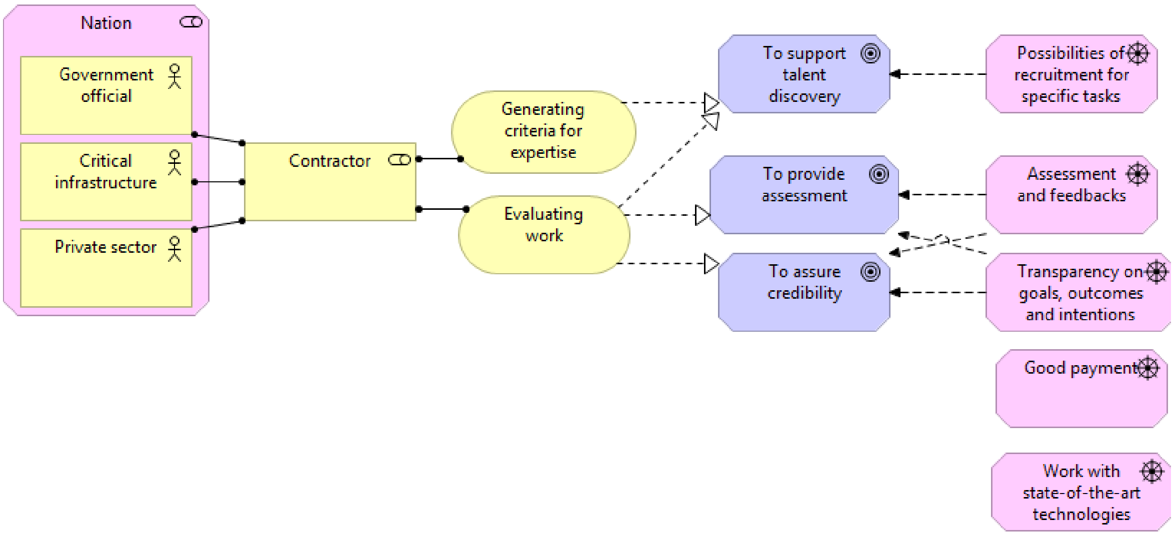


Figure 5.6: Excerpt of expertise discovery in national cyber security

Because some services may provide input information for other services, they should be arranged in such a way as to identify dependencies between all the identified services. Such an arrangement may imply selecting additional business roles and services when creating the ontology chart. Moreover, arranging the services by considering the information flow provides a holistic view of the service inventory for the CS community and facilitates designing the DECYSE method. The service inventory serves as a repository of business concepts that may even be used to conceptualise solutions for other problems affecting the CS community (i.e. if those services that are not associated to actor roles). The candidate services for DECYSE (cf. Table 5.1) are specified in terms of processes and functions in a “to be” context (Figure 5.7).

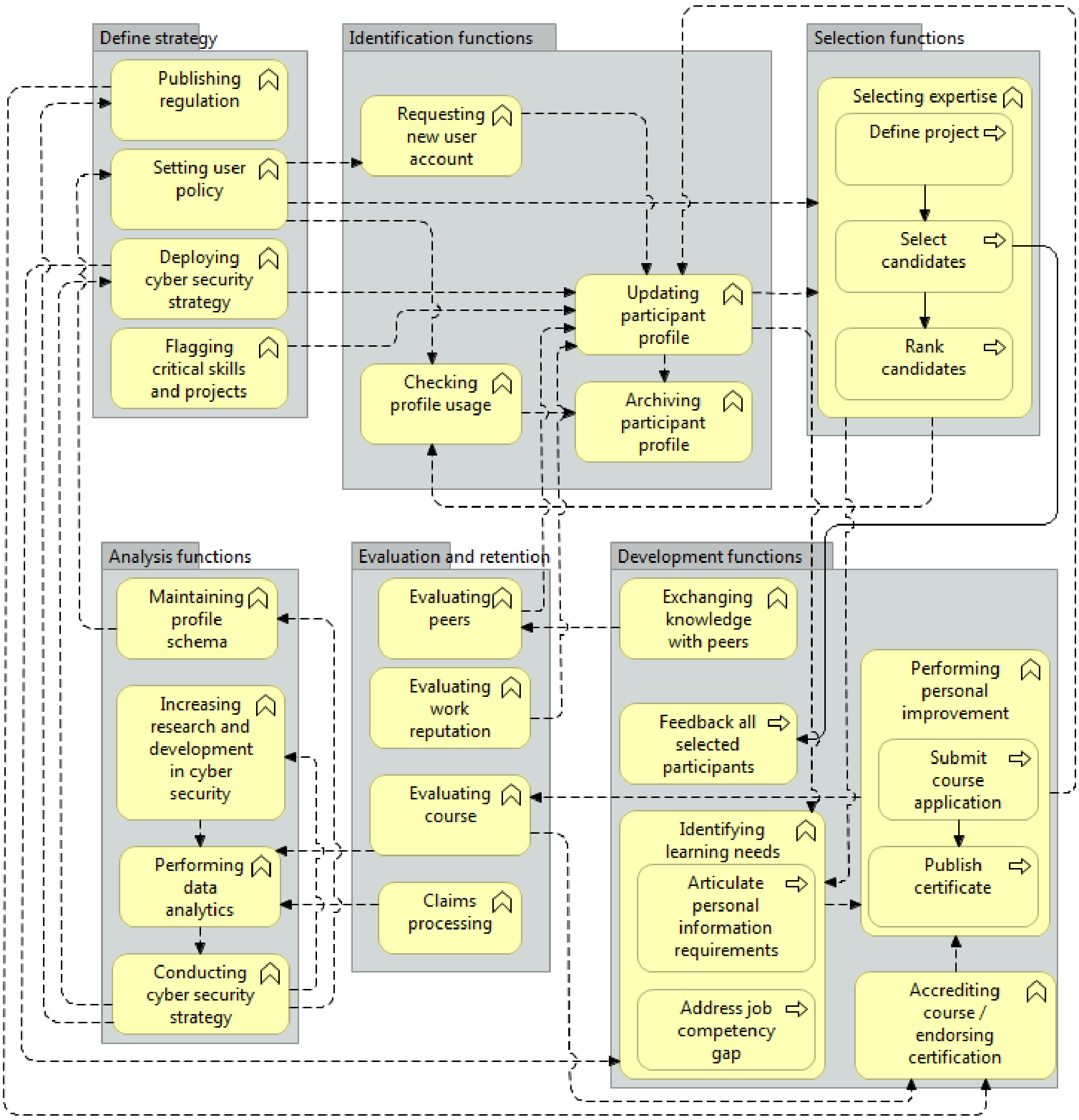


Figure 5.7: Conceptual and expanded view of the services for DECYSE

Those concepts can be clustered according to the expertise discovery principle that they seem more related. Grouping such concepts enables the solution architect and clients to visualise the information flow, how each principle is being addressed and eventually refine the system before performing the semantic analysis. In stark contrast to the “as is” context in the CS community (cf. Section 4.2.4), DECYSE offers integration of expertise discovery principles (i.e. improving the expertise discovery lifecycle). Each business process to be selected for ontological modelling is profiled in order to provide attributes for IT support. A template for profiling business processes can be seen in Table 5.3. Some profiled information (e.g. goal, CS service name and role) is already available in this step. Other information (e.g. outputs and norms) are provided when the process is ontologically modelled and analytically described through its norms.

Table 5.3: Business process profile template

Business Process Profile			
Process name:		Version:	Date:
Goal(s):		CS service name:	
Description:		Function(s):	
Role(s):	Input(s):	Output(s):	Norm(s):
Constraint(s):			

When selected services are fully described, it is possible to create the candidate affordances using stakeholders, roles (both drawn from Table 5.1) and processes (cf. Figure 5.7). Then, candidate grouping can be performed by structuring the selected semantic units, leading to the development of the DECYSE ontological model.

It is worth noting that only those services and roles related to stakeholders categorised as actors (cf. Table 5.1) are the ones to be ontologically modelled in order to comply with the problem scope. For illustration purposes, Appendix D presents conceptual views of business services, some of which guided ontological modelling and others that did not fall into the scope of this work.

5.1.4 Determining the Types of Criteria and Metrics for CS Expertise Discovery

The final step in articulating the research problem (cf. Step 3 in Figure 4.3) involves determining the types of criteria and metrics supporting CS expertise discovery, which addresses the second research question and supports the design of profiles. This step encompasses identifying the types of criteria to find suitable experts for projects (Table 5.4) and determining metrics to measure the "capability warehouse" over time. These types of criteria were determined via empirical research (cf. third row in Table 4.2).

Table 5.4: Nature and type of criterion for expertise discovery

Nature of criterion	Type of criterion	Observation
Qualitative	Competence Academic area Certification (either professional certifications or short courses)	Explicitly described in a project profile; also known as capability types; support definition of KRI

Nature of criterion	Type of criterion	Observation
Quantitative	Experience (competence level associated to each given competence) Academic degree (associated to each academic area) Number of requested certifications (existing in the project and participant profiles) Reputation (peer reputation) Waiting time	Implicitly determined from qualitative criteria or other profile features

DECYSE was developed in such a way that criteria can be combined in different types and quantities. Project specification actually depends on the three types of capabilities (i.e. the criteria with qualitative nature). For example, a project may request three competencies and one academic area and disregard certifications. The quantitative criteria (e.g. three competency levels, one academic degree, peer reputation and waiting time) are automatically determined according to the selection for expertise processes, as further described. The qualitative criteria, which are subject to ranking, have a deliberate correspondence in the quantitative criteria (e.g. competence with experience). Such a design decision ensures that no qualitative criterion is cast aside before the *optimise choice* process takes place, as detailed in the following section. For the sake of simplicity, the quantitative criteria do not need to be explicitly defined in a project, since the algorithm automatically selects the optimum values.

With regard to measuring the expertise, a metric was created for each type of capability (which is also the qualitative criteria for expertise discovery). Such a metric is the basis of some KRIs, which calculate the general expertise gap or shortage. These indicators represent meaningful patterns in profiled data (by using analytics) and serve to raise awareness on expertise in the CS community. Hence, DECYSE measures the expertise either for the benefit of specific actors (e.g. contracting organisations) or for the CS community as a whole (via the KRI).

5.2 The Description and Functions of DECYSE Ontological Model

5.2.1 The articulated DECYSE ontological model

The articulated requirements are used to design the DECYSE methodology (cf. Steps 4 to 6 in Figure 4.3) as shown in Figure 5.8. This model adapted the ontological approach, which is defined by a number of semantic units performing the expertise discovery in CS.

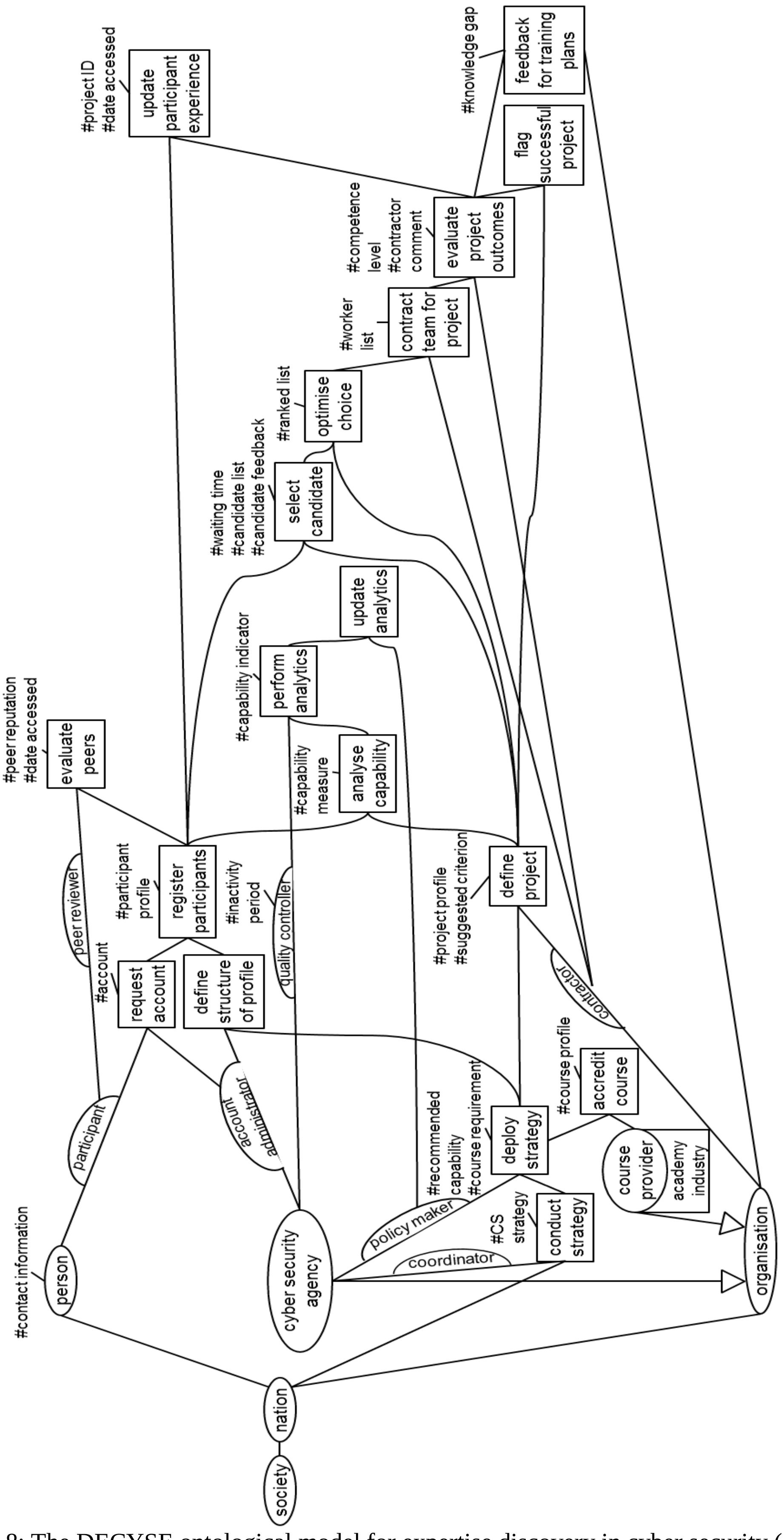


Figure 5.8: The DECYSE ontological model for expertise discovery in cyber security (adapted from Fontenele and Sun, 2016)

In this model, the concept of society is afforded by nation, which encompasses a CS context. Within such a context, there are two further actors: organisation and person. An organisation can be specialised as an agency providing either academic or industry accredited courses. Based on the nature of CS initiatives, an organisation can also be specialised as a CS agency. Each of these actors has some roles to play with regard to expertise discovery. The model transforms the seven CS expertise discovery principles (cf. Figure 5.4) into five groups of semantic units as follows:

- Define project. This is an affordance which ontologically depends on the semantic units of *conduct strategy*, *deploy strategy* and *accredit course* on its left. The affordance of *conduct strategy* is contributed by the antecedents of *cyber security agency* and *nation*. The CS agency acting as a *coordinator* elaborates and publicises the national *#CS strategy* via *conduct strategy*. The CS strategy contains the main drivers, goals and other concepts that describe the CS community. Aligned with such a strategy, the *policy maker* sets *#course requirements* for accreditation purposes and defines *#recommended capabilities* (e.g. courses and competencies) via *deploy strategy*. These requirements and capabilities are mapped so that a course provider can ensure that a course is suited to the CS community's needs for expertise, via *accredit course*. Those accredited courses are documented in a *#course profile*. An organisation has a role of contractor who is responsible for issuing and managing projects via *define project*. Such a process defines the CS project requirements that will be used to derive criteria for expertise discovery, which are described in *#project profile*. Priorities for criteria are assorted according to Appendix F. The contractor may specify his / her own set of criteria for the project and additionally rely on *#recommended capabilities* or *#suggested criterion* derived from similar projects.
- Register participants. This affordance is contributed by the semantic units of *request account*, *define structure of profile* and *evaluate peers*. The *request account* enables a person to apply for a participant account through the antecedent of an account administrator. The *define structure of profile* indicates the type of information to be collected from a participant, with the antecedents of account administrator and *deploy strategy*. The latter provides a list of recommended capabilities for the registering process. The affordance of *register participants* forms a semantic unit, which consists of the antecedents of *request account* and

define structure of profile. The *register participants* carries out activities to create a talent pool, which contains the *participant profile*s. Such an affordance, along with peer reviewers, are encompassed by the semantic unit of *evaluate peers*. In order to ensure trust and encourage knowledge exchange within the pool, the peer reviewer contributes by rating the interactions with other participants for reputation purposes; whereas a *participant profile* is dynamically co-created through appraisals by *evaluate peers*.

- Select candidate. The main affordances on the previous groups contribute to *select candidate*, which antecedes the affordances *optimise choice*, *contract team for project* and *evaluate project outcomes*. The *select candidate* workflow adopts a multi-criteria data analysis, i.e. Data Envelopment Analysis, to discover suitable candidates for the defined CS project. If more candidates are discovered than the CS project requires, the *optimise choice* adopts the decision-making method, i.e. the Analytic Hierarchy Process, to rank the most suitable candidates who are likely to be considered for participating in the CS project. The affordance of a *contract team for project* determines formally a team of talents and captures personal features required for the CS project. The performance of a contracted team is appraised after the completion of a project via *evaluate project outcomes* (Fontenele and Sun, 2016).

- Feedback for training plans. When appraisal on the contracted team is performed, the workflows of *update participant experience*, *flag successful project* and *feedback for training plans* are triggered. The affordance of the *update participant experience* is the remaining workflow that contributes to the co-creation of the *participant profile.* The affordance *flag successful project* identifies completed projects that can be further reused to improve future project definition. The affordance *feedback for training plans* provides feedback based on results delivered by the *evaluate project outcomes*, after completion of a CS project that supports identification of organisational knowledge gaps (Fontenele and Sun, 2016).

- Perform analytics. This affordance is contributed by *analyse capability* and delivers its outputs via *update analytics*. The semantic unit of *analyse capability* scales both talent and project pools to efficiency by removing obsolete profiles. Moreover, *analyse capability* computes the demanded (i.e. capabilities in project profile) and available expertise (i.e. capabilities in participant profile) to determine CS

#capability measure in a specific moment in time. Then, a quality controller triggers KRIs to acquire knowledge from those metrics via *perform analytics*. Such workflow calculates those *#capability indicators* in order to enable business intelligence and raise awareness on CS expertise. The *update analytics* workflow forwards the resulting *#capability indicators* to the policy maker.

The following subsections describe the model via these five groups of semantic units and detail some of its most representative affordances.

5.2.2 The Representation of "Define Project"

The semantic unit of *define project* is derived from three other major semantic units of *conduct strategy*, *deploy strategy*, and *accredit course* (Figure 5.9). The purpose of these semantic units is to set policies for expertise development and structure requests for expertise. Such context agrees with the need to ensure strategy alignment and to include an efficient coordination role within business ecosystems. The three leftmost semantic units support the policy setting principle in the expertise discovery lifecycle, whereas *define project* support the selection for expertise principle.

A national CS strategy is published via *conduct strategy* by a CS agency playing a *coordinator* role. Such a strategy contains the elements from which an initial set of *#recommended capabilities* is instantiated and *#course requirement* for accrediting CS courses are set via the *deploy strategy*. By abiding with these requirements, industry certifications and academic courses can be properly advertised within the CS community and become a *recommended capability*.

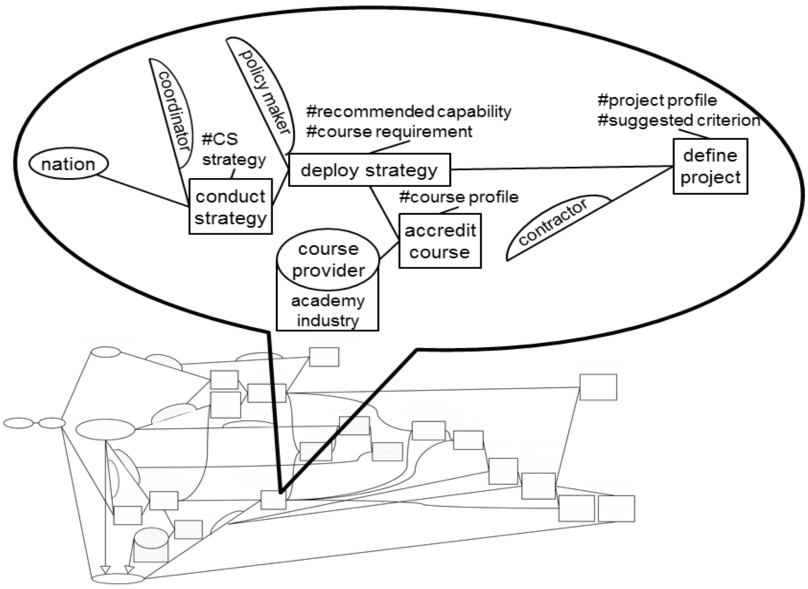


Figure 5.9: The affordance of *define project*

The data in *#recommended capability* represent competencies, academic areas and industry certificates (i.e. the qualitative criterion in Table 5.4). In our approach, those capabilities are suggested by a policy maker and structured according to Table 5.5.

Table 5.5: The structure of CS recommended capabilities

Responsible for defining	**Capability**	**Structured in terms of a set of**	**Defined in (affordance)**
Policy maker	Competence	Knowledge, skills, abilities (KSA) and tasks	*deploy strategy*
Course provider	Academic area (with respective degree) and certification	Competencies	*accredit course*

Certain competencies, academic areas or professional certifications may be considered as a recommended capability as a result of analytics feedback. In addition, a *course provider* may also offer and recommend training services for the CS community with the *accredit course* workflow. These courses (either certifications or academic courses) should develop a set of competencies, which in turn are defined as a set of KSA and tasks. The structure of those concepts is inspired by CS initiatives presented in Chapter 2 and considers the types of criteria in Table 5.4. However, instantiating KSA and tasks while structuring such concepts is not within the scope of this research work. Hence, the three types of capabilities are investigated as concepts with the highest granularity in this thesis.

The purpose of setting policies for the CS community is twofold. On the one hand, recommending capabilities provides conceptualisation of such properties for ontological commitment. For example, a *participant*, a *contractor* and a *course provider* should have a common understanding of what a competence means before claiming it, requesting it for a project or offering courses to develop it, respectively. On the other hand, the set of *recommended capabilities* is updated by the *course provider* (via *accredit course*) and by the *policy maker* based on analytics feedbacks (provided by *update analytics*).

Some discussion is required on finding the proper sources for developing competencies, since this is a major driver for discovering the CS expertise (cf. Table 5.2). The affordance of *accredit course* is contributed by the antecedents of *course provider* and *deploy strategy*. The workflow of *accredit course* articulates training services from a *Course provider* with recommended CS competencies defined in *deploy strategy*. If compliance with *#course requirement* is satisfied and the course offers development of at least one *#recommended capability*, a *Course profile* is created and made available to course applicants.

The affordance of *accredit course* has a property of *#Course profile* as shown in Figure

5.10. Its structure contains a set of elements, i.e. *Course ID, Course title, Description, Competence, Provider ID, Authentication* and *Course assessment*.

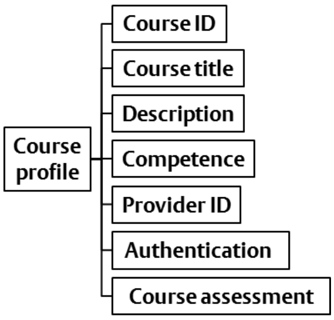


Figure 5.10: The course profile

The *Course ID* is the unique identification for each course. *Course title, Description* and *Provider ID* serve the purpose of providing basic information about the course. *Competence* includes the set of recommended capabilities that the course aims to develop. *Authentication* contains the URL pointing out the list of course graduates in order to ensure trustworthiness in the information provided by the *participant*. The *Course assessment* captures the accrued evaluation performed by the *course applicant* after the completion of the course calculated in a similar way to peer reputation and competence level. However, the articulation of the *course profile* with the DECYSE ontological model in terms of course application and evaluation lies beyond the scope of this work. The purpose is to present the required set of features for profiling accredited courses and enable expansion of the DECYSE method to include course evaluation in future work (see the "Evaluating course and course applicant" example in Appendix D).

The process of the *define project* guides the creation of a *project profile* as a preliminary step in selection for expertise. A *contractor* specifies a project *P* to address a particular problem in the CS domain. The project profile data can use *#recommended capability* and *#suggested criterion* to perform the expertise discovery. The latter utilises additional criteria drawn from similar and successful projects.

A project *P* can be specified in *Project profile* with the features of *Project ID, Title, Goal, Description, Keyword, Capability, Start date, Finish date, Contractor information* and *Completed project* (Figure 5.11).

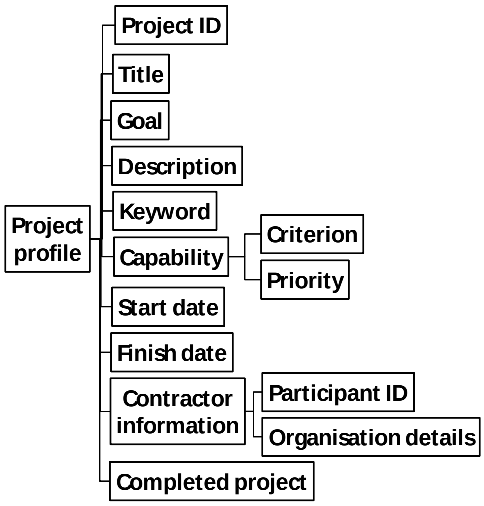


Figure 5.11: The project profile (adapted from Fontenele and Sun, 2016)

The *Project ID* and *Title* identify *P*. The project *Goal* and *Description* define a purpose and details of *P*, which guide an evaluation of participants. The *Keyword* enables searching for similar projects for feature reuse and derivation from the *Goal*. The property of *Capability* holds a number of qualitative criteria (cf. Table 5.4) for seeking suitable expertise for the *P*. Each *Criterion* can be assigned with a *Priority* according to the nature of the *P* and its required skill sets. The *Start date* and *Finish date* determine the duration of *P* (Fontenele and Sun, 2016). *Contractor information* ensures that the *contractor* is already registered via a *Participant ID* and indexes the *Organisation details* to the profile. The property of *Completed project* is flagged to indicate if the project has been successfully concluded.

A project *P* comprises a set of capabilities as a resource to solve similar problems, as an Exemplar-Based Reasoning application. The workflow in Figure 5.12 aids a *contractor* to refine the project requirements by suggesting additional criteria (and their priorities) present in similar projects that were successfully concluded. There are two major differences in our approach that contrast to examples of CBR in the literature review. Firstly, the project solution in our approach (i.e. the selected candidate) is not saved due to privacy reasons and to comply with the dynamic skill evolution of the talent pool. Therefore, this approach is oriented to reusing project requirements rather than the same individuals. Secondly, the choice for comparing the *P.Keyword* instead of *P.Capability* is to focus on retrieving projects with similar goals instead of projects with similar content. Jaccard coefficient (cf. Eq. (2)) is suitable for binary data that does not require numerical attributes; thus, it is used to determine similarity with other projects. When a *contractor* defines *P.Capability.Criterion*, the respective *P.Capability.Priority* is calculated via a pairwise comparison (cf. Appendix F). If the similarity rate among keywords from a P^{new}

and a $P^{retrieved}$ reaches a threshold (initially arbitrated as 80%), then exclusive criteria from $P^{retrieved}$ is suggested to the contractor, giving the opportunity to refine project requirements.

<Affordance define_project>
<StartNorm ID=N1.1>
<whenever>expertise is required for a *P*</whenever>
<if>*deploy strategy* exists AND *contractor* exists</if>
<then>*contractor*</then>
<is>permitted</is>
<to>structure expertise requirements in *P*</to>
</StartNorm>

<FinishNorm ID=N1.2>
<whenever>expertise is required for a *P*</whenever>
<if> *deploy strategy* ceased OR *contractor* terminated</if>
<then>*contractor*</then>
<is>obliged</is>
<to>discontinue definition of *P*</to>
</FinishNorm>

<OperationalNorm ID=N1.3>
<whenever>defining *P*</whenever>
<if>a set of *P.Capability.Criterion* is specified</if>
<then>*contractor*</then>
<is>obliged</is>
<to>perform a comparison analysis between *P.Capability* criteria according to Appendix F AND generate *P*.{*Title, Goal, Description, Keyword, Start Date, Finish Date*}</to>
</OperationalNorm>

<OperationalNorm ID=N1.4>
<whenever>N1.3 is satisfied</whenever>
<if>a set of *P.Capability*.{*Criterion*, *Priority*} is specified</if>
<then>*project creation engine*</then>
<is>obliged</is>
<to>find matching $P^{retrieved}$ by calculating *Simil* ($P^{new}, P^{retrieved}$) = $\frac{n(P^{new}.Keyword \cap P^{retrieved}.Keyword)}{n(P^{new}.Keyword \cup P^{retrieved}.Keyword)} \times 100\%$,
where $P^{retrieved}.Completed\ project$ is flagged AND $n(P^{new}.Keyword\ \cup\ P^{retrieved}.Keyword) > 0$
AND
present *suggested criterion* (*Criterion, Priority*) = {$P^{retrieved}$.*Capability*−P^{new}.*Capability*, *Simil* ($P^{new}, P^{retrieved}$)}, where *Simil* ($P^{new}, P^{retrieved}$)≥80%</to>
</OperationalNorm>

<OperationalNorm ID=N1.5>
<whenever>N1.4 is satisfied</whenever>
<if>*contractor* is satisfied AND weightings set for *P.Capability.Criterion* is determined</if>
<then>*project creation engine*</then>
<is>obliged</is>
<to>store resulting weights in *P.Capability.Priority* AND store *contractor#description* in *P.Contractor_information* AND create *P.Project_ID*</to>
</OperationalNorm>

<OperationalNorm ID=N1.6>
<whenever>N1.4 is satisfied</whenever>
<if>*contractor* is not satisfied AND weightings set for *P.Capability.Criterion* is determined</if>
<then>*contractor*</then>
<is>permitted</is>

```
<to>update criteria AND execute N1.3</to>
</OperationalNorm>
</Affordance define_project>
```

Figure 5.12: The norms of *define project*

5.2.3 The Affordance of "Register Participants"

The affordances of *define structure of profile*, *request account* and *evaluate peers* contribute to *register participants* (Figure 5.13), which outputs a *#participant profile*. The purpose of these semantic units is twofold. Firstly, a talent pool for the CS community using a single participant profile is created, which realises the expertise identification principle. Secondly, the affordance *evaluate peers* promotes contributions between peers, which is aligned with the need to promote innovation and for establishing contribution links within business ecosystems. There are two additional affordances (i.e. *update participant experience* and *analyse capability*), which contribute to the expertise identification principle and to the *register participants*. Profile updates may also be derived from the participant's performance and experience based on CS project results provided by *update participant experience*. Moreover, the *analyse capability* workflow archives a *participant profile* due to inactivity, leaving the talent pool only with active participants. However, the semantic units containing these affordances are discussed in the following sections. This section discusses the semantic units in Figure 5.13 and the properties used to profile participants for expertise discovery.

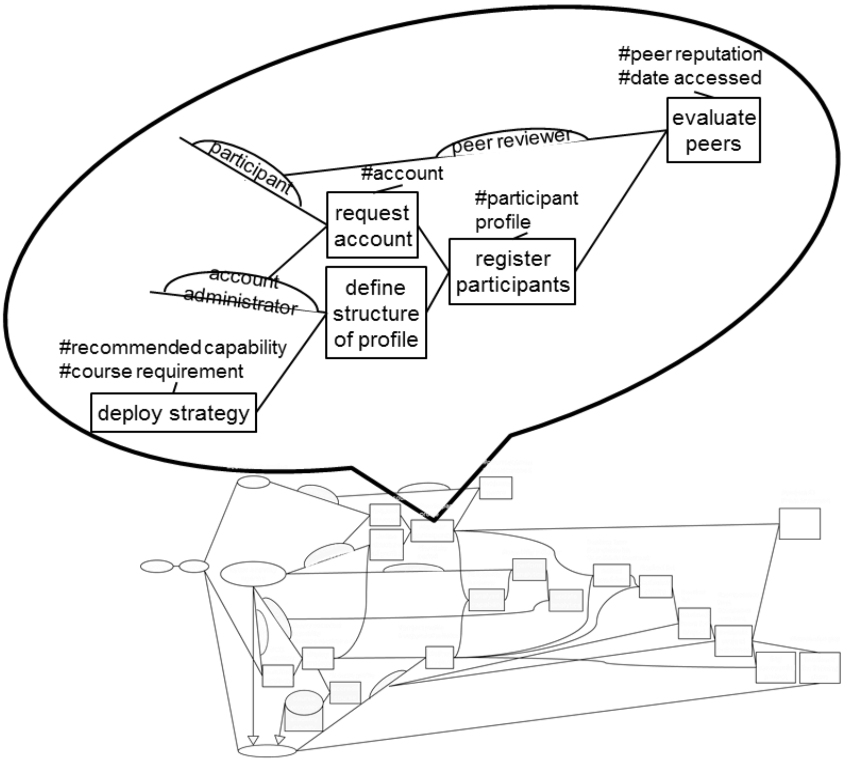


Figure 5.13: The affordance of *register participants*

In order to become available for expertise discovery, a person has to apply for a user *#account* through the *request account* workflow so as to become a *participant*. The purpose of such process is to check for authenticity (e.g. to define password and to restrict assignment of multiple accounts per individuals). The *account administrator* performs the *define structure of profile* workflow to enable *recommended capability* values to be registered in a *participant profile*. Given that the information matches the requirements set by an *account administrator* in *define structure of profile*, a *Participant profile* is created as an outcome of *register participants*.

The profile schema structures the information needs, which articulate with the DECYSE ontological model. The *participant profile* (Figure 5.14) contains the major properties of *Participant ID*, *Contact information*, *Capability*, *Availability date*, *Complementary info*, *Candidate feedback* and *Date accessed*.

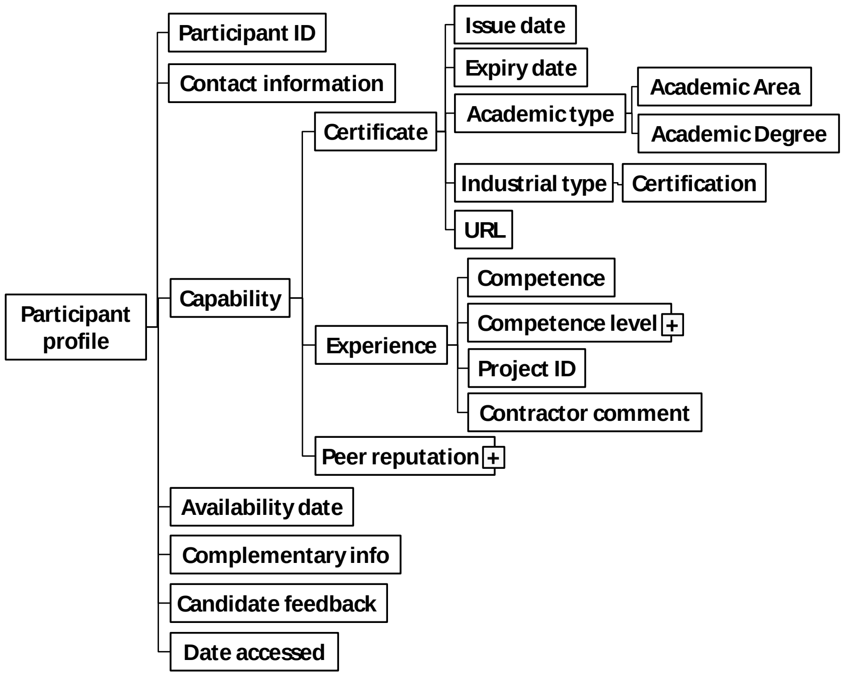


Figure 5.14: The participant profile (adapted from Fontenele and Sun, 2016)

A *Participant ID* is uniquely assigned to each *participant*. The *Contact information* is disclosed to a contractor after each *selection for expertise* process takes place. The *Availability date* indicates when the participant is available to join a CS project. In the participant profile, *Complementary info* and *Candidate feedback* serve a purpose of self-managing the expertise. The former contains additional information suitable for contractors, while the latter obtains details of the CS project for which the participant has been successfully selected. The profile's *Date accessed* refers to when the profile was created. Such a feature is updated during workflows under the evaluation of expertise principle and may serve to optimise the talent pool with *analyse capability* process (Fontenele and Sun, 2016).

Capability is expanded to define *Certificate* of achievements in types (academic or industrial), awarding institutions, dates when they were awarded and a URL (enabled by *accredit course*) to ensure authentication of information. *Capability* is also defined in terms of *Experience*, comprising a set of *Competencies*; a respective *Competence level* that captures an objective evaluation result after being involved in projects (i.e. *Project ID*); and *Contractor comment* containing additional subjective evaluation (for future contractors' appreciation). *Peer reputation* captures appraisals from *peer reviewers* during online interactions in a closed social media, so that a *participant* can continuously contribute within the CS community, even when not involved in projects. A *Participant profile* is thus co-created by a participant itself (e.g. *Certificate* and *Competence*) and external evaluations (e.g. *Competence level* and *Peer reputation*) that iteratively build the participant's professional online image. In order to measure a participant's performance and reputation, the participant profile captures ratings generated from interactions with contractors (e.g. *Competence level*) and peers (e.g. *Peer reputation*). A *competence level* is connected with each *competence* held in *participant profile* and embraces the work score (*WS*) and total work score (*TWS*) based on which the work reputation (*WR*) is determined. The level of each competence claimed is hence defined by practical experience. *Peer reputation* (PR) is similarly defined by peer score (*PS*) and total peer score (*TPS*). These values represent a reinforcement of trust on personal information in order to ensure higher quality experts being selected when discovering CS talents (Fontenele and Sun, 2016). Score variables (i.e. *WS*, *TWS*, *PS* and *TPS*) for each competence claimed start with value 1, which results in reputation variables (i.e. *PR* and *WR*) with an initial value of 0.2242.

The *participant* updates the qualitative details in the *participant profile* to be used during selection for expertise processes. Moreover, the *register participants* (Figure 5.15) sets initial quantitative values for *Date accessed*, peer reputation and for each competence identified by *participant*. This affordance ensures that each *participant* has the minimum set of data required for the talent pool.

<Affordance register_participants>
<StartNorm ID=N2.1>
<whenever>information regarding *participant* requires updating</whenever>
<if>*request account* exists AND *define structure of profile* exists</if>
<then>*participant*</then>
<is>permitted</is>
<to>create a *participant profile*</to>

</StartNorm>

<FinishNorm ID=N2.2>
<whenever>information regarding *participant* requires updating</whenever>
<if>*request account* ceased OR *define structure of profile* terminated</if>
<then>*participant*</then>
<is>obliged</is>
<to>discontinue definition of *participant profile*</to>
</FinishNorm>

<OperationalNorm ID=N2.3>
<whenever>*participant profile* is specified with requirements in *define structure of profile*</whenever>
<if>a *participant* account is created</if>
<then>*participant*</then>
<is>obliged</is>
<to>update *Contact information* AND *Capability.Certificate* AND *Capability.Experience.Competence* AND *Availability date*</to>
</OperationalNorm>

<OperationalNorm ID=N2.4>
<whenever>a *participant* account is created</whenever>
<if>*participant profile* is specified AND *Peer reputation* is nil</if>
<then>*participant profile engine*</then>
<is>obliged</is>
<to>create (*Participant ID*, *Date accessed*) AND store {1,1,0.2242} in *Capability.Peer reputation*</to>
</OperationalNorm>

<OperationalNorm ID=N2.5>
<whenever>a new *Capability.Experience.Competence* is added to *participant profile*</whenever>
<if>*Capability.Experience.Competence_level* is nil</if>
<then>*participant profile engine*</then>
<is>obliged</is>
<to>store {1,1,0.2242} in *Capability.Experience.Competence_level*</to>
</OperationalNorm>
</Affordance register_participants>

Figure 5.15: The norms of *register participants*

Before the selection for expertise workflows occurs, the *evaluate peers* is one of the processes in the expertise identification, evaluation and retention categories. This process considers outputs from *exchanging knowledge with peers* service, which is in the service inventory (cf. Table 5.1); however it is not within the scope of our DECYSE method. Such a service is devised to enable a closed social media platform (e.g. the discussion forum in COMSIC) dedicated to exchange CS knowledge with peers and with rating capabilities on those interactions. Figure 5.16 defines the algorithm of *evaluate peers* that reinforces each registered participant with regard to a level of trustworthiness and commitment. The operational norm has employed the Reputation System in Eq. (6), (7) and (8) for the processing tasks. In order to ensure trusted evaluations, the algorithm considers the *peer reviewer*'s own reputation to determine the evaluation impact (Fontenele and Sun, 2016). In our approach, $PS_{participant}^{t+1}$ should be greater than zero, since the algorithm is not able to determine a negative reputation and maintain the quality standards within the pool of experts. Moreover, *PR* is further used to determine efficient units using Data Envelopment Analysis, which requires positive values. Hence, N3.5 assures that both *PS* and *PR* have a minimum value of 0.001.

<Affordance evaluate_peers>
<StartNorm ID=N3.1>
<whenever>a *participant* is subject to evaluation by peers</whenever>
<if>a peer reviewer exists AND *Participant profile* exists</if>
<then>*peer reviewer* </then>
<is>permitted</is>
<to>rate reputation of participants</to_action>
</StartNorm>

<FinishNorm ID=N3.2>
<whenever>a *participant* is subject to evaluation by peers</whenever>
<if>peer reviewers ceased their roles OR *Participant profile* ceased</if>
<then>*peer reviewer*</then>
<is>obliged</is>
<to>discontinue rating tasks</to>
</FinishNorm>

<OperationalNorm ID=N3.3>
<whenever>a *participant* is subject to evaluation by peers</whenever>
<if>a positive or a negative rating is attributed to a *participant*</if>
<then>*peer reviewer*</then>
<is>obliged</is>
<to>assess a *participant* as follows:
calculate $TPS^{t+1}_{participant} = TPS^{t}_{participant} + PR^{t}_{peer\ reviewer}$
AND
$$PS^{t+1}_{participant} = \begin{cases} PS^{t}_{participant} + PR^{t}_{peer\ reviewer} \text{ (positive rating) OR} \\ PS^{t}_{participant} + (-1/100) \times TPS^{t+1}_{participant} \times PR^{t}_{peer\ reviewer} \text{ (negative rating)} \end{cases}$$
</to>
</OperationalNorm>

<OperationalNorm ID=N3.4>
<whenever>a *participant* is subject to evaluation by peers</whenever>
<if>$PS^{t+1}_{participant} > 0$</if>
<then>*reputation engine*</then>
<is>obliged</is>
<to>calculate $PR^{t+1}_{participant} = \alpha \times w + (1-\alpha) \times y$,
$\alpha = 1/(1 + exp^{-(\log TPS^{t+1}_{participant} - k)})$, where *k*=2
$w = PS^{t+1}_{participant}/TPS^{t+1}_{participant}$
$y = 1/\left(1 + exp^{-(\log PS^{t+1}_{participant} - k)}\right)$, where *k*=2)
AND
update $TPS^{t+1}_{participant}$, $PS^{t+1}_{participant}$, $PR^{t+1}_{participant}$ and #*date_accessed* in *Participant profile*</to>
</OperationalNorm>

<OperationalNorm ID=N3.5>
<whenever>a *participant* is subject to evaluation by peers</whenever>
<if>$PS^{t+1}_{participant} \leq 0$</if>
<then>*reputation engine*</then>
<is>obliged</is>
<to>update $TPS^{t+1}_{participant}$, $PS^{t+1}_{participant}$=0.001, $PR^{t+1}_{participant}$=0.001 and #*date_accessed* in *Participant profile*
</to>

```
</OperationalNorm>
</Affordance evaluate_peers>
```

Figure 5.16: The norms of *evaluate peers* (adapted from Fontenele and Sun, 2016)

The outcomes of $TPS_{participant}^{t+1}$, $PS_{participant}^{t+1}$ and $PR_{participant}^{t+1}$ update the existing value of *Participant profile#Peer reputation* along with a time stamp in *Participant profile#Date accessed*. For example, a participant with a *#peer reputation* defined by $\{TPS_{participant}^{t}, PS_{participant}^{t}, PR_{participant}^{t}\}$ provide knowledge contributions in a discussion forum and receives a rating from a peer with a given $PR_{peer\ reviewer}^{t}$. If the rating is positive, $PS_{participant}^{t}$ is slightly increased when state changes from *t* to *t+1*. Otherwise, $PS_{participant}^{t}$ is significantly decreased in *t+1* when receiving a negative rating (Fontenele and Sun, 2016). Those peer-evaluated participants build their reputation, which can improve their chances to be considered as a candidate for a project *P*.

5.2.4 The Workflow of "Select Candidate"

The affordances of *select candidate, optimise choice, contract team for project* and *evaluate project outcomes* are embarked on in the context of the *select candidate* affordance (Figure 5.17). The purpose of the semantic units discussed in this section is threefold. Firstly, the expertise selection takes the criteria defined in *#project profile* to search for candidates (i.e. *select candidate* and *optimise choice*) to work in CS projects (i.e. *contract team for project*). Secondly, provision of feedback raises awareness on suitable expertise for candidates (via *select candidate*). Thirdly, evaluation of expertise in projects (i.e. *evaluate project outcomes*) occurs post to the conclusion of the project to determine new competence levels for participants. The semantic units discussed in this section are portrayed in Figure 5.17.

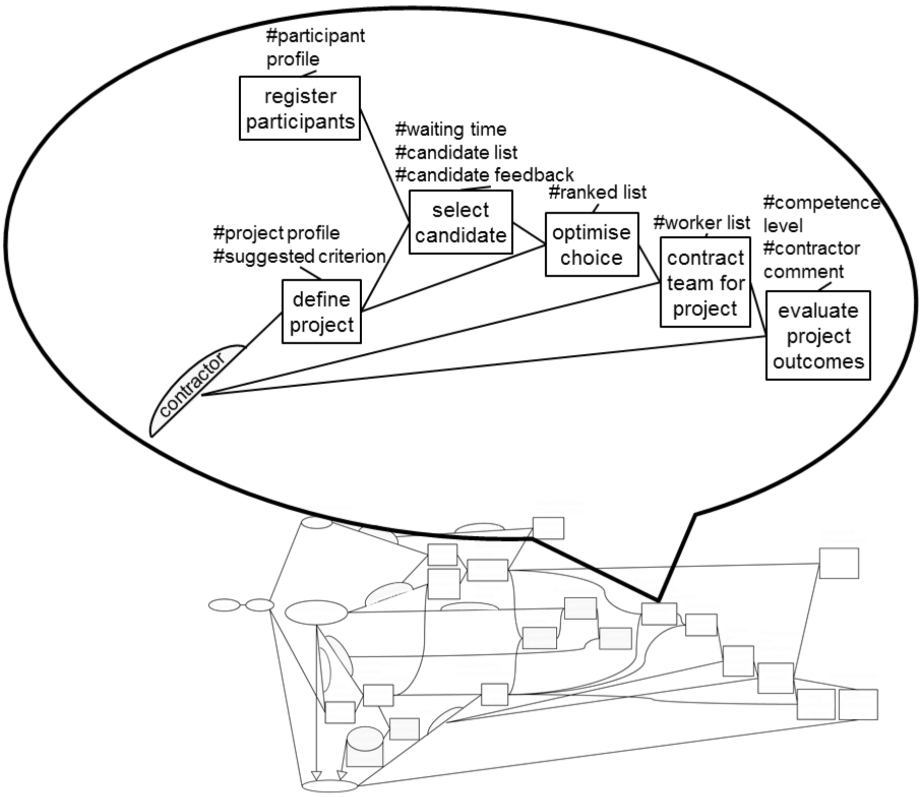


Figure 5.17: The affordance of *select candidate*

Selecting candidates for a project *P* requires the outputs of *define project* and *register participants*. Figure 5.18 determines how the selection is performed during the *select candidate* process applying Data Envelopment Analysis on quantitative criteria specified in Table 5.4. Some of these criteria are obtained from the participant profile (i.e. *competence level*, *academic degree* and *peer reputation*). Other quantitative criteria (i.e. *waiting time* and *number of requested certifications*) are calculated on demand.

Prior to running the selection through the participants in the profile, #*waiting time* is calculated as shown in N4.3. This condition ensures that only those candidates whose *availability date* maps with the project *start date* enter the selection. The two exceptions to this condition involve the cases where there are no available candidates with a project requirement or when a candidate who excels in the required capabilities is going to be available in a short period. Moreover, the algorithm establishes a minimum value of 1 for the *waiting time* to be further used as a Data Envelopment Analysis input. The number of requested certifications present in a *Participant profile* is also calculated (cf. N4.4) as an output to ensure selection of the most fit in terms of certification. A symbolical level value of 0.001 is assigned when the requested capabilities (i.e. y_r and x_i in N4.5) are not presented in a *Participant profile* (Fontenele and Sun, 2016). The peer reputation as a Data Envelopment Analysis output ensures that, amongst participants matching the same project criteria, those with a higher reputation are preferred as candidates. In line with the provision of feedback principle and to support complementary capabilities in the CS ecosystem, the *select candidate* process updates the *participant profile* with the requirements in the latest project in which the participant is a candidate (cf. N4.6).

<Affordance select_candidate>
<StartNorm ID=N4.1>
<whenever>a new project defines procurement for CS expertise as *P.Capability*</whenever>
<if>*Participant profile* exists AND *Project profile* exists</if>
<then>select engine</then>
<is>permitted</is>
<to>select suitable candidate experts</to>
</StartNorm>

<FinishNorm ID=N4.2>
<whenever>a new project defines procurement for CS expertise as *P.Capability*</whenever>
<if>*Participant profile* ceased OR *Project profile* ceased</if>
<then>select engine</then>
<is>obliged</is>
<to>terminate the selection process</to>
</FinishNorm>

<OperationalNorm ID=N4.3>
<whenever>pre-select candidates' availabilities meet *P.Start_date*</whenever>
<if>*P.Start_date* is specified AND *Availability_date* in *Participant profile* is provided</if>
<then>select engine</then>
<is>obliged</is>
<to>calculate $Waiting\_time_{Participant\ ID}$ as $P.Start\_date\ - 1 \geq Availability\_date_{Participant\ ID}$ AND update *#waiting time*</to>
</OperationalNorm>

<OperationalNorm ID=N4.4>
<whenever>pre-select candidates' availabilities meet *P.Start_date*</whenever>
<if>*P.Start_date* is specified AND *Certification* in *Participant profile* is provided</if>
<then>select engine</then>
<is>obliged</is>
<to>calculate n(*Cert*), s.t. *Cert*={$P.Capability.Criterion \cap Participant\ profile(Certification)$}, where *Cert* is the set of certifications in both participant and project profiles</to>
</OperationalNorm>

<OperationalNorm ID=N4.5>
<whenever>select candidates meet *P.Capability* AND N4.4 is satisfied</whenever>
<if>the multi-criteria (input and output variables) AND the weights set according to the *P.Capability*</if>
<then>select engine</then>
<is>obliged</is>
<to>calculate e_0 for each DMU, where $e_0 = \{max \sum_r u_r y_{r0} / \sum_i v_i x_{i0}$, s.t. $\sum_r u_r y_{rj} - \sum_i v_i x_{ij} \leq 0, \forall j \in \mathbb{N}^* \mid y_{r0}, x_{i0} \geq 0.001$ AND $u_r, v_i \geq \varepsilon, \forall r, i \in \mathbb{N}^*\}$
AND
create the *#candidate list* with the DMU ($e_0 = 1$)</to>
</OperationalNorm>

<OperationalNorm ID=N4.6>
<whenever>feedback to unsuccessful candidates needs to be provided</whenever>
<if>*#candidate list* is not nil</if>
<then>select engine</then>
<is>obliged</is>

<to>generate feedback comments according to the *P.Goal, P.Capability* AND update *#candidate feedback* AND update *#date_accessed* in *Participant profile*</to>
</OperationalNorm>
</Affordance select_candidate>

Figure 5.18: The norms of *select candidate* (adapted from Fontenele and Sun, 2016)

The *select candidate* may produce a number of qualified Decision-Making Units. In this case, a next process of *optimise choice* performs further analysis with the purpose of ranking these Decision-Making Units for best fit. Figure 5.19 describes a pairwise comparison based on the Analytic Hierarchy Process to determine weight between the alternative Decision-Making Units. The specified qualitative criteria (cf. Table 5.4) and respective weights, as derived from the project specification, are used in this analysis to produce a ranked list of those candidates. Alternative Decision-Making Units refers to a Decision-Making Unit having or not having a required value in each criteria. The associated weights on having a required value is deemed as "extremely important" as opposed to not having such value (i.e. respectively 0.9 and 0.1 by using Appendix F). The *#ranked list* assists in decision-making on determining the most suitable for the project (Fontenele and Sun, 2016).

<Affordance optimise_choice>
<StartNorm ID=N5.1>
<whenever>ranking candidate DMU with $e_0 = 1$ for a *P*</whenever>
<if>*#candidate_list* exists AND *P.Capability* exists</if>
<then>*optimise engine*</then>
<is>obliged</is>
<to>carry out an optimisation process</to>
</StartNorm>

<FinishNorm ID=N5.2>
<whenever>ranking candidate DMU with $e_0 = 1$ for a *P*</whenever>
<if>*#candidate_list* ceased OR *P.Capability* terminated</if>
<then>*optimise engine*</then>
<is>obliged</is>
<to>discontinue an optimisation process</to>
</FinishNorm>

<OperationalNorm ID=N5.3>
<whenever>ranking candidate DMU with $e_0 = 1$ for the *P*</whenever>
<if>the required qualitative criteria are defined in *P.Capability* AND their weightings set is determined according to Appendix F</if>
<then>*optimise engine*</then>
<is>obliged</is>
<to>perform a comparison analysis of the DMU in *#candidate list* according to respective alternative priority
AND
create *#ranked list* for the DMU</to>
</OperationalNorm>
</Affordance optimise_choice>

Figure 5.19: The norms of *optimise choice* (Fontenele and Sun, 2016)

Prior to starting a project *P*, a team of experts is assembled by the *contractor* based on the options provided on *#ranked list*. A *#worker list* is created containing *Competence levels* of team members with regard to *competencies* in *P*, during *contract team for project*.

After the end of the project, a contractor performs an assessment via the *evaluate project outcomes*. Feedback can be generated for both individual participants involved in the project and the organisers who defined the project. Figure 5.20 defines the assessments and measures (e.g. high, medium or low) with regard to the participants' performance. *WR* has also employed Eq. (6), (7) and (8) for the analysis, but in a different way from *PR*. First, while a participant has a single *PR* because of multiple evaluations from peers, *WR* is associated with each competence for the evaluated participant. Thus, its purpose is to measure experience gained over time in certain competencies while working on projects. Second, *WR* deliberately does not consider the evaluator's (i.e. contractor) reputation, as this work is talent and project oriented rather than focused on employers. Hence, each contractor has the same weight (Fontenele and Sun, 2016).

It is worth noting that each competence is associated to a respective set of tasks (cf. Table 5.5) for the purpose of providing an objective rating against a standard (Kanij et al., 2014). If a worker (whether employee or contracted on demand) delivers all the requested tasks associated with a given competence, he/she receives a medium rating. If a worker excels the tasks for a given competence (or performs tasks above his competence level), the worker receives a high rating for that particular competence. However, if a worker is not able to perform tasks or demonstrate KSA associated with a particular competence, he/she may receive a low evaluation. Medium and high evaluations increase $WS^{t}_{competence}$ at different rates, while low evaluations decrease $WS^{t}_{competence}$ in a given competence (Fontenele and Sun, 2016). Gradation rates are defined by the variable $M \in (0,2]$. Using such an interval for M implies that $TWS^{t}/WS^{t} \leq 50$. If a participant near such a threshold receives a negative evaluation, WS^{t+1} can reach a null or negative value and the fomula is not able to determine a negative reputation. Likewise in *evaluate peers*, if $WS^{t+1}_{competence}$ reaches a non-positive value due to continuous negative evaluations from contractors, $WR^{t+1}_{competence}$ receives a symbolical value (i.e. 0.001) until the participant receives positive work ratings (cf. N6.5). The outputs $TWS^{t+1}_{competence}$, $WS^{t+1}_{competence}$ and $WR^{t+1}_{competence}$ are stored as *#competence level.* An additional qualitative appraisal may be delivered, which is stored as *#contractor comment*.

<Affordance evaluate_project_outcomes>

<StartNorm ID=N6.1>
<whenever>evaluation of a *P* is required after the completion of the *P*</whenever>
<if>*#worker_list* of the *P* exists AND the *contractor* role *exists*</if>
<then>*contractor*</then>
<is>obliged</is>
<to>appraise performance of each contracted expert within the *P*</to>
</StartNorm>

<FinishNorm ID=N6.2>
<whenever>evaluation of a *P* is required after the completion of the *P*</whenever>
<if>*#worker_list* of the *P* terminated OR the *contractor* role ceased</if>
<then>*contractor*</then>
<is>obliged</is>
<to>discontinue appraisal tasks</to>
</FinishNorm>

<OperationalNorm ID=N6.3>
<whenever>contracted experts within *P* are required to be evaluated</whenever>
<if>the expert's project log exists</if>
<then>*contractor*</then>
<is>obliged</is>
<to>assess each expert in the *#worker list* as follows:
calculate $TWS^{t+1}_{competence}$ = $TWS^{t}_{competence}$+1
AND

$$WS^{t+1}_{competence} = \begin{cases} WS^{t}_{competence} + M, \text{where } M = 2 \text{ (high rating) OR} \\ WS^{t}_{competence} + M, \text{where } M = 1 \text{ (medium rating) OR} \\ WS^{t}_{competence} + (-1/100) \times TWS^{t+1}_{competence} \times M, \text{where } M = 2 \text{ (low rating)} \end{cases}$$

</to>
</OperationalNorm>

<OperationalNorm ID=N6.4>
<whenever>a *participant* is subject to evaluation by contractors</whenever>
<if>$WS^{t+1}_{competence} > 0$</if>
<then>*reputation engine*</then>
<is>obliged</is>
<to>calculate

$$WR^{t+1}_{competence} = \alpha \times w + (1 - \alpha) \times y$$

where:
$\alpha = 1/(1 + exp^{-(\log TWS^{t+1}_{competence} - k)})$, where *k*=2
$w = WS^{t+1}_{competence} / TWS^{t+1}_{competence}$
$y = 1/\left(1 + exp^{-(\log WS^{t+1}_{competence} - k)}\right)$, where *k*=2
AND
update *#competence level*{$TWS^{t+1}_{competence}$, $WS^{t+1}_{competence}$, $WR^{t+1}_{competence}$} and *#contractor comment*
</to>
</OperationalNorm>

<OperationalNorm ID=N6.5>
<whenever>a *participant* is subject to evaluation by contractors</whenever>
<if>$WS^{t+1}_{competence} \leq 0$</if>
<then>*reputation engine*</then>
<is>obliged</is>
<to>update *#competence level* with $TWS^{t+1}_{participant}$, $WS^{t+1}_{participant}$=0.001, $WR^{t+1}_{participant}$=0.001 AND *#contractor comment*</to>
</OperationalNorm>
</Affordance evaluate_project_outcomes>

Figure 5.20: The norms of *evaluate project outcomes* (adapted from Fontenele and Sun, 2016)

5.2.5 The Context of "Feedback for Training Plans"

This is the category of semantic units in which opportunities for training, improving capabilities and evaluation feedback are highlighted via *update participant experience*, *flag successful project* and *feedback for training plans* (Figure 5.21). In line with the need for feedback within ecosystems, the purpose of these affordances is to benefit multiple actors and roles (e.g. participant, contractor and organisation) with up-to-date information.

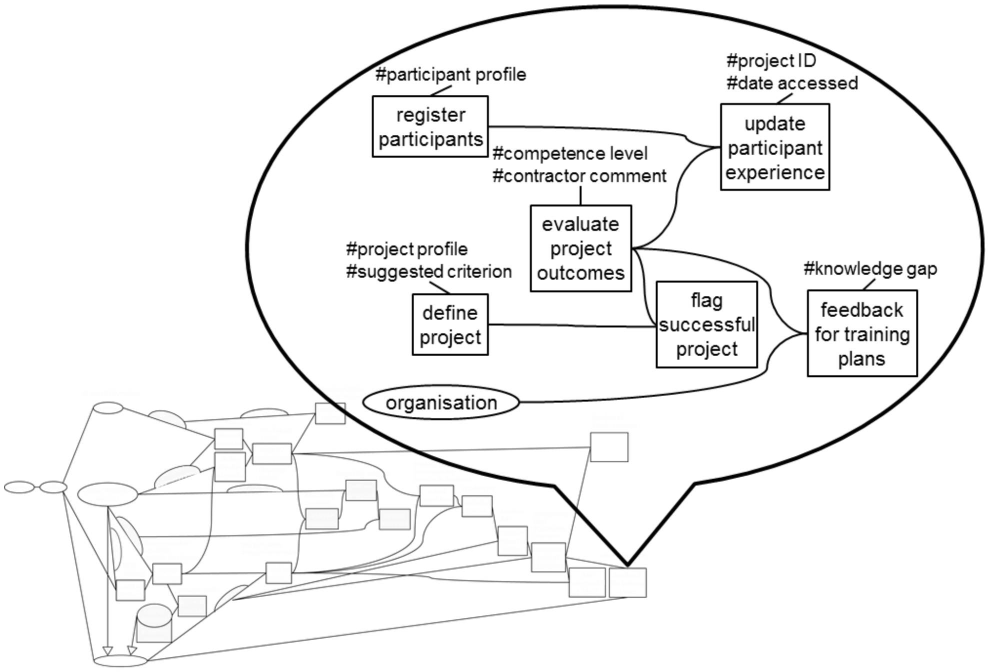


Figure 5.21: The affordance of *feedback for training plans*

The first semantic unit updates the *Experience* (i.e. *Competence level* and *Contractor comment*) in the *participant profile* because of the performance evaluation in a CS project. The affordance *update participant experience* also updates *#date accessed* and uploads the *#project ID* into the *participant profile*. Iterative evaluations maintain the quality and reputation of an expert. The affordance *flag successful project* changes the status of *P.Completed project*. Such a change of status enables project *P* to participate in the matching process described in *define project*. In the third semantic unit, *#contractor comment* will be processed by *feedback for training plans* to the actor of *organisation* to produce recommendations in order to improve further expertise discovery as well as to enhance CS projects in the future (Fontenele and Sun, 2016). Those criteria in a *project profile* that were not met with the *#candidate list* are similarly identified as a *#knowledge gap* and forwarded to the organisation. Therefore, such an affordance supports development of complementary capabilities, which is a challenge for business ecosystems.

5.2.6 The Semantic Unit of "Perform Analytics"

The semantic units containing the workflows of *analyse capability*, *perform analytics* and *update analytics* fall into this context (Figure 5.22). The purpose of *perform analytics* is to promote the ongoing improvement of the dynamic CS knowledge environment through awareness of the expertise gap and shortage based on descriptive analytics. Such an awareness enables the steering of the CS community to take proper action and improve the CS strategy employment. In line with business ecosystem's challenges, these affordances deliver feedback (via a set of *#capability indicators*) to ensure alignment with the strategy. It is worth noting that disclosure of such indicators within the CS community is recommended according to business drivers (e.g. "transparency on goals, outcomes and intentions" and "feedback", cf. Table 5.2).

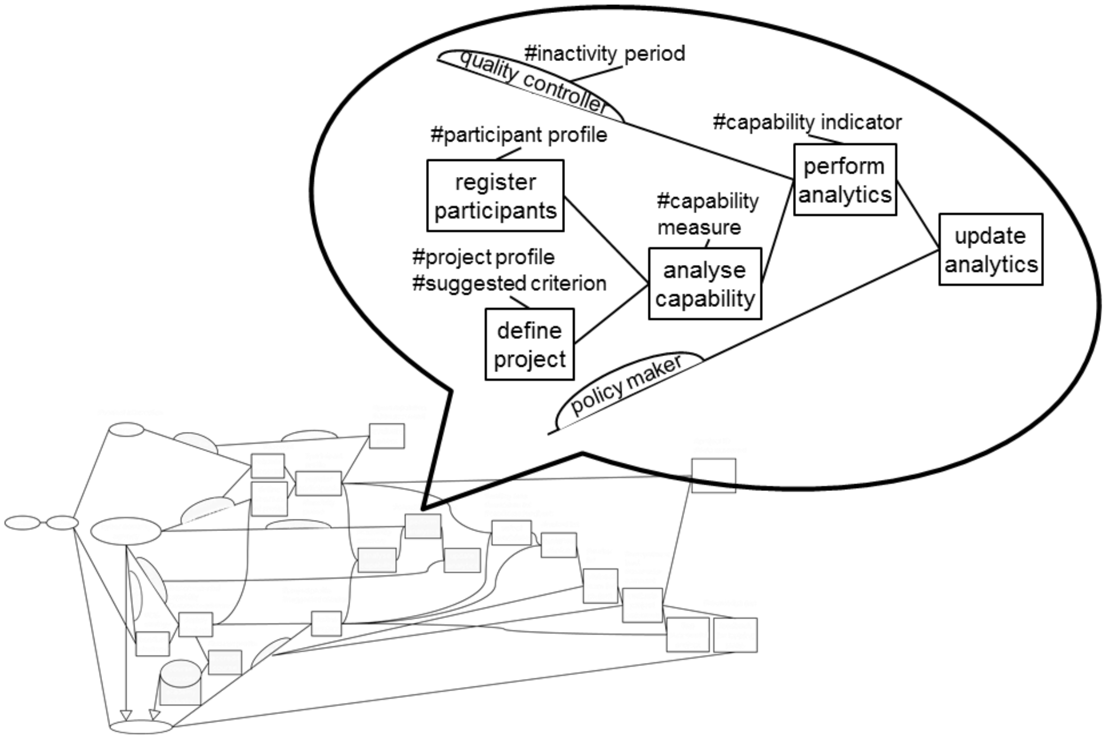


Figure 5.22: The affordance of *perform analytics*

DECYSE enables project and participant profiles as a valuable resource to perform information analytics in order to measure the expertise and the knowledge gap and shortage. Since these profiles are constantly changing or becoming obsolete, in this thesis, the existing capabilities in both participant and talent pools (i.e. transactional data) are measured on a daily basis. However, the indicators, which constitute summarised data, can be provided at a longer rate.

Prior to determining the *#capability measure*, unused participant profiles and obsolete/unsuccessful project profiles are filtered through the *analyse capability* workflow.

When talent and project pools are scaled to efficiency, the workflow computes the existing capabilities in the respective profiles. Therefore, the available expertise in the pool of experts and the capabilities requested in projects can be periodically measured in terms of supply and demand. Such values are stored as a *#capability measure*. A *quality controller* queries such measures via *perform analytics* in order to compute a *#capability indicator* (i.e. a result indicator). Those indicators are forwarded to the *policy maker* via *update analytics*. The purpose of such an affordance is to update the recommended CS capabilities further. Those analysed capabilities are defined and suggested in *deploy strategy*, although new values can be created in *define project* because of new knowledge demands.

As previously discussed, competence, academic area and certification constitute major types of (qualitative) capabilities for CS (cf. Table 5.4 and Figure 5.14). Such types of capabilities appear in a number of participant profiles and in a number of projects. A capability measure (*CpM*) profiles the number of participants and projects holding a given capability in a given time. *CpM* is the output of *analyse capability* and a result indicator formally represented as a tuple in Eq. (9). Those values are linked to a timestamp, since *analyse capability* captures a snapshot of the available and requested CS expertise. The timestamp can be used to perform static and dynamic expertise analysis. The former enables to analyse different capabilities in a given moment in time, while the latter enables to analyse a given capability during a period of time.

$$CpM = \{CpT, CpV, nPa, nP, ts\} \tag{9}$$

where:
CpM is the capability measure
CpT is the capability type
CpV is the capability value
nPa is the number of participants with a given *CpV*
nP is the number of projects with a given *CpV*
ts is the timestamp in which the analysis was performed

In order to improve quality in data analytics and to scale the talent and project pools to retain only the relevant profiles, the *analyse capability* workflow (Figure 5.23) periodically archives unused participant profiles and obsolete project profiles. Such a process analyses values in the talent pool (i.e. *#date accessed*) and project pool (i.e. *P.Start date*) in comparison with the current date (i.e. the timestamp *ts*). Since the creation of the profile (cf. Figure 5.15), each time a *participant* interacts with CS actors (cf. Figure 5.16, Figure 5.19 and *update participant experience*), *participant profile#date accessed* is updated. Any profile instance that has not been accessed over a given *#inactivity period* (specified by the quality controller) is then removed from the talent pool. Therefore, participants are encouraged to become active and useful actors for the CS community. Likewise, if *P.Start* *date* is older than the *inactivity period*, the project profile is archived. It is suggested that the affordance *analyse capability* should be performed on a daily basis, due to the dynamics of the *#date accessed* variable and because the *inactivity date* variable is given in a number of days, but such a recommendation is not mandatory. In addition, the granularity of *#capability indicator* depends on the *#capability measure*.

All available capabilities ($CpV_{Participant}$) under a given type (*CpT*) are further counted as *nPa*, while capabilities existing in successful projects ($CpV_{Project}$) are counted as *nP*. The affordance *analyse capability* merges the outputs of both antecedents, where $CpV_{Project}$ = $CpV_{Participant}$, creating a *CpM* tuple (cf. Eq. (9)).

<Affordance analyse_capability>
<StartNorm ID=N7.1>
<whenever>*participant profile* and *project profile* are subject to analysis of capabilities</whenever>
<if>*participant profile* exists AND *project profile* exists</if>
<then>*analytics engine*</then>
<is>permitted</is>
<to>analyse activity of *participant profile* and *project profile*</to>
</StartNorm>

<FinishNorm ID=N7.2>
<whenever>*participant profile* and *project profile* are subject to analysis of capabilities</whenever>
<if> *participant profile* ceased OR *project profile* terminated</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>discontinue activity analysis of *participant profile* and *project profile*</to>
</FinishNorm>

<OperationalNorm ID=N7.3>
<whenever>*participant profile* is subject to analysis for archiving</whenever>
<if>*inactivity period* is specified</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>retrieve the list {*Participant profile.Profile ID, Participant profile.Date accessed*}</to>
</OperationalNorm>

<OperationalNorm ID=N7.4>
<whenever>N7.3 is satisfied</whenever>
<if>*inactivity period* is specified</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>archive $Participant\ profile_{Profile\ ID}$ | $Participant\ profile_{Profile\ ID}$.*Date accessed* < *current date* - *inactivity period,* where *current date* refers to when the analysis is carried out </to>
</OperationalNorm>

<OperationalNorm ID=N7.5>
<whenever>*project profile* is subject to analysis for archiving</whenever>
<if>*inactivity period* is specified</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>retrieve the list {*P.Project Id, P.Start date*}</to>
</OperationalNorm>

<OperationalNorm ID=N7.6>
<whenever>N7.5 is satisfied</whenever>
<if>*inactivity period* is specified</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>archive $Project\ profile_{Project\ Id}$ | *P.Start date < current date - inactivity period,* where *current date* refers to when the analysis is carried out</to>
</OperationalNorm>

<OperationalNorm ID=N7.7>
<whenever>CS capabilities are analysed in both *participant profile* and *project profile*</whenever>
<if>N7.4 is satisfied AND N7.6 is satisfied</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>retrieve {$CpT, CpV_{Participant}$} = ({*Participant profile#Competence*} | *CpT*="competence") ∪ ({*Participant profile#Academic Area*} | *CpT*="academic area") ∪ ({*Participant profile#Certification*} | *CpT*="certification")
AND
retrieve $CpV_{Project}$ = ({*P.Capability.Criterion*}, where *P.Completed project* is flagged
AND
calculate *nPa*=n($CpV_{Participant}$)
AND
calculate *nP*=n($CpV_{Project}$)
AND
store *CpM* ={$CpT, (CpV_{Project} \cup CpV_{Participant})$*, nPa, nP, ts*}, where *ts* is the *current date*
</to>
</OperationalNorm>
</Affordance analyse_capability>

Figure 5.23: The norms of *analyse capability*

The choice for periodically creating a tuple with the values instead of performing the queries directly on the project and participant pools is due to keep track of the changes in capabilities over time. Such a decision complies with the constant changes in both types of profiles.

The *CpM* tuples are a valuable resource to create capability indicators (i.e. the KRI in the context of this thesis). In compliance with the research problem, these indicators measure the CS capability gap or shortage, which are calculated and as a demand / supply ratio and ranked accordingly. The adopted capability indicators and respective norm ID are presented in Table 5.6 to illustrate the use of *CpM*. Those capability indicators are mapped when the *quality controller* queries *#capability measure* via *perform analytics* and computes knowledge gaps and shortages on the different types of capabilities.

Table 5.6: List of capability indicators

Norm ID	Capability indicator description
N8.3	Current demanded capabilities in projects with no available expertise
N8.4	Competence within highest demand/supply ratio
N8.5	Academic area within highest demand/supply ratio

Norm ID	Capability indicator description
N8.6	Certification within highest demand/supply ratio
N8.7	Adoption (in projects and by participants) of the competency with the highest demand/supply ratio during a fortnight

If the expertise supply (i.e. the number of experts with a given capability) equals zero and a demand occurs (i.e. at least one project requesting such capability), there is an expertise gap (cf. N8.3). Capability gaps require priority action in the CS community. If there is a similarly high demand (i.e. requested in numerous projects) and a low supply of a given capability (i.e. at least one expert), this is considered an expertise shortage. In this case, the demand / supply ratio is calculated according to each type of capability (i.e. the qualitative criteria for expertise discovery in Table 5.4), which generates three indicators (cf. N8.4, N8.5 and N8.6). Since competencies are the building blocks of the other capabilities, in the context of this thesis, an indicator is devised to analyse the behaviour of the competence with the highest demand/supply ratio during a certain period (i.e. N8.7). These five capability indicators demonstrate the flexibility provided by *#capability measure*. In this thesis, the KRI are structured through norms, so that pieces of code may be reused when appropriate. For example, N8.7 relies on the first result from N8.4.

The *perform analytics* workflow (Figure 5.24) determines capability indicators using persistent data from *#capability measure*. The process is triggered according to a *quality controller* in order to produce summarised data. As previously discussed in Chapter 3, the use of dashboards to present the indicators can facilitate visualisation and support further actions to balance the demand/supply ratio for CS expertise.

```
<Affordance perform_analytics>
<StartNorm ID=N8.1>
<whenever>capabilities are measured to support analytics</whenever>
<if>analyse capability exists AND quality controller exists</if>
<then>analytics engine</then>
<is>obliged</is>
<to>provide indicators based on capability measure</to>
</StartNorm>

<FinishNorm ID=N8.2>
<whenever>capabilities are measured to support analytics</whenever>
<if>analyse capability ceased OR quality controller terminated</if>
<then>analytics engine</then>
<is>obliged</is>
<to>discontinue creation of indicators based on capability measure</to>
</FinishNorm>

<OperationalNorm ID=N8.3>
<whenever>analytics are performed on capability measures</whenever>
<if>new CpM instances are created AND quality controller requests capability indicators</if>
<then>analytics engine</then>
```

<is>obliged</is>
<to>retrieve *CpM* tuples AND sort *CpM* by *nP* in descending order AND retrieve Capability ID and {*CpV, nP*}, where *nPa*=0 and *ts*=*t*</to>
</OperationalNorm>

<OperationalNorm ID=N8.4>
<whenever>analytics are performed on capability measures</whenever>
<if>N8.3 is satisfied</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>retrieve *CpM* tuples AND Calculate *ratio*={*nP/nPa* | (*nP*, *nPa*)>0 and *ts*(*t*)} AND sort *CpM* by *ratio* in descending order AND retrieve Capability ID and the {*CpV*, *nPa*, *nP*}, where *CpM.CpT*="competence"</to>
</OperationalNorm>

<OperationalNorm ID=N8.5>
<whenever>analytics are performed on capability measures</whenever>
<if>N8.4 is satisfied</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>retrieve *CpM* tuples AND calculate *ratio*={*nP/nPa* | (*nP*, *nPa*)>0 and *ts*(*t*)} AND sort *CpM* by *ratio* in descending order AND retrieve Capability ID and the {*CpV*, *nPa*, *nP*}, where *CpM.CpT*="academic area"</to>
</OperationalNorm>

<OperationalNorm ID=N8.6>
<whenever>analytics are performed on capability measures</whenever>
<if>N8.5 is satisfied</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>retrieve *CpM* tuples AND calculate *ratio*={*nP/nPa* | (*nP*, *nPa*)>0 and *ts*(*t*)} AND sort *CpM* by *ratio* in descending order AND retrieve Capability ID and the {*CpV*, *nPa*, *nP*}, where *CpM.CpT*="certification"</to>
</OperationalNorm>

<OperationalNorm ID=N8.7>
<whenever>analytics are performed on capability measures</whenever>
<if>N8.6 is satisfied</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>retrieve *CpM* tuples AND Retrieve Capability ID and {*nP, nPa, ts*}, where *CpM.CpV*=N8.4(CpV_1) and *ts*(*t*-*m*)|*m* ∈ ℕ and 0≤*m*≤15}</to>
</OperationalNorm>

<OperationalNorm ID=N8.8>
<whenever>analytics are performed on capability measures</whenever>
<if>N8.7 is satisfied</if>
<then>*analytics engine*</then>
<is>obliged</is>
<to>compute the capability indicators based on N8.3 to N8.7 AND present results on a dashboard</to>
</OperationalNorm>
<Affordance perform_analytics>

Figure 5.24: The norms of *perform analytics*

After updating the dashboard, results are forwarded to *policy maker* via *update analytics*. Those results serve manifold purposes such as updating *#recommended capability*, developing actions to balance demand and supply for knowledge and promoting self-awareness for the CS community as a whole.

5.3 Summary

In this chapter, there has been a detailed discussion about the development and presentation of DECYSE. The use of enterprise architecture illustrated the contrast between the “as is” context and the conceptual elements “to be” integrated through the DECYSE ontological model. In order to set the former context, an empirical research was employed to investigate the current practices for expertise discovery and existing gaps within the CS expertise discovery lifecycle. The main types of criteria required to discover talents (e.g. competence, competence level, academic area and certification) were identified and can be used with DECYSE. Moreover, expertise discovery for the CS community requires a thorough understanding of the stakeholders’ needs and the services they can provide for each other. Thus, the problem articulation was structured in steps to develop the DECYSE method (i.e. the meta-processes) so that they can be revisited in the future in order to ensure continuous improvement on the method itself.

An overview on the DECYSE ontological model unravelled how the expertise discovery lifecycle is covered and integrated by our solution via five major affordances. The DECYSE ontological model was discussed through norm-based processes, their purpose, algorithms, underpinning techniques and profiles. More than simply providing inputs for affordances, the processes in DECYSE constitute a series of information loops that contribute to the ongoing improvement of CS expertise. DECYSE is hence a knowledge management system that contributes to 1) integrate existing CS expertise discovery principles and actors’ information needs; 2) scale up expertise discovery processes, which are normally used on a single agency scale, to the CS community and 3) embark feedback loops to keep information up-to-date with the dynamic CS environment. A discussion on some of the data collected to articulate the research problem and to test DECYSE is described in the following chapter.

Chapter 6

Applying DECYSE in Context

DECYSE is tested within the context of a Brazilian CS community. The empirical research was conducted for a threefold purpose: 1) triangulating findings and determining requirements to be conceptualised in the DECYSE method; 2) profiling data to experiment the DECYSE method; and 3) capturing appraisals from experts for the proposal. Prior to the method experimentation, a brief discussion introduces data collected for the articulation of the research problem (cf. Section 4.2.4) and for the experiment. Such a discussion contextualises and sheds light on the relevance of the problem. Then, DECYSE is tested according to selected affordances described in the previous chapter. Finally, the appraisal performed by a board of CS experts complements the validation of test results, which is presented in Chapter 7.

6.1 Data Collection for Articulating the Research Problem

There have been three sets of data collected for the experiment of DECYSE. The first set includes a pilot questionnaire (Appendix G), which had the purpose to design the context and triangulate findings in literature. Such questionnaire resulted in 238 answers. The second set, which relates to the articulation of the research problem, is introduced in this section and includes a questionnaire (Appendix H) and an interview (Appendix I). The latter was conducted with two senior military officers working in Brazilian CS projects. Its purpose was to complement the questionnaire's answers, by capturing the participants' unique views and experience in the CS community. The third set, related to the experiment, is introduced in the following section. A questionnaire in Appendix H is used to collect data from private sector, research centre, public agencies, academy and Armed Forces. These data are used to gain clear perceptions of a nature and characteristics of CS that leads to define and scope this study. Sixty-nine participants answered this questionnaire (Figure 6.1) and sixty-six valid answers were obtained (three participants either did not work or did not expect to work with cyber security).

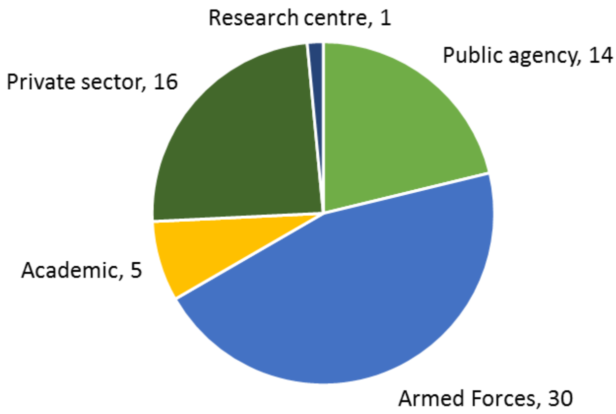


Figure 6.1: Participant's background for articulating the research problem

Some major issues in dealing with CS were identified during the development of the conceptual model. For example, 88% of the participants believe that the career of the CS professional is still not well defined. In addition, 68% of participants are not aware of any existing solution to manage talents to CS, although 27% admit that there are some non-integrated solutions available. There were 94% of participants who agreed that integrated Talent Management is crucial to increase national CS. The majority of participants think that regulations and scientific production regarding CS is scattered or decentralised (Figure 6.2a). Participants also believe that integration is a way of raising awareness on population (Figure 6.2b), identifying talents (Figure 6.2c) and creating job opportunities (Figure 6.2d). Analysis of variance was performed resulting in $F(3,260)=2.837$, $p<0.05$.

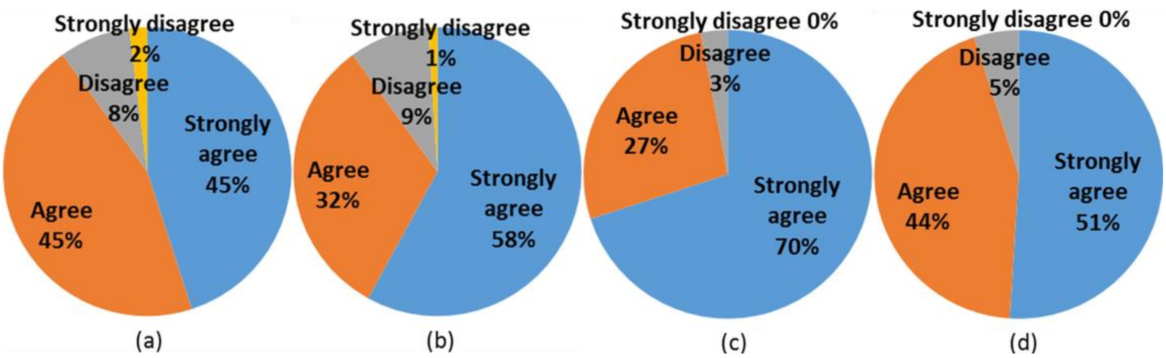


Figure 6.2: (a) Scattered documentation, (b) raising awareness, (c) identifying talents, (d) job opportunities

Also consistent with literature findings (cf. Figure 5.2), the key stakeholders in CS were identified by participants on the second questionnaire, as depicted in Figure 5.3.

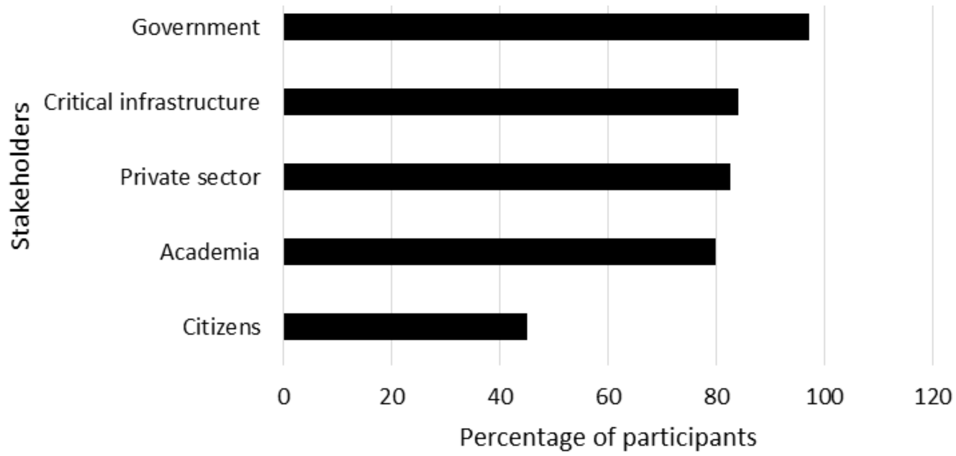


Figure 6.3: The most relevant cyber security stakeholders

Sourcing expertise for CS projects (Figure 6.4) was also investigated during the development of the conceptual model, in order to understand how structured the search for experts is. Since no evidence was found in documents, participants in the questionnaire (Appendix H) were asked which information sources do they use or recommend when searching for expertise. Results show that searching for CS expertise is mostly unstructured and builds on interactions and on a web of trust, since the majority of participants rely in networking and recommendations from peers.

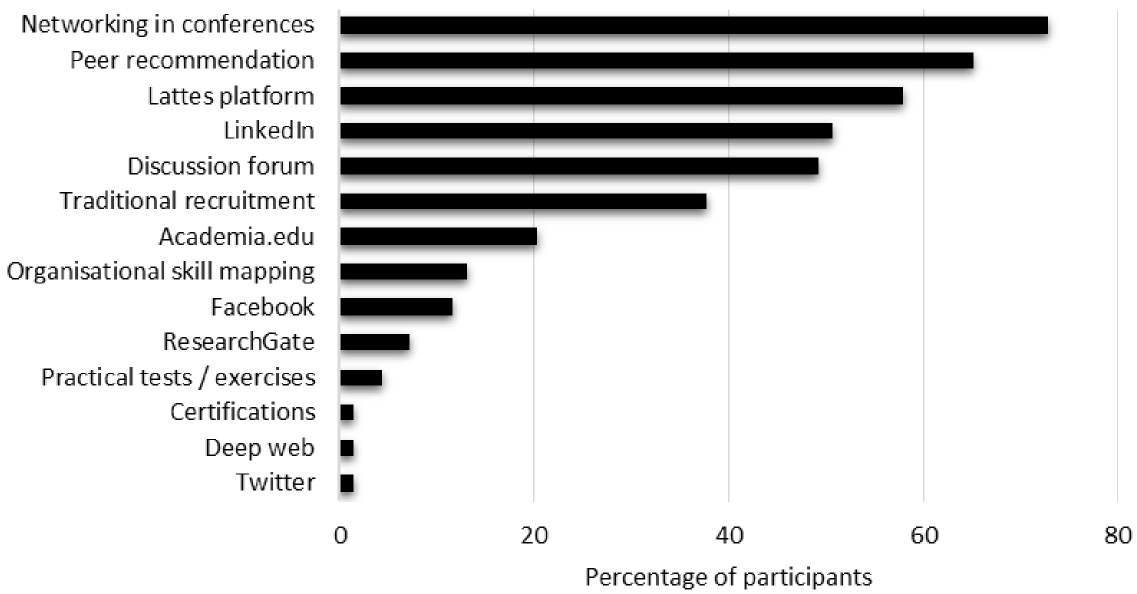


Figure 6.4: Sources for discovering cyber security experts

The criteria required for expertise discovery captures the most relevant features used by the CS community when contracting experts in the field. Such criteria represent the pervasive information needs underpinning discovery of relevant capabilities within the CS community. This section includes the results that determined the nature and types of the criteria (presented in Table 5.4) which led to the design of the profiles in the DECYSE method.

Participants in the first questionnaire emphasised the importance of competencies, skills, certifications, academic background and experience as important requirements for the CS professional and aligned with findings in Section 2.1. Participants in the second questionnaire also pointed out the most relevant criteria for selecting personnel to work in CS, as depicted in Figure 6.5.

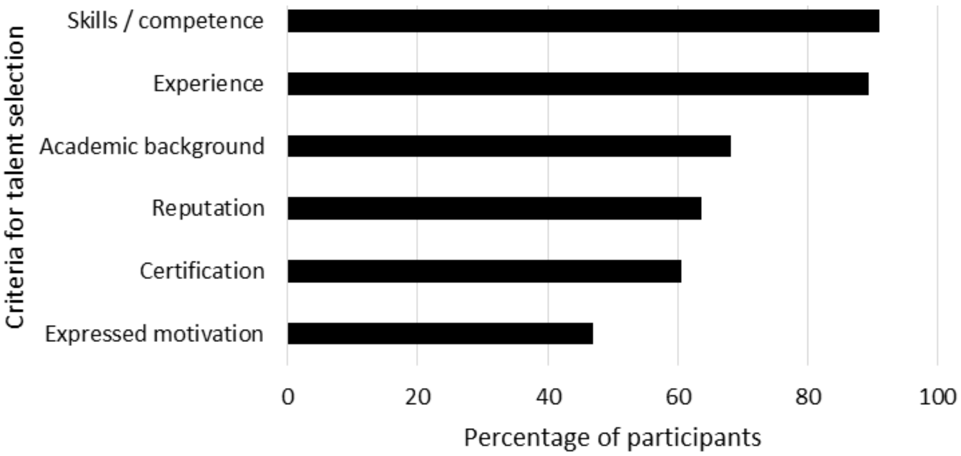


Figure 6.5: Most relevant criteria to select talent for cyber security

A couple of semi-structured interviews (cf. Appendix I) had the purpose to elicit further details of criteria for expertise discovery. Both interviews (I1 and I2), along with open-ended questions on the questionnaire (cf. Appendix H), contributed to define the current participant profile. The interviewees are retired and experienced senior military officers actively working in Brazilian CS projects (e.g. the Brazilian National School for Cyber Defence). The following paragraphs highlight discussions on the expertise discovery criteria.

Skills and competencies are the main criteria when searching for experts to work in CS projects (cf. Figure 6.5). In fact, the following interview excerpt emphasises the importance of such a capability:

> *"I would look mainly for skills and competencies ... competencies can be developed, but not some abilities. ... If I would have to select (a person), I would not look at his certifications, but first at ... skills, interests and his network. ... The most important (criterion) is skill." (I2)*

Our approach considers a set of CS competencies from which a participant can claim into his / her participant profile. Moreover, an initial set of competencies should be recommended to the CS community. In compliance with general practice (e.g. CS frameworks presented in Chapter 1), this thesis assumes that competencies should be defined in terms of a set of knowledge, abilities, skills and tasks for a common understanding.

Experience usually refers to the length of time or the number of times that a person has been actually performing a specific task. Experience becomes even more complex to measure when considering the quality of results. For example, when comparing two individuals who performed the same number of similar tasks, it becomes a challenge to measure how much one individual outperforms the other. The following excerpt highlights the importance of experience for CS and the challenge to measure it:

> *"I have elicited which are the activities that one does more often. ... This should be one of the criteria. For how long he has been working on the same organisation within a same level or within a same function? ... How will you know if the guy is experienced? ... You have to map the function he is in, because he may be there for 10 years, but he may be working just one year with security. ...You should look at (the individual's) working place to see if he had any kind of problem. ... You have to keep track of (one's career). ... The quantity of awards (an individual) has achieved in a given place doing a specific task is important, because you then have his real profile." (I1)*

I1 argues that awards are a means to recognise outstanding performance, but they can be highly subjective to bias and normally do not constitute an objective way of measuring experience in a given competence. Our method measures accrued experience for each competence claimed by an individual. The accrued experience results from participation in CS projects where a particular competence is required. Each contractor's assessment has an influence (positive or negative) on the experience gained in a particular competence. Our perspective on experience, therefore, considers project performance in which a given competence has been employed, rather than the duration in terms of time of such projects.

CS experts normally rely on knowledge gained from experience and competencies. However, the academic background seems important for longer projects where a specific body of knowledge is desirable and when the CS project involves research or managerial level requirements. In fact, a growing number of academic courses are dedicated to develop CS professionals (cf. Table 2.1). Despite having the knowledge, it is more important to know how to implement it in practice, as quoted in the following interview excerpts:

> *"I think (academic criterion) should be qualitative from both sides. Not only the level of the academic background, but also per area." (I1)*
>
> *"When you are searching for someone to work on a given area, if he is a doctor or if he is a master, it does not make much of a difference. The difference is if he knows how to do (the task)" (I1)*
>
> *"The research area that the person is involved... demonstrates (the type of work) that the person tends to go to… (the) area that he is more related, more than certification. Research area should be the second choice (to search for CS*

experts)... Higher (academic) degree when you consider management level is one thing, but in the operational level, it is another thing. ... In a managerial profile, an academic is relevant ...” (I2)

DECYSE considers both academic degree and area to search for candidates for CS projects. Although I1 focuses on a more pragmatic approach towards the criteria for CS, in our approach, the academic background is not subject to evaluation from contractors, but rather if a person has graduated or not in a given academic area. Therefore, in order to make sure that a candidate knows about the task (as requested by I1), contractors can also rely on other criteria (e.g. experience and reputation) when specifying CS project requirements.

CS involves dealing with sensible issues (e.g. security and defence). Hence, trustworthiness is a key asset when searching for experts. Even social networking sites do not seem to be the proper source for trust, despite being recognised as a valuable information source. Therefore, informal collaboration networks seem to currently play a crucial role to select people to work in CS projects and may be more important than having relevant courses in the curriculum, as described in the following interview excerpts:

“To work on the grand events, we have made a search basically upon our network. This means, an already known collaboration network ... I don´t remember to rely on any ... (social media) application. I didn´t use this. ... I have begun to meet people from certain places that also have important connections, and so your network grows… In this area of security and defence, you have to first establish a trusted relationship ... I can assure you that this is 90% (of the solution). Sometimes you know many people but you cannot establish a trusted relationship. ... Thus, it is the main criterion for me: to establish a trusted relationship within this network. ... You can put at least ... a basic endorsement ... without being invasive.” (I1)

“The relationship network was not asked about, right? ... The word reputation is properly employed. ... The social network in which the individual belongs is very important. Normally the network is more selective than courses” (I2)

I1 relies primarily on informal networks, recommendations and basic endorsements rather than information on social media. However, the behaviour of candidates in social media is one of the sources used for building a good reputation. Endorsement from agencies was also claimed as important information to be profiled. Thus, it seems that personal recommendations (e.g. reputation built on peer reviews) and collaboration network (e.g. work reputation) plays a major role for talent selection. In our approach, a participant’s profile determines reputation considering perceptions from peers (as a result from interactions in a closed social media) and employers (because of the expert’s performance in CS projects).

Certifications are a valuable source of skills for CS, as discussed in Chapter 2 and cf. Figure 6.5. Certifications can serve as a confirmation that an individual has mastered specific competencies and are preferred over academic courses when dealing with technical issues that require rapid action (e.g. incident response). The Brazilian Government (2013a) lists a set of recommended certifications for information security professionals. An excerpt from such a list is presented in Appendix E.

> *"(There) are good courses that provide you with the background focused in a given area. It helps you on your daily work, but is still not a certification… for example, CERT.br ... (provide a set) of short courses that lasts around a week ... where the work is very intense ... and then at the end you can master the skills." (I1)*

> *"Courses related to security, ... certified or not (are relevant). ... I always tell about CQI ..., which means capacity, qualification and indication. ... Qualification is achieved through courses ... and capacity is what (a person) is already born with or develops ... to execute an activity. ... I would suggest up-to-date certifications ..., in this case, it would be ok, I would use them (to search for experts) if they could be renewed... In this case, they are relevant. ... Certification is important as long as it is kept up-to-date." (I2)*

Certifications in our approach include professional industry standards (e.g. the list in Appendix E) and may include short courses that provide specific skills (as referred to in I1). Notwithstanding, no list of officially recommended short courses were found during the research work. In contrast to Fontenele and Sun (2016), the number of certifications held by a participant who matches project requirements is also considered as a type of criterion.

Expressed motivation in Figure 6.5 refers to the availability of an expert to join a CS project. In order to adopt a more flexible approach in this criterion, availability is defined as a date rather than a binary "yes" or "no". For example, let us suppose that a project starts on a day d and a highly experienced candidate with a required capability is available on $d+2$. It might be more appropriate for a contractor to hire such a candidate instead of relying on a candidate with less experience. In another scenario, there might be no candidates holding a given criteria by the start of the project, so the candidate available within the least amount of time could be suitable for selection. The DECYSE method articulates the dates in which a candidate is available with the project start in order to balance experience levels with availability.

Therefore, results in the second questionnaire confirmed the need to manage talents for CS and the stakeholders involved as presented in the literature review. There were no academic works found during this research devising a strategic Talent Management solution to support CS and offering assessment functionalities.

One of the major challenges in conceptualising the CS domain is defining the set of instances for the required capabilities. Indeed, take for example the following interview excerpt:

> *"One of the first things that should be provided [in the near future] is the recommended skills and competencies list. (...) We do not have a skills list, which is necessary." (I2)*

Some agencies already defined their own set of CS competencies and skills (e.g. US Government, 2014). However, findings have shown that this is a task yet to be performed in most nations (such as in the Brazilian CS setting). Therefore, DECYSE acknowledges such request as a service (i.e. "deploying CS strategy" in Table 5.1) and uses a given set of values when profiles are further discussed. However, it is beyond the scope of this research work to provide a detailed list of skills and competencies for manifold reasons. Amongst these motivations, there are already frameworks providing such values and the relevant topics are dynamically defined by expertise supply and demand relations within the CS community. The competencies and other capabilities used in this research solely fulfil the purpose to test DECYSE. The method also includes analytic processes to determine knowledge supply and demand over time.

Some additional criteria elicited during the interviews and questionnaires refer to personal characteristics that could be measured as a general competence, but could affect sensitive or private information. For example, when asked about any additional requirements or restrictions to select personnel for CS projects, the following answer was given:

> *"I think teamwork is critical" (I1)*

In such case, a person can be technically qualified for a project; however, he may not be suited to work in a team. If such information leaks, it could affect his personal image. Moreover, some participants in the second questionnaire emphasised the need to conduct further investigation into talents depending on the sensibility level of the information to which they might have access. Although "teamwork ability" can be measured as a competence, such a type of personal feature along with further investigation on candidates may lead to a disclosure of sensible and personal information, which is beyond the scope of this research. In summary, the capabilities adopted as criteria for expertise discovery are listed in Table 5.4.

6.2 Data Collection for the Experiment of DECYSE

In order to experiment the DECYSE method, a dataset for the talent pool and project profiles was collected using Appendix J. During this round of experiments, 66 participants have joined the talent pool (i.e. answered the questionnaire) during a time span of 19 days (from 28 January until 15 February), as depicted in Figure 6.6. The dates in which participants joined are particularly important for capability indicator results, as further discussed.

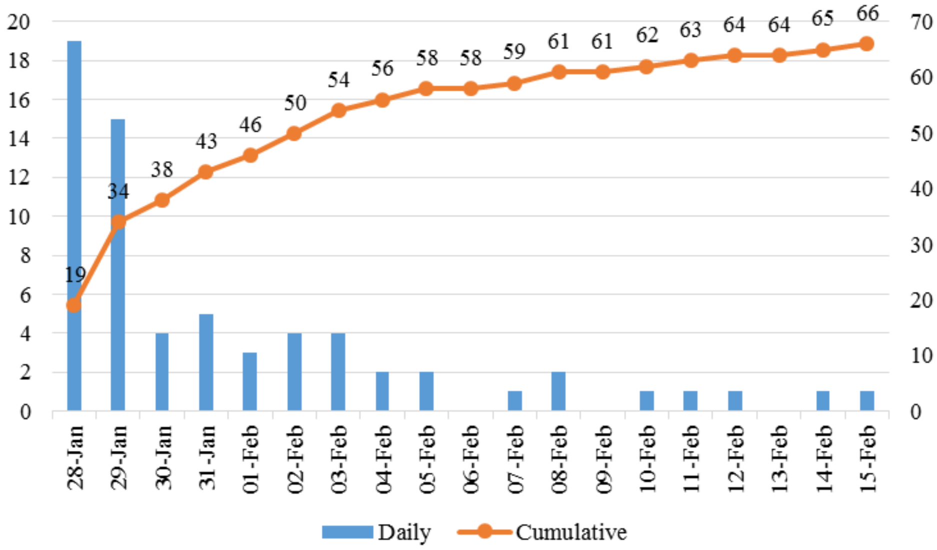


Figure 6.6: Participants joining the talent pool

Data cleansing was performed to remove those incomplete or inaccurate answers. It is worth noting that much of the results hereby discussed are tightly connected to the policy setting, expertise identification and selection for expertise principles, which are discussed in the following section. The participants within the CS community are representing different sectors of society, as shown in Figure 6.7a.

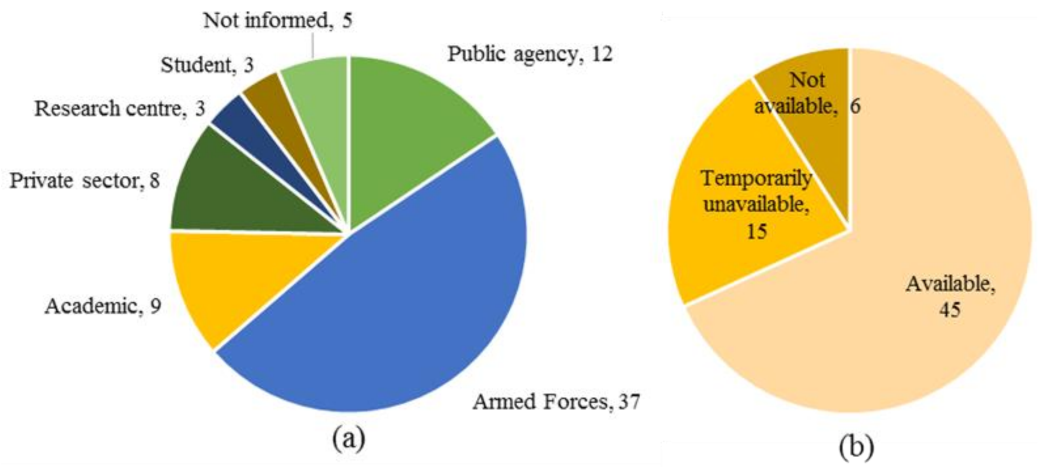


Figure 6.7: Status of participants' (a) background and (b) availability within the cyber security community

It is worth noting that participants were able to inform more than one activity background. Instead of asking about their *availability date*, which would be compared to *P.Start date*, they were asked directly for the number of days until they would be available (i.e. *waiting time*). This procedure was adopted to disregard the time that each participant needed to answer the questionnaire for the *select candidate* affordance. The availability of participants to join CS projects is presented in Figure 6.7b.

Instead of simply removing the six participants who claimed to be unavailable, a *waiting time* value of 1000 days was deliberately attributed to test the sample and check if any of them would be selected for a project. For those temporarily unavailable who did not inform as to when they would change the status to available, a *waiting time* value ranging from 2 to 10 days was randomly assigned.

Participants were asked to check their knowledge background from a list of recommended competencies. They were encouraged to provide additional *Capability* values (cf. Appendix J). The frequency of competencies to be used in the experiment are presented in Figure 6.8. All listed competencies were assigned to at least one profile. Nine additional unlisted competencies were included by participants.

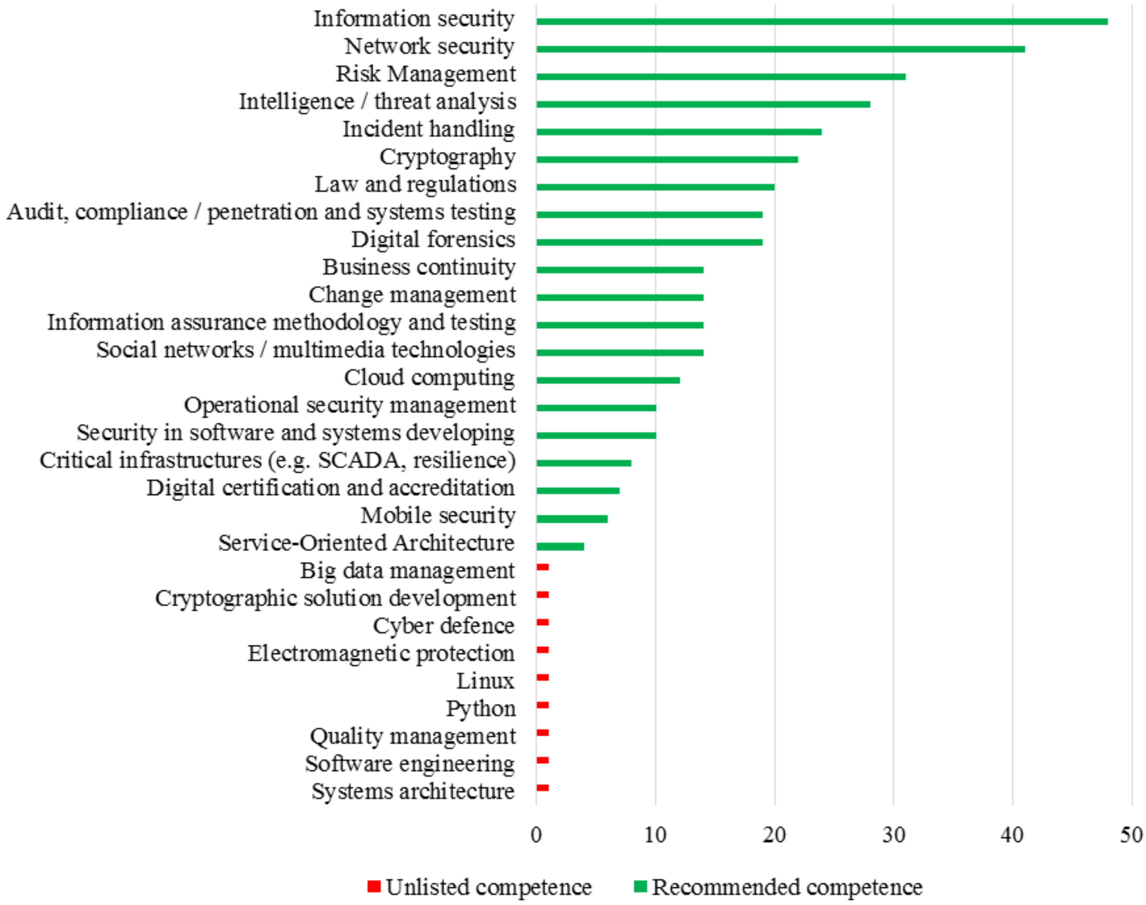


Figure 6.8: Frequency of competencies in participant profiles

Participants similarly informed their certifications from a pre-defined list of recommended capabilities (green bars in Figure 6.9) and informed other unlisted certifications, which they believed to be relevant to the CS community (red bars in Figure 6.9). For instance, three participants are OSCP certified, although such a certification is (still) not considered to be a recommended capability. On the other hand, 23 recommended certifications (e.g. CASP and CBCI) were not assigned to any of the participant profile instances.

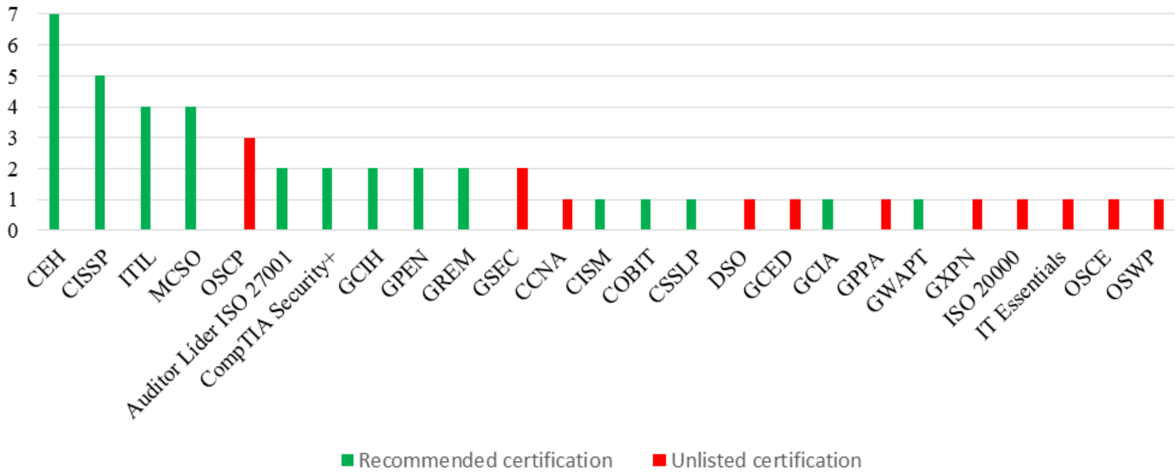


Figure 6.9: Frequency of certifications in participant profiles

Participants were also asked to inform their academic background (i.e. area and degree) according to recommended academic areas resulting in the frequencies in Figure 6.10. Additional academic areas not listed as recommended capabilities (e.g. statistics and management science) were also added.

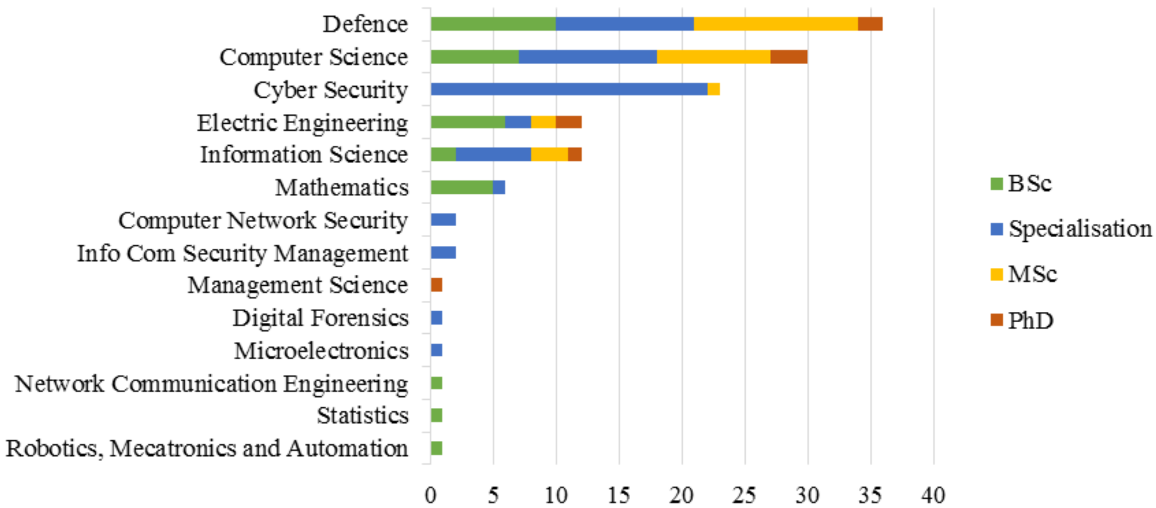


Figure 6.10: Frequency of academic areas and levels in participant profiles

Some profile features (e.g. *Peer reputation* and *Competence level*) cannot feasibly be acquired from real world data, since such data is iteratively improved according to interactions in the CS community. Additional data cleansing was hence performed in order to fill such gaps within the talent pool and to provide a more realistic scenario where evaluations have already occurred. Those values were randomly generated within participant profiles.

The questionnaire also captured excerpts from actual projects related to CS. One of these excerpts was selected to test the method for a twofold reason. First, the criteria derived from different types of capabilities. Second, the selected project included recommended values (e.g. professional certifications in Appendix E) and other values that were not pre-defined. Both reasons add to demonstrating the power and flexibility of DECYSE as presented in the following section.

6.3 Experiment on DECYSE

This section highlights the main results when applying DECYSE in the adopted context. The discussion is conducted in line with DECYSE's major affordances.

6.3.1 Specifying a Project through "Define Project"

In this work, the initial set of *#recommended capabilities* were suggested by the *policy maker* via *deploy strategy* as follows. Eight recommended values for the *Academic area* are mainly drawn from CNPq (2015). Corresponding *Academic degrees* are arbitrated as 0.001 (none), 0.3 (BSc), 0.4 (specialisation[9]), 0.5 (MSc) or 0.7 (PhD). Thirty-seven certifications for IT professionals (Appendix E) constitute recommended values for *Certification*. Likewise, twenty competencies obtained from Table 2.1 comprise the recommended values for *Competence*.

For the sake of simplicity, *#recommended capabilities* is provided as a list, disregarding the conceptual relations in Table 5.5; instantiating such relationships is beyond the scope of this thesis (cf. Section 5.2.2) and does not affect the tests with DECYSE.

A CS project *P.Project Id (1_2016)* was specified in terms of *P.Capability.Criterion* and *P.Capability.Priority* described in Figure 6.11. Priorities were defined as the result of a pairwise comparison in Table 6.1, with a consistency ratio of 0.022. The weights in Table 6.1 were randomly determined since no project in the questionnaire results defined weights to their criteria.

[9] A Brazilian graduate degree, which represents a specialisation in a specific area and takes one to two years to achieve. It is also known as a "*lato sensu* postgraduate" degree in contrast to "*stricto sensu* postgraduate" degree (e.g. MSc and PhD).

```xml
<?xml version="1.0" ?>

<Project>
<Project_profile Project_ID="1_2016">
		<Capability Criterion=”Digital forensics”>
			<Priority>0.302954709</Priority>
		</Capability>
		<Capability Criterion=”Audit, compliance and systems testing”>
			<Priority>0.302954709</Priority>
		</Capability>
		<Capability Criterion=”Information science”>
			<Priority>0.123817925</Priority>
		</Capability>
		<Capability Criterion=”OSCP”>
			<Priority>0.066821716</Priority>
		</Capability>
		<Capability Criterion=”CEH”>
			<Priority>0.123817925</Priority>
		</Capability>
		<Capability Criterion=”CHFI”>
			<Priority>0.037904103</Priority>
		</Capability>
		<Capability Criterion=”OSWP”>
			<Priority>0.041728915</Priority>
		</Capability>
</Project_profile>
</Project>
```


Figure 6.11: Excerpt of the project specification for 1_2016

The project 1_2016 requires knowledge of two competencies (i.e. “digital forensics” and “audit, compliance and systems testing”) and an academic background in Information Science. In addition, four certifications (i.e. OSCP, CEH, CHFI and OSWP) were requested. It is worth noting that the remaining criteria (i.e. competence level, peer reputation, waiting time and the number of certifications, cf. Table 5.4) are dynamically weighted according to Data Envelopment Analysis principles. The academic degree is also used as an ordinal variable; thus, it is considered in order to determine the efficiency of the Decision-Making Units.

Table 6.1: Pairwise comparison among project 1_2016 capabilities

Capability	**Competence A (and level)**	**Competence B (and level)**	**Academic area and degree**	**Certification A**	**Certification B**	**Certification C**	**Certification D**
Competence A (and level)	1.00	1.00000	3.00000	5.00000	3.00000	7.00000	5.00000
Competence B (and level)	1.00	1.00	3.00000	5.00000	3.00000	7.00000	5.00000
Academic background	0.33	0.33	1.00	3.00000	1.00000	3.00000	3.00000
Certification A	0.20	0.20	0.33	1.00	0.33333	3.00000	3.00000
Certification B	0.33	0.33	1.00	3.00	1.00	3.00000	3.00000
Certification C	0.14	0.14	0.33	0.33	0.33	1.00	1.00000
Certification D	0.20	0.20	0.33	0.33	0.33	1.00	1.00

It is worth noting that during the *define project* workflow; no additional capabilities were suggested since no *P.Keyword* was established to determine similarity with other projects.

However, a brief example of a suggestion for additional capabilities seems appropriate. Table 6.2 describes the variables and respective arbitrated values to illustrate such a feature. Considering both sets of keywords, P^{7_2015} is a successfully concluded project that is suitable for reuse due to its similarity (80% is acceptable) with project 1_2016. Amongst the values in $P^{7\_2015}.Capability$, three capabilities are suggested as additional criteria for project 1_2016 (i.e. Forensics, Law and Regulations and Cryptography). The contractor decides whether these suggested criteria are going to be incorporated into project 1_2016 (cf. Figure 6.11) or not.

Table 6.2: Example of suggesting additional criteria for project 1_2016

Variable	**Value**
$P^{1\_2016}.Keyword$	{audit, password recovery, encrypted container, forensics}
$P^{7\_2015}.Keyword$	{audit, password recovery, encrypted container, forensics, identity theft}
Simil (P^{1_2016}, P^{7_2015})	80%
$P^{7\_2015}.Capability$	{Forensics, Digital Forensics, Law and Regulations, Information Science, Cryptography, CHFI}
#suggested criterion	{Forensics, Law and Regulations, Cryptography}

6.3.2 Providing Expertise Identification with "Register Participants"

The *#recommended capabilities* are afforded during *participant profile* creation via the *define structure of profile*. In this thesis, such an initial set of *#recommended capabilities* was presented in the first four questions in Appendix J. During the *register participants* process, participants were able to enter their *Contact information*, *Availability date*, *Competence*, *Certification*, *Academic Area* and respective *Academic Degree* in their *Participant profiles*. The *Date accessed* was defined according to the day that the account was created (i.e. when participants answered the questionnaire). The results from applying the *register participants* process stands for the data collection for the experiment, which is presented in Section 6.2. Figure 6.12 illustrates an excerpt of profiled information from participant 42, which is captured in a XML file. The participant holds two academic degrees, four certifications and four competencies. Some of these capabilities match the project requirements, as discussed in the following section.

```
<?xml version="1.0" ?>

<Participant>
<Participant_profile Participant_ID="42">
        <Capability>
            <Certificate>
                <Academic_type>
                    <Academic_area>Defence</Academic_area>
                    <Academic_degree>Specialisation</Academic_degree>
```

```
                <\/Academic_type>
        </Certificate>
        <Certificate>
                <Academic_type>
                        <Academic_area>Cyber_Security</Academic_area>
                        <Academic_degree>Specialisation</Academic_degree>
                </Academic_type>
        </Certificate>
        <Certificate>
                <Industrial_type>
                        <Certification>GCIH</Certification>
                </Industrial_type>
        </Certificate>
        <Certificate>
                <Industrial_type>
                        <Certification>GPEN</Certification>
                </Industrial_type>
        </Certificate>
        <Certificate>
                <Industrial_type>
                        <Certification>OSCP</Certification>
                </Industrial_type>
        </Certificate>
        <Certificate>
                <Industrial_type>
                        <Certification>CEH</Certification>
                </Industrial_type>
        </Certificate>
        <Experience>
                <Competence>Network_security</Competence>
                <Competence_level>{60,62,0.679}</Competence_level>
        </Experience>
        <Experience>
                <Competence>Incident_handling</Competence>
                <Competence_level>{21,24,0.525}</Competence_level>
        </Experience>
        <Experience>
                <Competence>Digital_forensics</Competence>
                <Competence_level>{3,46,0.131}</Competence_level>
        </Experience>
        <Experience>
                <Competence>Critical_infrastructures</Competence>
                <Competence_level>{30,33,0.577}</Competence_level>
        </Experience>
        <Peer_reputation>{6,47,0.186}</Peer_reputation>
    </Capability>
    <Date_accessed>31/1/2016</Date_accessed>
</Participant_profile>
</Participant>
```


Figure 6.12: Excerpt from participant profile instance

After the completion of the *register participants* process, experts may be evaluated according to their contributions and knowledge sharing in social media. Since deploying such a platform was deemed beyond the scope of this research, Table 6.3 illustrates possible outcomes when the *evaluate peers* workflow is performed. Such an appraisal shows how three peer reviewers (e.g. participants 1, 22 and 45) with different reputations affect $PR_{participant\ 13}$ according to positive or negative ratings.

Table 6.3: Possible peer evaluation for participant 13

Peer reviewer	$PR^t_{peer\ reviewer}$	Before appraisal: Participant 13 $\{PS^t, TPS^t, PR^t\}$	After appraisal: Participant 13 $\{PS^{t+1}, TPS^{t+1}, PR^{t+1}\}$	
			Positive rating	Negative rating
1	0.997	{102, 121.79, 0.677}	{102.997, 122.787,0.6785}	{100.7858, 122.787, 0.66796}
22	0.509	{102, 121.79, 0.677}	{102.509, 122.299, 0.6778}	{101.3801, 122.299, 0.67237}
45	0.258	{102, 121.79, 0.677}	{102.258, 122.048, 0.6774}	{101.6858, 122.048, 0.67465}

Figure 6.13 illustrates the changes in $PR_{participant\ 13}$ according to Table 6.3. It is noticeable that the higher the peer reviewer's reputation, the higher the impact on PR^{t+1}. Negative ratings significantly decreases $PR_{participant\ 13}$ in contrast to a positive rating, as expected. Therefore, the appraisal from Participant 1, who has the highest *PR*, delivers the most significant changes for $PR_{participant\ 13}$ in relation to participants 22 and 45.

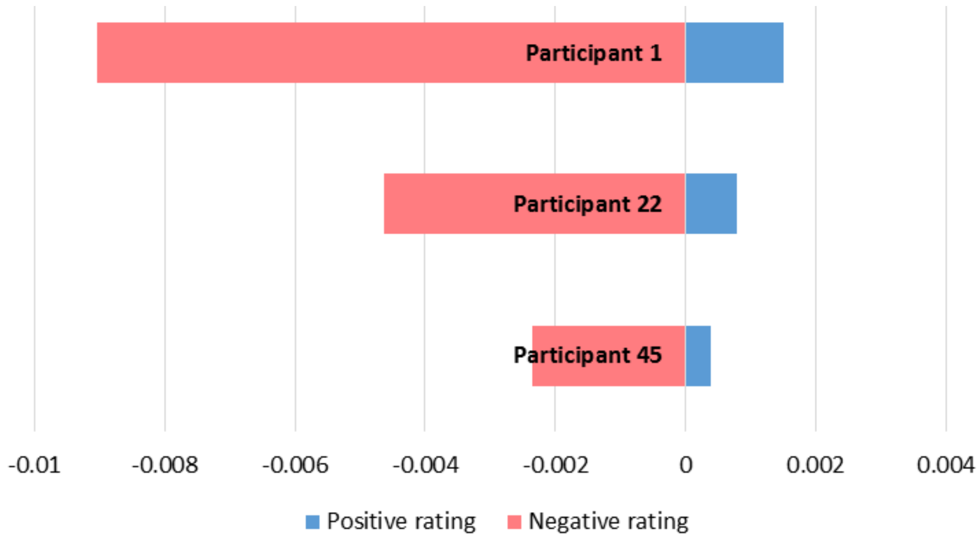


Figure 6.13: Impact of ratings from different peer reviewers in a *PR*

6.3.3 Performing Selection and Evaluation of Expertise via "Select Candidate"

A set of one input and five outputs is used to perform the selection of expertise for project 1_2016. The outputs for the *select candidate* correspond to the participant's levels in requested competencies (i.e. "Digital Forensics" and "Audit, compliance and systems testing") and a degree from the required academic area (i.e. "information science") obtained from project 1_2016. In addition, n(*Cert*), *Peer reputation* and the *Waiting time* are added as default outputs and input, respectively. Among the 66 participants registered in the talent pool, 13 efficient Decision-Making Units were discovered. Table 6.4 presents a subset of Decision-Making Units and the values used to compute their relative efficiency.

Table 6.4: The partial results from *Select candidate* process for project 1_2016

DMU	Input: Waiting time x_1	Output: Peer reputation (PR) y_1	Digital Forensics (WR) y_2	Audit and Compliance (WR) y_3	Academic degree y_4	Number of certifications y_5	Weighted: Input $\sum_i v_i x_{i0}$	Output $\sum_r u_r y_{r0}$	Maximum efficiency $\max \frac{\sum_r u_r y_{r0}}{\sum_i v_i x_{i0}}$
1	1	0.997	0.001	0.001	0.001	0.001	1	1	1
2	1	0.812	0.001	0.54	0.001	1	1	1	1
4	1	0.53	0.464	0.001	0.001	1	0.842	1	0.842
7	1	0.831	0.001	0.001	0.001	0.001	0.834	1	0.834
10	1	0.65	0.881	0.937	0.001	0.001	1	1	1
11	2	0.33	0.001	0.001	0.001	1	0.302	1	0.302
12	1	0.769	0.689	0.281	0.001	0.001	0.951	1	0.951
13	180	0.677	0.001	0.724	0.001	0.001	0.005	1	0.005
15	1	0.337	0.278	0.849	0.001	0.001	0.888	1	0.888
17	1	0.315	0.001	0.001	0.7	0.001	1	1	1
18	300	0.186	0.001	0.001	0.001	0.001	0.001	1	0.001
22	1000	0.509	0.001	0.581	0.001	0.001	0.001	1	0.001
24	1	0.542	0.606	0.956	0.001	1	1	1	1
26	365	0.12	0.718	0.001	0.001	0.001	0.002	1	0.002
28	4	0.291	0.001	0.001	0.001	0.001	0.073	1	0.073
29	1	0.825	0.035	0.001	0.001	0.001	0.836	1	0.836
31	1	0.81	0.169	0.057	0.001	0.001	0.856	1	0.856
33	1	0.148	0.001	0.001	0.5	0.001	0.715	1	0.715
38	1	0.886	0.001	0.001	0.001	0.001	0.889	1	0.889
41	1	0.926	0.001	0.001	0.5	0.001	1	1	1
42	1	0.186	0.131	0.001	0.001	2	1	1	1
43	1	0.836	0.593	0.44	0.001	0.001	1	1	1
45	1	0.258	0.9	0.001	0.001	1	1	1	1
51	1	0.954	0.001	0.001	0.001	0.001	0.957	1	0.957
57	1	0.528	0.962	0.001	0.001	0.001	1	1	1
58	1	0.756	0.933	0.323	0.001	0.001	1	1	1
59	1	0.911	0.001	0.171	0.001	1	1	1	1
63	1	0.072	0.001	0.001	0.5	0.001	0.715	1	0.715
64	1	0.359	0.716	0.225	0.001	2	1	1	1
66	1000	0.455	0.893	0.366	0.001	0.001	0.001	1	0.001

During the *select candidate* process, those 13 efficient Decision-Making Units (cf. Table 6.4) have had their profile updated (via *Date accessed* and *Candidate feedback*) with the current date and information about the project 1_2016 (e.g. goal and requested capabilities). Such an activity follows the provision of feedback principle and enables the candidates to know that they are in a recruiting process and what expertise they are missing.

Computing peer reputation can improve the quality of candidates, although such criterion does not need to be explicitly declared in project profiles. For example, there were two efficient candidates who graduated in information science: Decision-Making Units 17 and 41. The former scores the highest academic degree value (i.e. PhD), while the latter holds an MSc in the area. Decision-Making Unit 41 was deemed efficient in contrast to the other candidates holding an MSc (i.e. Decision-Making Units 33 and 63) due to their higher reputation.

Despite having both competencies and an average peer reputation, Decision-Making Unit 66 received a very low efficiency ratio mainly due to the declared unavailability (i.e. the *waiting time* value of 1000). All other participants who were temporarily unavailable also resulted in inefficient Decision-Making Units (e.g. Decision-Making Units 11 and 13). In fact, only those currently available participants were deemed efficient candidates.

More than one Decision-Making Unit satisfied the selection criteria. These Decision-Making Units were then ranked according to the best fit as per the project requirements and were updated on the project details. Table 6.5 presents the results that were produced according to the qualitative criteria in order to enable further ranking and analysis of these Decision-Making Units.

Table 6.5: The ranking about optimal expert choice for the project 1_2016

Candidate	**Digital forensics**	**Audit and compliance**	**Information science**	**OSCP**	**CEH**	**CHFI**	**OSWP**	**Alternative weight**	**Rank**
Priority weight	0.30295	0.30295	0.12382	0.06682	0.12382	0.03790	0.04173		
64	0.9	0.9	0.1	0.9	0.1	0.1	0.9	0.671559	1
24	0.9	0.9	0.1	0.9	0.1	0.1	0.1	0.638175	2
10	0.9	0.9	0.1	0.1	0.1	0.1	0.1	0.584719	3
43	0.9	0.9	0.1	0.1	0.1	0.1	0.1	0.584719	3
58	0.9	0.9	0.1	0.1	0.1	0.1	0.1	0.584719	3
42	0.9	0.1	0.1	0.9	0.9	0.1	0.1	0.494871	4
2	0.1	0.9	0.1	0.1	0.9	0.1	0.1	0.441415	5
45	0.9	0.1	0.1	0.1	0.9	0.1	0.1	0.441415	5
59	0.1	0.9	0.1	0.1	0.9	0.1	0.1	0.441415	5
57	0.9	0.1	0.1	0.1	0.1	0.1	0.1	0.342359	6
17	0.1	0.1	0.9	0.1	0.1	0.1	0.1	0.199055	7
41	0.1	0.1	0.9	0.1	0.1	0.1	0.1	0.199055	7
1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.099999	8

As one may notice, Candidate 64 ranked first, while being the most fit candidate for the project and matching four out of seven explicit criteria. If the project were restricted to only one expert, Candidate 64 would be the best fit for the job. The list continues until it reaches the lowest ranked Candidate 1, who fulfilled no explicit criterion despite achieving a high peer reputation. Since the contractor opted to create a team to address all criteria, Candidates 17, 42 and 64 were considered as the best-fit experts for the project 1_2016, even though no candidate matched CHFI. In fact, no single participant has achieved such a certification so far (cf. Figure 6.9). Candidate 64 is the first choice for contraction, since their total alternative weight outranked the other candidates. There were two candidates with an Information Science background. Candidate 17 overlapped capabilities and alternative weights with Candidate 41. As a matter of illustration, the former was selected due to a preference for a higher academic degree over the higher peer reputation of the latter. Finally, concerning the remaining criterion, Candidate 42 was chosen as the highest priority holding CEH. Hence, Candidates 17, 42 and 64 become susceptible to evaluation after the conclusion of *Project Id*(1_2016).

When the project was successfully completed, the contractor evaluated the project outcomes by rating 42 and 64 referring to the competencies based upon which they were selected. Candidate 17 was not rated, since the type of criterion by which he was selected (i.e. an academic area) is not subject to evaluation. Notwithstanding, the team's appraisals regarding the application of the remaining criteria (e.g. academic background and certifications) are stored as *Contractor comment*. Possible evaluation results for each competence generated the set of values in Table 6.6.

Table 6.6: Possible competence evaluation for members of project 1_2016 team

Team member	Competence before start of project (*t*)		Competence evaluation after the end of project (*t+1*)		
	Digital Forensics $\{WS, TWS, WR\}$	Audit and Compliance $\{WS, TWS, WR\}$	High rating $\{WS, TWS, WR\}$	Medium rating $\{WS, TWS, WR\}$	Low rating $\{WS, TWS, WR\}$
42	{3,46,0.131}	-	{5,48,0.168}	{4,47,0.151}	{2.08,48,0.109}
64	{100,108,0.716}	-	{102,110,0.719}	{101,109,0.717}	{97.84,110,0.698}
	-	{18,130,0.225}	{20,132,0.236}	{19,131,0.231}	{15.4,132,0.206}

The impact of different ratings on each resulting evaluation is depicted in Figure 6.14. The smoothing factor of the algorithm ensures that *WR* values remain between 0 and 1. It is noticeable how lower levelled competencies (e.g. 0.131) have a stronger significance in terms of absolute values, whether regarding positive or negative appraisals. Positive ratings in lower *WR* (e.g. 0.131) therefore have a higher impact in a similar way to negative ratings in higher *WR* (e.g. 0.716). For example, participant 42 would have an increase of 0.037 in $WR_{Digital\ Forensics}$ because of a high rating in contrast to a decrease of 0.022 with a low rating. Participant 64 similarly would have an increase of only 0.003 in $WR_{Digital\ Forensics}$ because of a high rating against a decrease of 0.018 with a low rating.

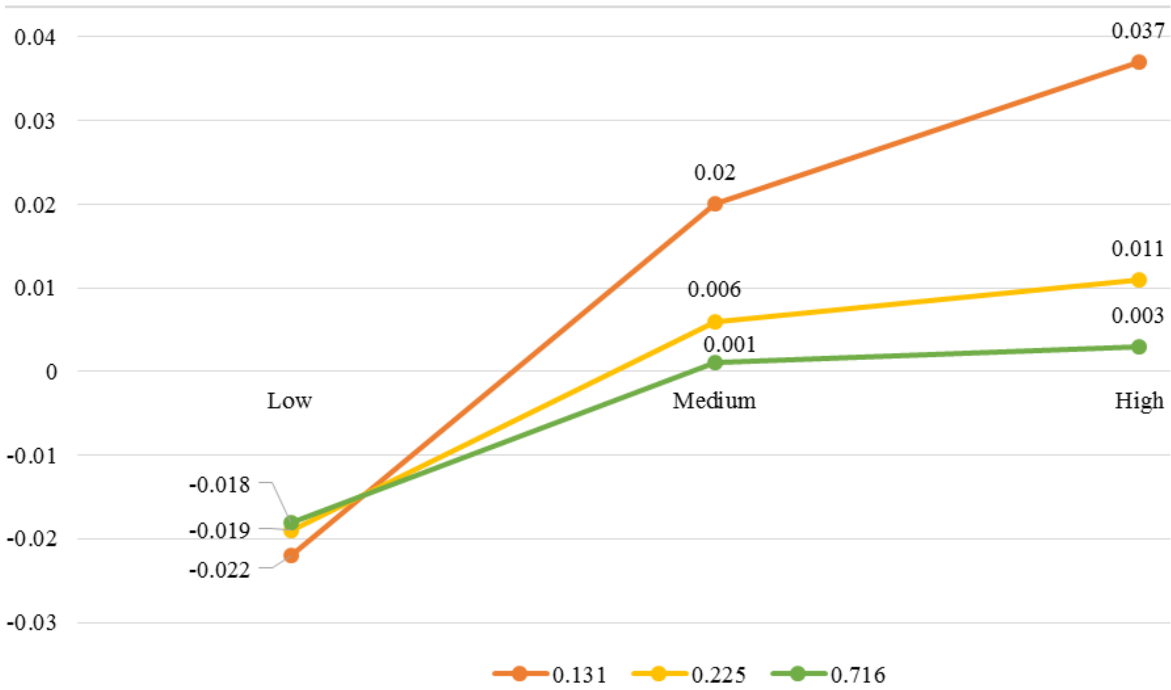


Figure 6.14: Impact of ratings on different *WR* values in *t+1*

6.3.4 Delivering Provision of Feedback with "Feedback for Training Plans"

Three actions are triggered post the evaluation of team members. Firstly, the new *Competence levels* (cf. Table 6.6), *Contractor comments*, *Project Id*(1_2016) and *Date accessed* are updated to the team members' profiles via *update participant experience*. Secondly, $P^{1\_2016}.Completed\ project$ is flagged via *flag successful project*, which makes features in project 1_2016 susceptible for reuse in future project definitions. Thirdly, although project 1_2016 was completed, CHFI was identified as a *#knowledge gap* in the pool of experts, since it was the only missing criterion in project 1_2016. Such information is provided to the contracting organisation via *feedback for training plans*.

Whether a missing criterion may only affect the quality of project outcomes or jeopardise the whole project, it is up to the contractor to decide. Carrying on with the project and leaving CHFI as a requested criterion in project 1_2016 not only alerts the contracting organisation, but also improves awareness of the CS knowledge gap for other actors, as discussed in the following section.

6.3.5 Improving Expertise Analysis using "Perform Analytics"

When performing analytics, the *analyse capability* workflow generates a set of capability measure (*CpM*) tuples on a daily basis. The *ts* values for the experiment consider the dates on which participants answered the questionnaire, which illustrates participants joining the talent pool and creating their profiles. Each *participant profile* was updated once during data collection and no participant joined the talent pool on 6, 9 and 13 February. This means that the set of *CpM* tuples generated on 5, 8 and 12 February respectively have the same values for {*CpT*, *CpV*, *nPa*, *nP*}.

The *perform analytics* workflow was triggered by the *quality controller*. The *analytics engine* queries the capability measure *CpM* tuples according to the norms that describe the indicators (cf. Table 5.6). Such a procedure enables the creation the dashboard containing a set of capability indicators (i.e. the KRI in the context of this thesis) which resulted in Figure 6.15, Figure 6.16 and Figure 6.17. The graphs in the first two figures represent the status and comparison of the main types of CS capabilities, while the graph in the latter shows the competency with highest demand/supply ratio (i.e. Network Security) along a timeline.

The dashboard highlights the capabilities that constitute a knowledge gap (cf. N8.3 and Figure 6.15), i.e. presents those capabilities requested in projects that were not found in participant profiles. For example, CHFI was requested in *Project Id*(1_2016) and in one more project, but no participant currently has such a certification. Such an indicator acts on the creation of actions and measures to improve those capabilities and balance the demand/supply ratio.

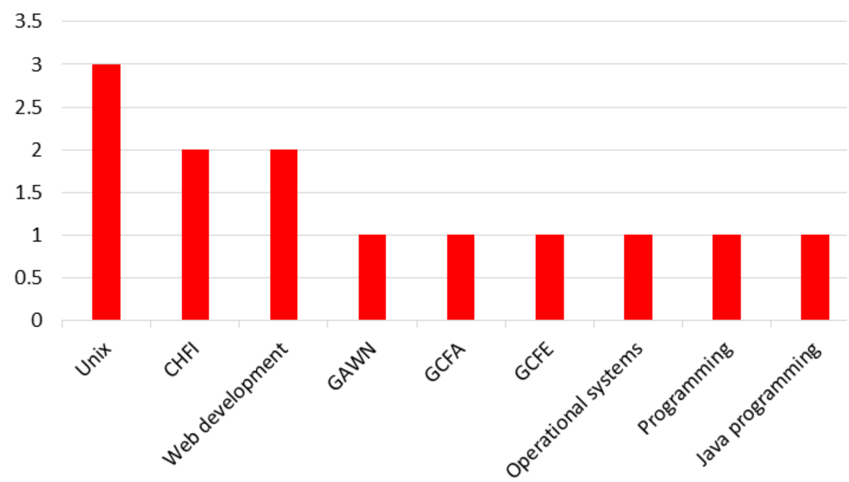
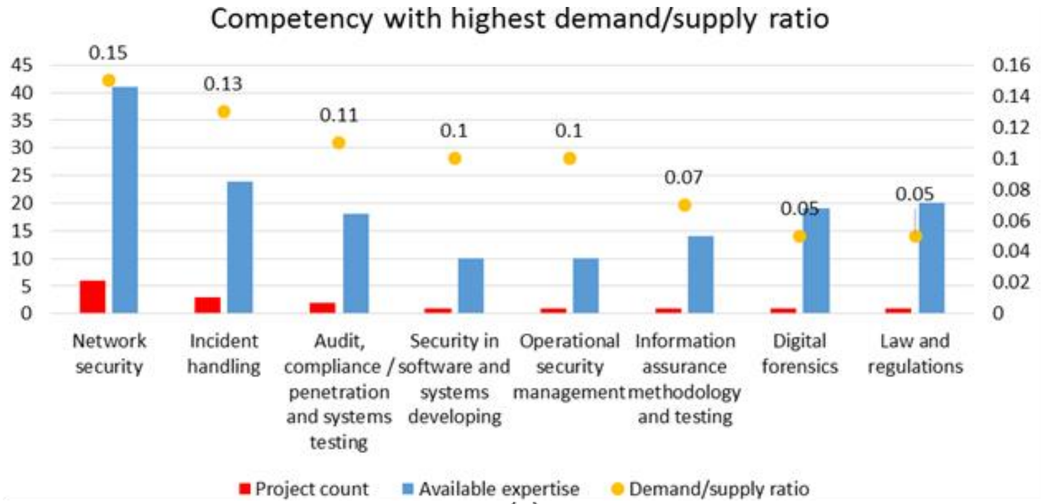
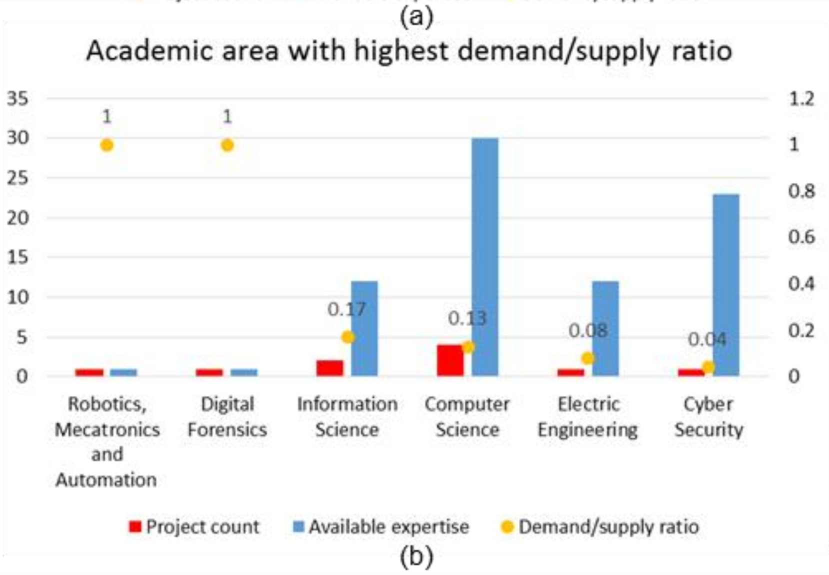


Figure 6.15: The dashboard based on project capabilities with no available expertise (cf. Table 5.6)

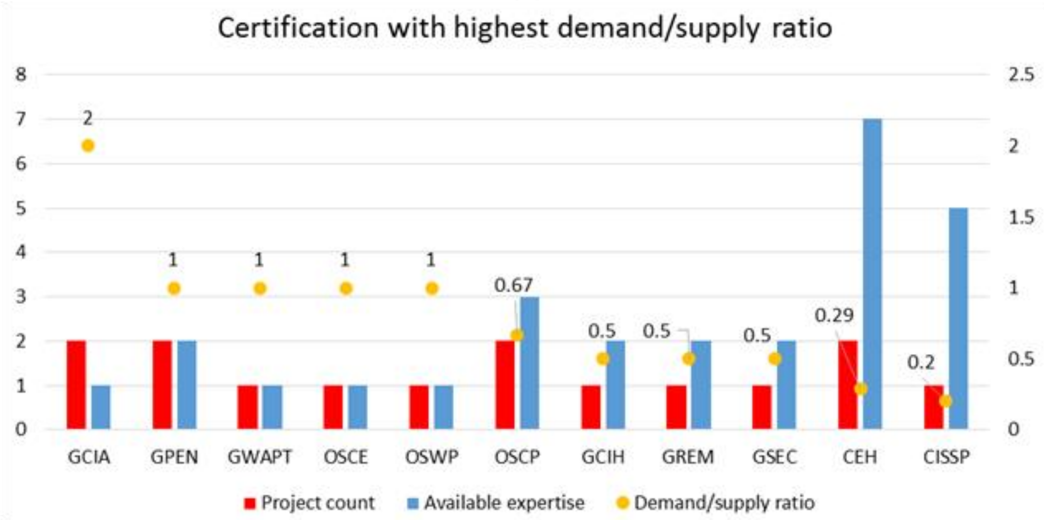
The following three graphs respectively emphasise the most underdeveloped competencies (Figure 6.16a), academic areas (Figure 6.16b) and certifications (Figure 6.16c) within the CS talent pool, which represent the knowledge shortage. The number of participants currently holding a given capability (*nPa*) is depicted as a blue column. The number of projects similarly requesting such capability (*nP*) is depicted as a red column. As shown in those three graphs and specified in N8.4, N8.5 and N8.6, the capabilities are ranked according to their demand/supply ratio, represented by the orange colour on the dashboard. Each *CpV* was therefore classified according to the respective *CpT*.



(a)



(b)



(c)

Figure 6.16: The dashboard based on capabilities with the highest demand/supply ratio (cf. Table 5.6)

Figure 6.17 depicts the demand/supply ratio during a fortnight for Network Security, which is currently the highest relatively demanded competence (cf. Figure 6.16a).

Figure 6.17: Participants and projects with Network Security (measured daily during a fortnight)

The dashboard with capability indicators is presented to the *policy maker* via *update analytics* so that those emerging capabilities can be further conceptualised and structured as a new *#recommended capability*. For example, both digital forensics (cf. Figure 6.16b) and OSCE (cf. Figure 6.16c) are requested in one project and held by one participant each. However, they are not listed as a *#recommended capability* (cf. Appendix E).

6.4 Summary

In this chapter, an application on the DECYSE method has been discussed. Initially, a brief discussion on the data used to develop the method highlighted some of the underpinning ideas behind the method. Afterwards, profile data for the experiment was presented. The collected profiles represent experts from different backgrounds and projects in the CS community. Data cleansing was performed on values that were not feasibly attainable (e.g. reputation). However, the method was capable of processing different types of criteria (i.e. certifications, academic area and competencies) specified in the project profile. The problem of emerging capabilities is managed by further structuring those values as recommended capabilities. Then, the experiment was presented in terms of the actual results and other possible outcomes in order to illustrate the efficiency of the solution. The expertise discovery principles were realised through DECYSE's semantic units. The following chapter builds on the findings to perform the validation of DECYSE and highlights the contributions to the research fields and to industry as well as the method's limitations.

Chapter 7

Critical Evaluation

The empirical validation demonstrated the use of DECYSE as a tool that enables automated support for expertise discovery for the CS community, which received positive appraisals from CS experts. This chapter critically reviews the work performed to discover expertise for the CS community. The discussions comprise the validation of results from experiments, the contributions to theory and industry, the justifications for the adopted research design and an appraisal on the empirical research. In addition, the research evaluation presents some limitations of DECYSE.

7.1 Validating the Method of DECYSE

DECYSE is validated under a fourfold perspective comprising construct, external, internal and content validity. Those perspectives are respectively discussed in the following subsections.

7.1.1 Validating DECYSE with Different Datasets

The validation of a new method calls for a sufficient amount of testing. DECYSE relied on a series of tests and Design Science Research. During iterative refinements, affordances of the DECYSE method were tested/trialled with a test dataset, which was devised simulating real life values and enabled exercising the method under different conditions. For example, some of these experiments involved the *select candidate* process using a single input and two to four outputs (Table 7.1). The one-input and four-output experiment, along with other workflows, was published in the report by Fontenele and Sun (2016). This thesis uses five outputs, including the number of certifications as a default output and employs a different dataset, as described in Section 6.2.

Table 7.1: The outcomes of *select candidate* in different experiments

Number of profiles	Number of inputs	Number of outputs	Number of candidates	Observation
1000	1	2	2	Prior experiment
1000	1	3	6	Prior experiment
1000	1	4	18	(Fontenele and Sun, 2016)
66	1	5	13	Results in this thesis

It seems that the number of inputs and outputs in the *select candidate* process significantly influences the number of efficient candidates in contrast to the number of Decision-Making Units. Indeed, previous experiments using a test dataset with a thousand profiles resulted in eighteen, six and two efficient candidates using four, three and two outputs, respectively (cf. Table 7.1). Analysis of variance was performed in the number of profiles, outputs and candidates resulting on $F(2,9)=10.588$, $p<0.01$.

The reputation algorithms delivered expected results (cf. Figure 6.13 and Figure 6.14) in the same way that experiments were carried out in the report by Fontenele and Sun (2016). All those experiments show the scalability of DECYSE is in terms of the manageable number of resulting candidates (i.e. two to eighteen) and how the method delivers valid and consistent outputs with different datasets.

7.1.2 Validating DECYSE According to its Applicability in other Settings

It is well known that the adoption of an empirical research, if not properly used, may compromise the generalisability of the research solution (which relates to external validity). However, DECYSE was devised based on an investigation of common CS issues affecting different communities.

The empirical research was used with the purpose to triangulate such literature and involved participants with different backgrounds (cf. Figure 6.1). Data collected on which to test DECYSE similarly represented different sectors of society (cf. Figure 6.7a) which shape the CS community. For example, data used as the initial set of recommended competencies were drawn from recurring subjects in different CS frameworks and academic courses (cf. Table 2.1), while initial data for recommended certifications was drawn from an existing list (cf. Appendix E). The recommended competencies echoed in the majority of real-life profiles (i.e. the green bars in Figure 6.8) in contrast to the recommended certifications (cf. green bars in Figure 6.9). Such a comparison illustrates the depth in the investigation of CS expertise demands and the flexibility of DECYSE in handling new capabilities (i.e. the red bars in both figures). Hence, special care has been taken in order to ensure that DECYSE can be deployed in CS communities with different maturity levels and information needs on expertise discovery. Moreover, the method is purposely unbound to application and technology layers, social media platforms or existing CS frameworks, in order to favour its generalisability.

7.1.3 Validating DECYSE from its Underpinning Approaches

It is not clear if the current CS initiatives (cf. Section 2.1) are underpinned by academic research. Notwithstanding this and as previously discussed, those initiatives do not cover all the necessary aspects to perform expertise discovery, which are delivered by DECYSE. Inspired by Talent Management and Knowledge Management theories and in compliance with well-established methods and techniques, DECYSE offers a comprehensive information analytics solution to the CS community taking into account the scientific rigour for its development. This section highlights aspects of the theories, methods and techniques that underpin DECYSE and support internal validity.

Prior to the evaluation on the suitability and application of the adopted literature approaches, such an investigation was triangulated by an empirical research with primary data (cf. Section 6.1), due to the complexity of the CS field. The investigation also identified additional requirements to address the research questions. For example, the interviews along with open-ended answers in questionnaires supported specification of the criteria types for CS expertise discovery, since this is not clear in the literature. The set of variables created to determine the expertise shortage and gap (i.e. the *#capability indicator*) similarly relate to the criteria specified in a *project profile*, which make them adequate for the solution context. Other variables (e.g. *WR* and *PR*) are adaptations of the literature for the DECYSE method.

The pluralistic view of Talent Management and the co-evolution of CS actors are ensured because their interdependencies and information needs were mapped since requirements analysis (e.g. Figure 5.4). For example, the DECYSE method suggests relevant capabilities for individuals (in *select candidate*), contractors (in *flag successful projects* and *select candidate*), organisations (in *feedback for training plans*) and for the CS community as a whole via the *policy maker* (based on *update analytics*). Co-evolution is delivered according to an expertise discovery lifecycle (cf. Figure 4.5), which is realised through the affordances of the DECYSE ontological model. The DECYSE articulation (cf. Section 5.1) determines the services and information resources that each stakeholder should provide to or exchange with the CS community.

The complexity of the research problem and the identified viewpoints required the development of a conceptual model capable of testing and providing information analytics (as introduced in Section 3.1). Hence, DECYSE relies on a combination of selected KR and analytic techniques. For example, an application of Big Data Analytics allows the CS community members to systematically learn from the rich knowledge environment by measuring the knowledge gaps (cf. Figure 6.15) and shortages (cf. Figure 6.16), which addresses in part the research problem.

The use of Archimate in combination with OS supported the conceptualisation of national CS as an organisation concerned with managing its pool of talents. This combination builds the foundations to support the co-design of business and IT when deploying DECYSE in a real-life setting. Moreover, such an approach enabled a holistic view on the information flow and motivational elements, facilitating identification of the concepts needed for developing the DECYSE method. The use of profiling techniques under the light of co-creation improves trustworthiness, which is crucial to the domain context. This approach can be seen in the *Participant profile* and in the *Course profile#Course assessment*.

In terms of the information sources, DECYSE profiles both explicitly and implicitly provided data (cf. Section 3.2.3). The latter can be seen when new project profile capabilities are suggested based on patterns of similarity discovered in successful projects (cf. *define project* workflow). The solution, however, prioritises internal data sources rather than external data from social media to ensure reliability, agreement among concepts and because very little relevant information is made available via APIs.

In line with the benchmarking validation pattern, an investigation in DW, Big Data Analytics and KRI enabled the creation of a flexible capability metric to determine the expertise gap (cf. Figure 6.15) and shortage (cf. Figure 6.16) in the CS ecosystem. The merit of *CpM* relies in its simplicity as a tuple; scalability due to its lightness and because older instances can be archived according to its timestamp; flexibility to support a variety of capability indicators or even to measure relevance (in terms of supply and demand) in other settings; and suitability to the research problem, since it discovers the expertise over time.

The choice to use and adapt well-grounded analytic techniques (e.g. Analytic Hierarchy Process, Data Envelopment Analysis, Exemplar-Based Reasoning and Reputation System) within affordances enabled to produce sound experiment results. For example, the use of Data Envelopment Analysis for selecting candidates enables a scalable talent pool, since the number of participants (whether sixty-six or a thousand, as presented in Table 7.1) does not have a significant impact on the number of efficient candidates. The DECYSE method supports a knowledge management system capable of capturing and processing explicit knowledge. Moreover, the method enables computing claimed tacit knowledge (i.e. the *competence*) via an adapted Reputation System that captures objective ratings (i.e. the *competence level*).

7.1.4 Validating DECYSE from the Experts' Feedback

A major conceptual remark when validating decision support systems is that there is no feasible way to regard such a system as a definitive solution. Thus, validation under an interpretive approach seeks to verify whether the system behaves as expected and becomes an appropriate solution according to user's viewpoints (Finlay, 1989). In addition to the successful experiment results, DECYSE was content validated through appraisals captured from eight experts working in an agency involved in the CS coordination effort, which includes expertise management. Those professionals are potential users of the model. The appraisals were captured after presenting the model using a questionnaire that allowed open-ended answers. Questions were grouped in terms of usability, cost-effectiveness and acceptance, as shown in Figure 7.1.

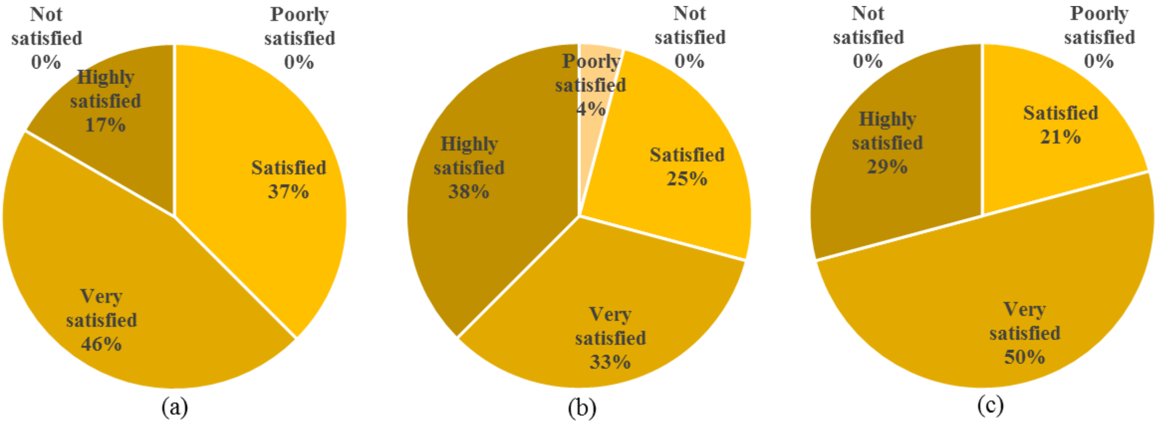


Figure 7.1: (a) Usability, (b) cost-effectiveness and (c) acceptance of expertise discovery approach

In contrast with the positive appraisal in the questionnaire, some criticisms have been identified. In terms of usability, an expert argued that developing ontologies for the capabilities would improve the quality of the model; however, this would demand a great effort. In terms of cost-effectiveness, an expert warned that the method might require high costs in terms of development and maintenance. In addition, experts argued that psychological features (e.g. ability for group work and leadership) should also be considered. Such a request is feasible in our approach as additional competencies. However, this could incur in sensitivity and privacy issues, since those features relate to moral and character and not to professional knowledge, skill or abilities. In terms of acceptance, an expert argued that the DECYSE method enables evaluating capabilities and managing people, but not talent as a select group of people. This work, however, adopts an inclusive approach, focusing on talent as proper allocation rather than exclusively individuals with natural ability or high capability in specific areas.

7.2 The Research Contributions to Theory and Practice

The novel approach and combination of theories and techniques presented in this research aimed to solve the problem of discovering expertise in order to promote national CS. In addition, several claims drawn from the literature are addressed, contributing for the evolution of the bodies of knowledge involved and to industry. Some of these contributions relate to the claims categorised under the aspects of theory and practice.

7.2.1 Research Contribution to Talent Management and Knowledge Management Theories

This work improved the understanding on challenges for the multidisciplinary CS field, through an overview of its current initiatives. The research has also summarised conceptual aspects, approaches and recent developments in theories and techniques in the fields of Talent Management and Knowledge Management. The following paragraphs present four contributions of DECYSE to literature.

Robust process for expertise discovery

This research work contributed to methodology in Knowledge Management (KM) by devising a robust process to perform expertise discovery in the CS domain. Findings in the literature determined the adaptation of a set of suitable principles to guide expertise discovery in the CS domain. These expertise discovery principles were integrated and analytically described using some robust processes based on solid techniques. Since no approach alone was capable to solve the research problem, some of these techniques were adapted and combined in a unique way. For example, a Reputation System was adapted as a solution to measure experience and as a recommender system for appraising knowledge exchange. The types of criteria for expertise discovery similarly had to be defined, while new variables (e.g. *WR* and *PR*) and metrics (e.g. *CpM* and *capability indicators*) had to be created for the methodology. The actors playing a role in expertise discovery for CS and their information needs were identified. In addition, some findings in the fields of CS, Talent Management and KM were obtained in terms of expertise discovery for CS and presented as viewpoints in Chapter 2.

In addition, the thesis delivers a literature review on the concepts related to CS (cf. Appendix B), to understand its fuzzy nature and from which core elements to support the research solution were obtained. Based on a thorough investigation into CS and Talent Management, DECYSE delivers seamless expertise discovery processes for the CS community, which promotes ongoing improvement and knowledge sharing. Such an achievement contrasts with current isolated and episodic initiatives that do not address the core of the research problem.

Delivering a pluralistic approach for Talent Management

DECYSE addresses calls in the literature and contributes towards a pluralistic approach of Talent Management (TM) theory by delivering a differentiated architecture for performing expertise discovery in the CS ecosystem. In contrast to the traditional focus of managing talents by a single agency, DECYSE adopts a business ecosystem perspective to model a complex system comprised of different organisations (e.g. cyber security agency and course providers). Hence, this work extended research on TM beyond private organisations, exploring the issues on organisational partnerships and its impact for society. The perspective considers the innovation that actors in the CS community can bring to the common goal of expertise discovery. The challenges found in deploying the CS community as an ecosystem are addressed through a systematic requirements elicitation approach (cf. Figure 5.4) and via the semantic units in the DECYSE ontological model. Hence, DECYSE provides a holistic view of the CS community, conceptualising the roles and the actions in the context of expertise discovery, structuring required information and supporting co-evolution of the actors for the benefit of the ecosystem as a whole. The method is both generic, because it suits the needs of CS communities in different stages of evolution (cf. Section 7.1.2), and repeatable, since it has been iteratively and successfully experimented using different datasets (cf. Section 7.1.1).

Measuring the expertise gap and shortage

This thesis contributes to the CS field and to Knowledge Management theory by delivering a mechanism to measure the expertise gap and shortage. The mechanism is devised to extend the concept of "competence warehouse" from simply mapping organisational knowledge to sort such knowledge in terms of relevance. Because of profiling the available expertise and CS projects, a set of indicators was conceived in order to measure the overall expertise gap and shortage within the CS community. Such information benefits directly the policy maker and indirectly other roles (e.g. participants and contractors) by raising general awareness on relevant CS expertise. Moreover, these metrics enable further implementation of predictive and prescriptive analytics in order to structure a proactive stance towards the CS expertise shortage.

Improving the CS ecosystem over time

The thesis presents a contribution to the theories of organisational modelling with regard to change management. Such a contribution is twofold and concerns the framework for developing the DECYSE method, as well as the DECYSE ontological model itself, which embraces an expertise discovery lifecycle. This work encompasses the iterative steps (i.e. the meta-processes) to devise DECYSE (cf. Figure 5.1), which combined requirements engineering with Archimate. The approach is capable of determining a service inventory and categorise services according to their relevance for expertise discovery. In line with Design Science Research's iterative development, those services that do not fall into the current scope can be used to extend the DECYSE method in future iterations or to solve other specific problems in the CS ecosystem. In addition, the complementary use of distinct methods (i.e. Archimate and semiotic-based ontology) enables revisiting the meta-processes in order to produce new requirements and services to update or customise the DECYSE method.

There is a semantic integration of feedbacks, which enables self-managing and self-evolving advice within the CS community. For example, the affordances *update participant experience*, *flag successful project* and *evaluate peers* automatically change the status of profiled information, which are the inputs used by the antecedents of those affordances. In line with the ecosystem approach, these feedbacks are delivered in multiple levels, e.g. for project candidates and contracting organisations. Feedbacks for the CS community as a whole are delivered via up-to-date *recommended capabilities*, which promote joint stewardship according to each member's responsibilities. Feedback also enables co-creation of profiles, which improve their trustworthiness, a valuable asset in the CS community. In addition, DECYSE provides a feedback of metrics improving efficiency on decision-making. These features realise the cyclical nature of DECYSE, which, along with its meta-processes (cf. Figure 4.5), makes it possible to keep the expertise and its discovery processes up-to-date, despite changes in the dynamic CS environment.

7.2.2 Contributions of DECYSE for Industry

DECYSE adapts and integrates a group of techniques capable of articulating the information needs from CS actors in the context of expertise discovery. The adoption of a pluralistic and ecosystemic approach for DECYSE generated innovation, which eventually benefited those actors in the CS community engaged with expertise discovery. There are four contributions of DECYSE to industry reaching different stakeholders.

1. Service-consuming organisations, i.e. contractors requesting expertise, can improve quality of search results through a flexible set of preferences. Structuring the criteria for searching expertise is facilitated by reusing features in similar projects and with a recommended and up-to-date set of capabilities. In addition, the types of criteria that were identified (cf. Table 5.4) loosely cover a wide range of expertise requirements for CS projects.

2. The contractor can extend the expertise pool by sharing talent resources who are available to work in projects even if the talent is employed in another organisation. DECYSE builds on the cooperation effort in the CS community and on the need to balance the resource-consuming field of Talent Management in order to structure and optimise such expertise-sharing network.

3. Opportunities for engagement in the CS effort are increased for the citizen able to participate in the community. Such opportunities can be represented through feedback providing self-awareness, job offers, accredited courses and knowledge exchange with peers.

4. Robust knowledge processes in the CS community are improved because DECYSE delivers structured, flexible, clear and integrated workflows. The benefits of this contribution reaches all actors involved, particularly those actors playing the role of contractors in need of expertise, and those performing a coordination role, which are in charge of steering the community. These benefits are supported with the appraisal in Figure 7.1.

7.3 Justification on the Research Design

Design Science Research was the approach adopted to conduct this work in order to design an artefact that could be improved over time for the dynamic CS community. Such an improvement can be the result of refinement on the method during further iterations or due to emerging needs from the community. Hence, the adoption of certain patterns under Design Science Research were appropriate to devise a sound solution, to ensure construct validity and to promote the research contributions. The justification and actions taken for designing the research encompassed by those patterns are discussed in Table 7.2. The actions also identified some challenges and shortcomings when applying the respective patterns.

Table 7.2: Selected patterns and respective actions for designing DECYSE

Used Patterns	Actions that were taken
Brain Storming and Framework Development	Brain storming was widely used as a framework to investigate and organise literature on the fields of CS, Talent Management and Knowledge Management. When the concepts and knowledge gaps were identified and structured, selected topics with regard to the research problem were selected to support development of DECYSE. The brain storming technique was also used to support the analysis of stakeholders and requirements for the CS community.
Wild Combinations; Solution-Scope Mismatch and Combining Partial Solutions	Some KR and analytical techniques were either adapted in an unconventional manner or expanded in order to suit the solution context. For example, the DECYSE ontological model includes semantically integrated feedback loops (e.g. via *update analytics* and *feedback for training plans* affordances) in order to support ongoing improvement. Another example is how a Reputation System for reciprocal transactions is adapted as a mechanism to benchmark experience. Moreover, the combination of many different information analytics techniques resulted in a unique solution catering the adopted principles in the expertise discovery lifecycle (cf. Figure 4.5). Such a combination entails seamless information flow on knowledge processes, which benefits all actors in the CS community. Therefore, DECYSE eventually becomes a robust and unique arrangement of the best-suited and well-grounded techniques that actually serves the purpose of managing CS expertise.
Research Domain Identification; Problem Formulation; Being Visionary and Bridging Research Communities	The problem addressed is considered relevant for the CS community (cf. results presented in Section 6.1) and delivers contributions for other fields of knowledge, which means that this research targets different audiences. For example, DECYSE promotes extension of talent pools for organisations while increases engagement opportunities for citizens. Since the CS context does not provide enough scientific resources and literature by itself, the fields of Talent Management and Knowledge Management had to be investigated to devise a sound solution based on literature. Bridging those fields to address the CS expertise discovery problem has provided insights and increased the research significance due to its interdisciplinary nature. Most insights regarding the research fields were presented through viewpoints (e.g. Figure 2.7 and Figure 2.13). Although much is already being done in terms of CS initiatives, DECYSE improves the existing solutions with a single extendable model. Moreover, it seems that DECYSE is applicable in similar communities facing problems with expertise discovery. However, some challenges have arisen due to the interdisciplinary nature of this research. Firstly, publishing became difficult since reviews from most journals did not favour reconciliation with other research areas. Secondly, most CS experts were not familiar with the analytic techniques used to support DECYSE. Therefore, explanation of the method required a preliminary approach on the methods and techniques involved.
Hierarchical Decomposition; Complex Systems Analysis; Hierarchical Design and Integrating Techniques	Due to its complexity, the CS community was broken down into manageable parts in order to analyse static (e.g. stakeholders and roles) and behavioural concepts (e.g. services and processes) with regard to the research problem. Thus, selected KR methods (e.g. Archimate and OS) played a key role to provide a graphic interface and a holistic view when discussing with stakeholders and when performing conceptualisation of the CS community. In our approach, the required concepts were mapped using Archimate and those relevant concepts were iteratively integrated using semantic analysis. The use of OS with profiles and norms allowed describing the different types of selected algorithms (e.g. Analytic Hierarchy Process, Data Envelopment Analysis and reputation) that govern the behaviour of the agents in the CS community. The integration of different algorithms, for example, improved flexibility in contractor’s decision-making by allowing a combination of qualitative and quantitative criteria for CS projects. DECYSE was conceived as a hierarchical design, where the processes are functionally categorised in compliance with expertise

Used Patterns	Actions that were taken
	discovery principles for the CS community (cf. Section 5.2). Such a functional approach enables the clustering of a particular category of processes for future improvement without significant impact on processes from other categories. Moreover, motivational concepts such as drivers and goals (cf. Figure 5.6) may support future improvement on the DECYSE method.
Emerging tasks	Only the tasks that contributed to solve the research problem were properly addressed in the DECYSE method. However, other related services (e.g. "exchanging knowledge with peers" and "evaluating a course" in Table 5.1 or Appendix D) were identified and left out of the DECYSE method. On the one hand, time constraints and simplicity drove the final scope of the problem to be addressed. On the other hand, those services and other emerging tasks can extend the DECYSE method in further iterations as future work in compliance with the Design Science Research approach. Either way, it is recommended to enable a communication channel and transparency of capability indicators to all registered participants in order to ensure constant innovation.

7.4 Justification on the Application of DECYSE

The research problem is complex, involves multiple research fields and cannot be mathematically proven as valid due to its strong social context. In contrast, the validation of the solution using solely real-life values is not feasible (i.e. time consuming) because some criteria in DECYSE (e.g. reputation) require interactions between users to generate those values. Thus, an empirical research was conducted in order to identify actual information needs and to test the method in a real-life scenario. However, special care has been taken in order to ensure generalisability of the solution.

The investigation conducted under the Design Science Research approach enabled the triangulation of findings in the literature and investigation of the patterns to devise DECYSE. Those patterns were obtained from CS documents used in different countries, common practices in the field of Talent Management and opinions of experts. The criteria used to test DECYSE reflect the urges and actual features of potential users, which were drawn from the questionnaires and interviews with experts. An initial suite of test data had to be developed to test DECYSE using a range of possible criteria identified as requirements. The range of values used for *recommended capabilities* were based on actual relevant criteria for the CS community drawn from the literature. Therefore, with regard to all that has been discussed in this section, generalisability of the adopted approach is assured (cf. Section 7.1.2).

There is no common agreement in the sample data size for the experiment as long as it enables exercising the method, representative testing of real-life situations and consideration of the of the solution goals (Vaishnavi and Jr, 2007). During the development of DECYSE, some test datasets were developed to test parts of the model and results were published by Fontenele and Sun (2016). The method's fundamental approaches and variables were appropriately selected, adapted or created to fit the research solution (cf. Section 7.1.3). The adoption of experimentation on a cleansed dataset seemed more appropriate to support validation of DECYSE. Therefore, after thorough experiments using the test dataset, data profiled from a Brazilian CS community was used with DECYSE, delivering satisfactory results as presented in Section 7.1.1. The empirical research ensured that different sectors of society were represented for the final experiment (cf. Figure 6.7a). Moreover, DECYSE was positively appraised in terms of usability, cost-effectiveness and acceptance (cf. Figure 7.1). Hence, the data analysis regarding the experimentation of DECYSE supports the validity of the solution to the research problem.

7.5 Limitations of DECYSE

There were limitations in this research, as follows:

1. The work reputation algorithm does not consider the duration of the project (cf. Section 6.1), which may favour participation in CS projects of shorter duration. For example, it might be worthwhile for a participant receiving *candidate feedback* from two projects to accept the one with a closer *P.Finish date*, in order to be available as soon as possible for another evaluation. In contrast, longer projects may offer more profit, which also affects the candidate's choice.

2. In this work, the perspective of capability types was adopted for the sake of simplicity in experiments, rather than exploring conceptual relations with higher granularity as suggested in Table 5.5. Although participants can clearly claim certification and academic area, other types of criteria such as competencies still seem to require further description in terms of tasks for proper evaluation.

3. There are still some concerns over the creation of new types of criteria besides those presented in Table 5.4. The adopted approach within the context of expertise discovery (cf. Section 5.2.4) requires caution when including new types of qualitative criteria, otherwise candidates may be prematurely excluded from the ranking process. As a design decision, the types of criteria used for *optimise choice* workflow should be somehow considered during *select candidate*. For example, competence, academic area and certification are respectively measured in quantitative values through experience, academic degree and number of requested certifications.

4. The research identified services for the CS community, which, for the sake of

relevance and time constraints, did not fall under the current scope of DECYSE (i.e. those associated to general roles other than "actors" in Table 5.1). Since the adopted iterative approach allows further improvement on the model, these services also should be considered in future work. Moreover, the creation of new affordances should concentrate on adding functionalities, rather than altering the existing information flow. Such an assertion aims to avoid inserting bias in the profiled information, which relies on constant feedback and co-creation.

5. Previous experience and evaluations from contractors that are not registered cannot be mapped in terms of *competence level*. On the one hand, all the work prior to *register participant* can only be subjectively described or pointed to other social media profiles via *Participant profile.Complementary Info*. On the other hand, *competence level* and *peer reputation* are not subject to bias by importing values from distinct social media, where the reputation algorithms are not clear. Moreover, this becomes a fresh start opportunity for newcomers.

7.6 Summary

This chapter conducted an evaluation on the research project by addressing the strengths and limitations of DECYSE. The empirical research provided valid results with regard to acceptable outputs from experiments and through favourable appraisals from possible users. The expected results were produced by an experiment using a real life dataset, which was further compared to previous experiments with a test dataset in order to avoid bias. Results also supported the validation of DECYSE in terms of the method's flexibility, generalisability and due to a robust combination of sound underpinning techniques and variables within the model's affordances. Finally, a board of experts appraised DECYSE with positive reviews and some additional valuable remarks. Hence, the results validate the experiment and the DECYSE method is regarded as an acceptable solution for its possible users. The design and development of DECYSE using Design Science Research patterns generated meaningful insights, which were transcribed for practitioners who intend to make use of or even extend this project. The method itself and its iterative development have brought contributions to theories and techniques in the fields of CS, Talent Management and Knowledge Management, as well as for the actors involved in expertise discovery for CS. The research work achieved the objectives and some of the identified limitations are addressed as opportunities for the future work.

Chapter 8

Conclusion and Future Work

This chapter presents some concluding remarks and summarises the research work for the thesis according to the research objectives. In addition, possible future research directions for this project are proposed.

8.1 Concluding Remarks

Discovering the right expertise in an effective fashion is paramount for the CS community. Nevertheless, its complex and dynamic knowledge environment has not been systematically addressed. The DECYSE method covers the information flow within the entire expertise lifecycle by using sound theoretical and methodological foundations. In addition to the research solution and aligned with the Design Science Research paradigm, special attention was dedicated to the meta-processes for devising the method itself. Hence, this research project delivers a solution capable of ongoing improvement on the expertise discovery and on the processes that govern the solution itself. The merit of DECYSE in addressing the research problems relies on how the method contributes to theory, methodology and practice (cf. detailed in Section 7.2) by answering the following research questions:

How can a methodology be developed and maintained to aid a robust discovery of expertise within a collaborative CS environment, where requirements for expertise are dynamic and evolving?

What criteria and metrics can be formulated in performing the expertise discovery workflow within the complex CS environment?

How can an expertise shortage be methodologically described in order to target skill development and to satisfy the fast-changing CS environment?

A methodology for aiding a robust discovery of expertise concerns more than just matching experts with the right capabilities for CS projects. Expertise discovery also includes keeping relevant expertise up-to-date in an environment that requires collaboration. Hence, DECYSE was conceived to provide seamless expertise discovery processes for the CS community by identifying their stakeholder's needs and integrating their information flow. These processes are robust and they encompass well-grounded analytic techniques, some of which were adapted to suit the research solution. Robustness is enabled with the use of norms which control behaviours and support automation. DECYSE aids the definition of project requirements by suggesting additional criteria. For example, contractors are capable of enhancing project requirements by reusing features in similar successful projects (via *#suggested criterion*) and having access to up-to-date recommended capabilities (based on metrics created for DECYSE). The initial set of suggested values for each type of criterion (i.e. *recommended capabilities*) was partly based on a set of recurring subjects arbitrated as competencies for testing the method. The initial set of recommended certifications and academic areas were similarly based on existing government documents. Although such a set of values was adopted to illustrate the power of DECYSE and will evolve over time because of expertise analysis, they require further updates and definitions before implementing DECYSE in a real setting. In addition, competencies should also be instantiated in terms of KSA and tasks to ensure conceptual understanding among CS actors. Notwithstanding, the experiment provided expected results.

Since collaboration is critical for CS, DECYSE embeds a systematic approach for identifying the stakeholder's needs under each expertise discovery principle, which encourages participation and promotes the overall expertise discovery lifecycle. The expertise discovery principles were structured based on a combination of talent practices and business ecosystem challenges. Participation in DECYSE includes constant evaluation of experts from different perspectives, which increases the reliability of profiles over time through co-creation. For example, using reputation as a measure of cumulative performance (i.e. experience) assures that the most experienced participants holding required competencies are selected as candidates for projects. Interactions for knowledge exchange are also rated (to determine a *peer reputation*) and encourage a proactive stance with other participants, which increases their chances to be selected for projects. While DECYSE embodies the challenges of business ecosystems in order to nurture innovation for expertise discovery, the method contributes to a pluralistic approach for Talent Management theory. Such an approach enabled contributions to industry as well. For example, mapping the expertise in the CS community enables extending a single agencies' expertise pool, while increasing visibility and consequent engagement opportunities for citizens. The DECYSE ontological model, therefore, provides a holistic view of a CS ecosystem, which enables a semantic integration of feedback. This feedback enhances awareness and dynamically improves the knowledge processes and the expertise within the CS community with reliable and structured information. Based on these features, DECYSE structures, measures and promotes a seamless information flow to transform the current CS environment, which requires coordination and collaboration, into a functional CS ecosystem. Moreover, DECYSE successfully combined Archimate with ontology modelling to identify requirements and conceptualise the research scope. Such combination is illustrated in the research framework for developing the DECYSE method (Figure 5.1). Hence, in addition to deliver the DECYSE ontological model providing feedback loops, this research deliberately described its development meta-processes in a conceptual model so that the solution itself can be revisited for future improvement.

There are two aspects concerning the expertise discovery workflow within DECYSE. The first aspect refers to criteria established as requirements to work in CS projects. DECYSE structures the CS expertise into three types of capabilities (i.e. competence, academic area and certification) for articulating qualitative criteria and metrics for expertise discovery. The DECYSE method processes four additional quantitative criteria as a means to select the most suitable candidates for a given CS project. The qualitative criteria is weighted (cf. Appendix F) to comply with the expertise requirements. The types of criteria (cf. Table 5.4) were drawn from literature and triangulated via the questionnaire and the interview, ensuring their relevance across different CS communities. Hence, the method enhances flexibility to process different types of criteria and quality in expertise search, which benefits contracting organisations. With regard to the second aspect, metrics for expertise discovery computes the overall expertise shortage and gap in the CS community. These metrics (i.e. *CpM*) are based on participant and project profiles and capture the supply and demand for each type of capability in a given time. *CpM* contributes to indicators that measure the expertise gap and shortage for each capability through a demand and supply ratio. These indicators depict a holistic, analytic and dynamic view of the expertise, which benefits members of the CS community.

The expertise shortage is identified and measured when capabilities are prioritised according to those ratios on a dashboard. An expertise gap is determined whenever there is a requested capability in the project pool that does not match with the participant profiles. Through the identification of expertise shortages and gaps, these capability indicators support keeping the *recommended capabilities* up-to-date to target skill development.

DECYSE is capable of guaranteeing expertise fitness over time, even in the evolving CS knowledge environment, by providing learning opportunities and structuring the model's meta-processes. The semantic units related to provision of feedback principle encompass learning opportunities and raising awareness on expertise for all actors in the CS community. For example, while participants are guided by advice on missing capabilities in *select candidate* workflow, contracting organisations can improve their project requirements over time with the *feedback for training plans* or *flag successful project* workflows. The latter workflow enables *define project* to offer advice for future contractors to improve their project requirements based on experiences. Another example to illustrate how DECYSE continuously satisfies the changes within the CS environment is how *recommended capabilities* are updated according to capability indicators via the *policy maker*. The expertise analysis dashboard delivers awareness on relevant capabilities for the whole CS community via the *update analytics* affordance, so actors can co-evolve according to the expertise discovery lifecycle. Such results depict the current expertise situation within the CS community, which provide the means for its players to take timely action according to their roles in order to improve expertise discovery over time.

This thesis began by outlining the background and motivation, along with the research problem, questions and the aim and objectives in Chapter 1. The following chapters performed discussions in line with the research objectives, leading to their achievement. After delivering an overview of the research scope of CS expertise discovery, a thorough discussion upon the fields of CS, Talent Management (TM), Knowledge Management (KM) and organisational modelling was conducted in Chapter 2. Some critiques with regard to the research problem were drawn for each of those fields, which enabled to identify the viewpoints to be addressed by DECYSE. CS was discussed according to a theoretical (via academic literature) and pragmatic stance (via government documents and PPP) to explain its current challenges for expertise discovery. Then, TM approaches, applications and practices were introduced as the means to enable suitable practices and expertise discovery principles guiding the CS community. The investigation provided viewpoints encompassing best practices and theoretical support to the adoption of a strategic and collaborative stance focused on the proper allocation of experts. Afterwards, the development of KM processes to support discovery of expertise in general was investigated through organisational modelling theories. The viewpoints obtained from those theories required approaches and analytical techniques to represent, source and process information required to perform expertise discovery. This led to the discussion in

Chapter 3 about methods and techniques for knowledge representation (e.g. ontology and enterprise architecture) and Big Data Analytics that could be used to support the aim and objectives of the research. Suitable methods and techniques focused on representation, measurement (e.g. DW, reputation and KRI) and analysis (e.g. decision support with Analytic Hierarchy Process, Data Envelopment Analysis and Exemplar-Based Reasoning) were selected with regard to their strengths and limitations.

Following discussions on the literature, Chapter 4 reviewed aspects of the research methodology in order to support the adoption of suitable research paradigms, methods and data collection techniques for this work. The investigation led to the adoption of mixed methods (i.e. empirical research and experiment) encompassed by the Radical Subjectivist Paradigm and Design Science Research. Then, a brief discussion on the adopted approach to develop DECYSE was performed comprising the data sources and collection, the development steps, validation and evaluation. Despite the complexity of the CS community, the scope of expertise discovery was reduced to a set of relevant functions, which in this thesis is realised through semantically connected processes. Those analytic techniques were properly adapted or combined in these processes in order to articulate profiled information and compute criteria of either a qualitative or quantitative nature. Aligned with the iterative essence of the adopted paradigms and the changing CS domain, particular emphasis was placed on questioning and revisiting the research solution over time. Hence, not only does the solution concentrate on an expertise discovery lifecycle, but the steps for improving the solution itself are also made available. Chapter 5 concentrated on introducing an overview of DECYSE, presenting aspects of its development, the relation among the DECYSE affordances with the expertise discovery lifecycle and a thorough discussion on the underlying analytical processes. In contrast with traditional single-viewed approaches focused on managing talents as individuals, DECYSE considers essential information needs and services that each actor can provide under the expertise discovery principles. Such an organic approach promotes interactions within the CS community, which eventually speeds up the expertise discovery processes and enables a CS ecosystem. DECYSE was trialled in Chapter 6, providing successful results. Prior to experimentation, some of the data collected to articulate the research problem and to devise the method was presented. Finally, Chapter 7 presented a critical evaluation on DECYSE. The solution was successfully validated according to the testing results, generalisability, suitability of underpinning techniques and appraisals from a board of experts. The evaluation also included DECYSE's contributions, details on how the method addresses the research questions, justification of soundness in terms of the adopted research methodology and a discussion of its shortcomings.

The human factor and change are crucial to CS. Thus, DECYSE was teleologically conceived to support the complex CS community as a living and self-nurturing ecosystem that should promote co-evolution of all of its members. The approach for eliciting mutual responsibilities under the expertise discovery principles (cf. Figure 5.4) illustrates how each member can contribute to innovation. Solutions for implementing ecosystem challenges within the CS expertise discovery context are embedded in the DECYSE method. The health of such ecosystem seems to rely on constant awareness of its participants' information needs.

8.2 Future Work

During the development of this project, several challenges emerged, some of which did not fall into the scope of this thesis and others were considered as limitations. Therefore, these challenges are selected as opportunities for extending the DECYSE artefact in future iterations. The recommendations for the future work are as follows.

Analysing how the duration of projects affect experience

It is still not clear how the time spent working on a single project can improve experience in a given competence. Although this work disregarded the duration of projects for the sake of simplicity, future research may study its relevance to extend the working reputation algorithm (cf. N6.3 and N6.4 in Figure 5.20) in order to balance evaluations between shorter and longer projects.

Instantiating an initial set of recommended capabilities

In the future, proper instantiation on the capability types should be one of the priorities for proper deployment of DECYSE in order to deliver ontological commitment among actors in the CS community. Such a call also resonated in interviews, as discussed in Chapter 5. Although this research work provided a sample set of recommended capabilities based on literature and initiatives, each nation should embrace their own priorities despite the existing CS framework. Hence, the initial set of recommended capabilities should reflect the nature of the main CS threats faced by actors in a given nation. DECYSE already provides the mechanism (i.e. feedback via the *update analytics* workflow) to keep those recommendations up-to-date. The capability instantiation encompasses defining the underpinning conceptual instances of those capabilities (cf. Table 5.5). For example, a competence should be defined in terms of its KSA and tasks for proper evaluation, while courses and certifications should be defined in terms of competencies. Such ontological dependencies may facilitate a more elaborate capability recommendation mechanism in future work.

Evaluating remaining actors in the CS ecosystem

The adoption of the ecosystem approach entails that CS community members have mutual responsibilities under the expertise discovery principles. Notwithstanding, some elicited responsibilities were not deemed essential to the research solution. For example, reporting a claim for the CS agency and course evaluation fall under the evaluation of expertise principle. However, these were considered as services placed beyond this research scope, since they were not associated to actor roles (cf. sixteenth and twentieth rows in Table 5.1). Future work can extend the DECYSE method to address evaluation of courses by participants in order to ensure quality in accredited courses over time. The property of *Course profile.Course assessment* is already defined in Figure 5.10 for this purpose. Moreover, Appendix D delivers valuable cues for further improvements of the DECYSE method.

Enhancing expertise analysis with predictive and prescriptive analytics

Another research direction is to extend the expertise analysis on the "capability warehouse" with predictive and prescriptive analytics. DECYSE already uses descriptive analytics, via the *#capability indicators*, to support determining *#recommended capability*. However, the prediction of trends for relevant capabilities can benefit with further development, for example, by using statistics on an indicator, which provides a history of demand and supply ratio (e.g. Figure 6.17). The use of prescriptive analytics can similarly assist in planning the course of actions to reduce the expertise shortage, which is already measured by DECYSE. Such a task can be performed with the creation of KPI on those expertise shortages or gaps.

Improving provision of feedback through personalised information provision

An additional research direction is to improve the provision of feedback principle through personalised information provision. Some of the literature (e.g. Ousmanou, 2007; Sun et al., 2010; Sun and Mushi, 2010) discuss such research topic using OS. An initial step should be mapping the *Contact information* and *Capability* from a *Participant profile* in

DECYSE with *Personal information* and *Portfolio* from a *User Profile* in (Ousmanou, 2007). Indeed, the *Capability* feature is thoroughly explored in this thesis, while the *Portfolio* feature is beyond the scope of that research. Therefore, the DECYSE method can be extended to incorporate such functionalities (e.g. via a "providing lifelong training for participants" service, presented in Table 5.1) in order to deliver tailored advice for expertise development according to participants' learning preferences. On the other hand, such a research direction can also extend the *User profile* presented in (Ousmanou, 2007).

Securing integration with external data sources to ensure information authentication

Despite the issues on privacy, copyright and limited content discussed in this work, APIs with social media platforms may improve trust in certain participant profile features by performing information retrieval on social profiles. DECYSE already considers authentication features as a condition for course providers to be accredited (i.e. by enabling *Course profile.Authentication*) and to ensure trust in information about experts (i.e. via *Participant profile.Capability.Certificate.URL*). Hence, future work can investigate the implementation of authentication mechanisms via APIs on relevant external social media platforms with DECYSE.

Extending DECYSE in other CS communities or different business settings

Another potential research direction is either extending DECYSE in other CS communities to improve the maturity of the model or apply the methodology in other similar multi-organisational business settings, in order to improve its versatility.

CS expertise discovery is a challenging and interdisciplinary subject with numerous possibilities of application. The DECYSE methodology is a comprehensive and flexible approach thoroughly devised for such subject. Hence, the opportunities for future work, whether to enhance DECYSE or to delve into other approaches, are not constrained by the aforementioned topics. Revisiting the requirements articulation (i.e. steps 1 to 3 in Figure 4.3) can eventually promote new insights and research directions to improve national CS and the well-being of society.

References

Aamodt, A., Plaza, E., 1994. Case-based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communications 7, 39–59.

Abdulrahman, R., 2012. Multi agent system for web database processing, on data extraction from online social networks (Ph.D.). University of Bradford.

ACM, 2015. The 2012 ACM Computing Classification System - Association for Computing Machinery [WWW Document]. Association for Computing Machinery. URL http://www.acm.org/about/class/class/2012 (accessed 11.18.15).

Adobor, H., 2004. Selecting management talent for joint ventures: A suggested framework. Human Resource Management Review 14, 161–178. doi:10.1016/j.hrmr.2004.05.001

Alavi, M., Leidner, D.E., 2001. Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. MIS Quarterly 25, 107–136. doi:10.2307/3250961

Altınöz, M., Çakıroğlu, D., Çöp, S., 2013. Effects of Talent Management on Organizational Trust: A Field Study. Procedia - Social and Behavioral Sciences 99, 843–851. doi:10.1016/j.sbspro.2013.10.556

Ananda Kumar, V., Pandey, K.K., Punia, D.K., 2014. Cyber security threats in the power sector: Need for a domain specific regulatory framework in India. Energy Policy 126–133. doi:10.1016/j.enpol.2013.10.025

Australian Government, 2009. Attorney-General’s Department: Cyber security [WWW Document]. URL http://www.ag.gov.au/RightsAndProtections/CyberSecurity/Pages/default.aspx (accessed 3.16.14).

Avolio, B.J., Sosik, J.J., Kahai, S.S., Baker, B., 2014. E-leadership: Re-examining transformations in leadership source and transmission. The Leadership Quarterly 25, 105–131. doi:10.1016/j.leaqua.2013.11.003

Barney, J., 1991. Firm Resources and Sustained Competitive Advantage. Journal of Management 17, 99–120. doi:10.1177/014920639101700108

Bayuk, J.L., 2012. Cyber security policy guidebook. Wiley, Hoboken, N.J.

BCS, 2015. IT skills framework (SFIAplus) [WWW Document]. SFIAplus | Business solutions | BCS - The Chartered Institute for IT. URL www.bcs.org/sfiaplus (accessed 6.1.15).

Belton, V., Stewart, T., 2002. Multiple Criteria Decision Analysis: An Integrated Approach. Springer.

Bernard, H.R., 2000. Social Research Methods: Qualitative and Quantitative Approaches. Sage Publications, Inc, Thousand Oaks, Calif.

BIS, 2010. Guidelines for Managing Projects.

Borst, W.N., 1997. Construction of Engineering Ontologies for Knowledge Sharing and Reuse (info:eu-repo/semantics/doctoralThesis). Universiteit Twente, Enschede.

Boselie, P., Dietz, G., Boon, C., 2005. Commonalities and contradictions in HRM and performance research. Human Resource Management Journal 15, 67–94. doi:10.1111/j.1748-8583.2005.tb00154.x

Brazilian Government, 2015. Information and Communications Security and Cyber Security Strategy for the Federal Public Administration 2015-2018.

Brazilian Government, 2014. Regulatory ordinance 2777-MD. Brazilian Federal Register 7.

Brazilian Government, 2013a. Complementary Norm #17 for Information and Communications Security (Norma Complementar 17 à IN Nº 01 GSI/PR/2008 - Segurança da Informação e Comunicações) [WWW Document]. Department of Information and Communications Security. URL http://dsic.planalto.gov.br/documentos/nc_17_profissionais_sic.pdf (accessed 4.1.14).

Brazilian Government, 2013b. Information Security and Cryptography National Network [WWW Document]. Rede Nacional em Segurança da Informação e Criptografia. URL http://www.renasic.org.br/ (accessed 3.19.14).

Brazilian Government, 2008a. National Strategy of Defence - Decrete nº 6703.

Brazilian Government, 2008b. Complementary Norms for Information and Communications Security (Normas Complementares à IN Nº 01 GSI/PR/2008 - Segurança da Informação e Comunicações) [WWW Document]. Department of Information and Communications Security. URL http://dsic.planalto.gov.br/legislacaodsic/23-dsic/legislacao/53-normas-complementares (accessed 4.1.14).

Brewster, C., O'Hara, K., 2007. Knowledge representation with ontologies: Present challenges—Future possibilities. International Journal of Human-Computer Studies 65, 563–568. doi:10.1016/j.ijhcs.2007.04.003

Broens, T., 2004. Context-aware, Ontology based, Semantic Service Discovery (Master's thesis). University of Twente, Enschende, The Netherlands.

Brown, J.S., Duguid, P., 1991. Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation. Organization Science 2, 40–57.

Burrell, G., Morgan, G., 1979. Sociological Paradigms and Organisational Analysis. Ashgate Publishing Limited, London.

Butler, T., Feller, J., Pope, A., Emerson, B., Murphy, C., 2008. Designing a core IT artefact for Knowledge Management Systems using participatory action research in a government and a non-government organisation. The Journal of Strategic Information Systems 17, 249–267. doi:10.1016/j.jsis.2007.10.002

Caldwell, T., 2013. Plugging the cyber-security skills gap. Computer Fraud & Security 2013, 5–10. doi:10.1016/S1361-3723(13)70062-9

Calegari, S., Pasi, G., 2013. Personal ontologies: Generation of user profiles based on the YAGO ontology. Information Processing & Management, Personalization and Recommendation in Information Access 49, 640–658. doi:10.1016/j.ipm.2012.07.010

Carayannis, E.G., Alexander, J., Ioannidis, A., 2000. Leveraging knowledge, learning, and innovation in forming strategic government–university–industry (GUI) R&D partnerships in the US, Germany, and France. Technovation 20, 477–488. doi:10.1016/S0166-4972(99)00162-5

Castaño, A.C., 2008. Populando ontologias através de informações em HTML - o caso do currículo lattes (text). Universidade de São Paulo.

Cebula, J.L., Young, L.R., 2010. A Taxonomy of Operational Cyber Security Risks. Carnegie Mellon University, Pittsburgh.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2, 429–444. doi:10.1016/0377-2217(78)90138-8

Chebat, J.-C., Filiatrault, P., Katz, A., Tal, S.M., 1994. Strategic auditing of human and financial resource allocation in marketing: An empirical study using data envelopment analysis. Journal of Business Research 31, 197–208. doi:10.1016/0148-2963(94)90083-3

Cheese, P., 2007. The Talent Powered Organization: Strategies for Globalization, Talent Management and High Performance. Kogan Page Publishers.

Chen, C.L.P., Zhang, C.-Y., 2014. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences 275, 314–347. doi:10.1016/j.ins.2014.01.015

Chen, Y., Kilgour, D.M., Hipel, K.W., 2008. Screening in multiple criteria decision analysis. Decision Support Systems, I.T. and Value Creation 45, 278–290. doi:10.1016/j.dss.2007.12.017

Choo, C.W., 2000. Working with knowledge: how information professionals help organisations manage what they know. Library Management 21, 395–403. doi:10.1108/01435120010342770

Choo, K.-K.R., 2011. The cyber threat landscape: Challenges and future research directions. Computers & Security 30, 719–731. doi:10.1016/j.cose.2011.08.004

Clarke, R.A., Knake, R.K., 2012. Cyber War: The Next Threat to National Security and What to Do About It, Reprint edition. ed. Ecco, New York.

CNPq, 2015. Lattes Panel - Statistics from Lattes Platform Curricula Database [WWW Document]. URL http://estatico.cnpq.br/painelLattes/ (accessed 2.19.15).

Collings, D.G., Mellahi, K., 2009. Strategic talent management: A review and research agenda. Human Resource Management Review 19, 304–313. doi:10.1016/j.hrmr.2009.04.001

Colucci, S., Tinelli, E., Di Sciascio, E., Donini, F.M., 2011. Automating competence management through non-standard reasoning. Engineering Applications of Artificial Intelligence 24, 1368–1384. doi:10.1016/j.engappai.2011.05.015

Conrad, J., 2012. Seeking help: the important role of ethical hackers. Network Security 2012, 5–8. doi:10.1016/S1353-4858(12)70071-5

Conti, G., Easterly, J., 2010. Recruiting, Development, and Retention of Cyber Warriors Despite an Inhospitable Culture. Small Wars Journal.

Cook, W.D., Seiford, L.M., 2009. Data envelopment analysis (DEA) – Thirty years on. European Journal of Operational Research 192, 1–17. doi:10.1016/j.ejor.2008.01.032

Cook, W.D., Tone, K., Zhu, J., 2014. Data envelopment analysis: Prior to choosing a model. Omega 44, 1–4. doi:10.1016/j.omega.2013.09.004

Corcho, O., Fernández-López, M., Gómez-Pérez, A., 2003. Methodologies, tools and languages for building ontologies. Where is their meeting point? Data & Knowledge Engineering 46, 41–64. doi:10.1016/S0169-023X(02)00195-7

Cranfield University, 2014. Cyber Defence and Information Assurance [WWW Document]. Cranfield University. URL http://www.cranfield.ac.uk/courses/masters/cyber-defence-and-information-assurance.html (accessed 7.24.14).

Davenport, T.H., Prusak, L., 2000. Working knowledge: how organizations manage what they know. Harvard Business School Press, Boston, Mass.

Davis, R., Shrobe, H., Szolovits, P., 1993. What is a knowledge representation? AI magazine 14, 17–33.

Deloitte, 2011. Talent edge 2020: Blueprints for the new normal | Organisation & Change | Human Capital | Consulting | Deloitte UK [WWW Document]. Deloitte. URL https://www.deloitte.com/view/en_GB/uk/services/consulting/5aa3f762c14be210VgnVCM1000001a56f00aRCRD.htm (accessed 4.1.14).

Demirkan, H., Kauffman, R.J., Vayghan, J.A., Fill, H.-G., Karagiannis, D., Maglio, P.P., 2008. Service-oriented technology and management: Perspectives on research and practice for the coming decade. Electronic Commerce Research and Applications 7, 356–376. doi:10.1016/j.elerap.2008.07.002

Dervin, B., Foreman-Wernet, L., Lauterbach, E., 2003. Sense-making Methodology Reader: Selected Writings of Brenda Dervin. Hampton Press, Cresskill, N.J.

Di Tria, F., Lefons, E., Tangorra, F., 2017. Cost-benefit analysis of data warehouse design methodologies. Information Systems 63, 47–62. doi:10.1016/j.is.2016.06.006

Dodd, F.J., Donegan, H.A., 1995. Comparison of Prioritization Techniques Using Interhierarchy Mappings. The Journal of the Operational Research Society 46, 492–498. doi:10.2307/2584596

Dodge Jr., R.C., Carver, C., Ferguson, A.J., 2007. Phishing for user security awareness. Computers & Security 26, 73–80. doi:10.1016/j.cose.2006.10.009

Dries, N., 2013. The psychology of talent management: A review and research agenda. Human Resource Management Review 23, 272–285. doi:10.1016/j.hrmr.2013.05.001

Edinburgh Napier University, 2014. Advanced Security and Digital Forensics - MSc at Edinburgh Napier University [WWW Document]. URL http://www.courses.napier.ac.uk/w56732.htm (accessed 12.14.14).

El Ouirdi, M., Segers, J., El Ouirdi, A., Pais, I., 2015. Predictors of job seekers' self-disclosure on social media. Computers in Human Behavior 53, 1–12. doi:10.1016/j.chb.2015.06.039

ENISA, 2014. An evaluation framework for Cyber Security Strategies — ENISA. European Union Agency for Network and Information Security.

Erl, T., 2007. SOA Principles of Service Design, 1 edition. ed. Prentice Hall, Upper Saddle River, NJ.

e-skills UK, 2014. Cyber Academy [WWW Document]. URL http://www.e-skills.com/professional-development/cyber-security/ (accessed 7.21.14).

Farndale, E., Pai, A., Sparrow, P., Scullion, H., 2014. Balancing individual and organizational goals in global talent management: A mutual-benefits perspective. Journal of World Business 49, 204–214. doi:10.1016/j.jwb.2013.11.004

Feng, Y. j., Lu, H., Bi, K., 2004. An AHP/DEA method for measurement of the efficiency of R&D management activities in universities. International Transactions in Operational Research 11, 181–191. doi:10.1111/j.1475-3995.2004.00450.x

Ferguson, M., 2014. The Hadoop Data Refinery and Enterprise Data Hub [WWW Document]. MAPR. URL https://www.mapr.com/resources/hadoop-data-refinery-and-enterprise-data-hub (accessed 10.6.14).

Ferris, G.R., Perrewé, P.L., Ranft, A.L., Zinko, R., Stoner, J.S., Brouer, R.L., Laird, M.D., 2007. Human resources reputation and effectiveness. Human Resource Management Review, The Status of Theory and Research in Human Resource Management: Where Have We Been and Where Should We Go From Here? 17, 117–130. doi:10.1016/j.hrmr.2007.03.003

Festing, M., Schäfer, L., 2014. Generational challenges to talent management: A framework for talent retention based on the psychological-contract perspective. Journal of World Business, Talent Management 49, 262–271. doi:10.1016/j.jwb.2013.11.010

Figueira, J., Greco, S., Ehrgott, M., 2005. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer.

Finlay, P.N., 1989. Introducing Decision Support Systems. Blackwell Publishers, Manchester.

Foddy, W., 1993. Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research, 1 edition. ed. Cambridge University Press, Cambridge, UK ; New York, NY, USA.

Fontenele, M., Sampaio, R.B., da Silva, A.I.B., Fernandes, J.H.C., Sun, L., 2014. Performing Knowledge Requirements Analysis for Public Organisations in a

Virtual Learning Environment: A Social Network Analysis Approach. J Inform Tech Softw Eng 4. doi:10.4172/2165-7866.1000134

Fontenele, M., Sun, L., 2016. Knowledge management of cyber security expertise: an ontological approach to the talent discovery, in: Cyber Science 2016: Pioneering Research & Innovation in Cyber Situational Awareness. Presented at the International Conference on Cyber Security and Protection of Digital Services (Cyber Security 2016), Centre For Multidisciplinary Research, Innovation And Collaboration, London, pp. 466–490. doi:10.1109/CyberSecPODS.2016.7502356

Freeman, R.E., Harrison, J.S., Wicks, A.C., 2007. Managing for Stakeholders: Survival, Reputation, and Success. Yale University Press.

French Government, 2011. French Secretariat-General for Defence and National Security - Information systems defence and security. France's strategy [WWW Document]. URL http://www.ssi.gouv.fr/IMG/pdf/2011-02-15_Information_system_defence_and_security_-_France_s_strategy.pdfIn (accessed 3.17.14).

Gable, G.G., 1994. Integrating case study and survey research methods: an example in information systems. European Journal of Information Systems 3, 112–126.

Gagne, F., 2000. Understanding the complex choreography of talent development through DMGT-based analysis, in: International Handbook of Giftedness and Talent. pp. 67–79.

Gallardo-Gallardo, E., Dries, N., González-Cruz, T.F., 2013. What is the meaning of "talent" in the world of work? Human Resource Management Review 23, 290–300. doi:10.1016/j.hrmr.2013.05.002

García-Sánchez, F., Martínez-Béjar, R., Contreras, L., Fernández-Breis, J.T., Castellanos-Nieves, D., 2006. An ontology-based intelligent system for recruitment. Expert Systems with Applications 31, 248–263. doi:10.1016/j.eswa.2005.09.023

Geers, K., 2010. The challenge of cyber attack deterrence. Computer Law & Security Review 26, 298–303. doi:10.1016/j.clsr.2010.03.003

Gelens, J., Dries, N., Hofmans, J., Pepermans, R., 2013. The role of perceived organizational justice in shaping the outcomes of talent management: A research agenda. Human Resource Management Review 23, 341–353. doi:10.1016/j.hrmr.2013.05.005

German Federal Ministry of the Interior, 2011. IT-Beauftragte der Bundesregierung - Downloads [WWW Document]. Cyber Security Strategy for Germany. URL http://www.cio.bund.de/SharedDocs/Publikationen/DE/Strategische-Themen/css_engl_download.pdf (accessed 3.17.14).

Get Safe Online | Free online security advice [WWW Document], 2016. URL https://www.getsafeonline.org/ (accessed 3.14.14).

Gewin, V., 2010. Collaboration: Social networking seeks critical mass. Nature 468, 993–994. doi:10.1038/nj7326-993a

Gewin, V., 2009. Networking in VIVO. Nature 462, 123–123. doi:10.1038/nj7269-123a

Giuffrida, R., Dittrich, Y., 2013. Empirical studies on the use of social software in global software development – A systematic mapping study. Information and Software Technology 55, 1143–1164. doi:10.1016/j.infsof.2013.01.004

Gjelten, T., 2011. Volunteer Cyber Army Emerges In Estonia [WWW Document]. NPR.org. URL http://www.npr.org/2011/01/04/132634099/in-estonia-volunteer-cyber-army-defends-nation (accessed 4.1.14).

Golemati, M., Katifori, A., Vassilakis, C., Lepouras, G., Halatsis, C., 2007. Creating an Ontology for the User Profile: Method and Applications, in: Proceedings of the First International Conference on Research Challenges in Information Science. Presented at the RCIS 2007, Ouarzazate, Morocco, pp. 407–412.

Gómez-Pérez, A., Fernández-López, M., Corcho, O., 2004. Ontological engineering with examples from the areas of knowledge management, e-commerce and the Semantic Web. Springer, London; New York.

Gruber, T.R., 1993. A translation approach to portable ontology specifications. Knowledge Acquisition 5, 199–220. doi:10.1006/knac.1993.1008

Guba, E.G., 1990. The Paradigm Dialog. SAGE Publications, Inc, Newbury Park, Calif.

Guba, E.G., Lincoln, Y.S., 1994. Competing paradigms in qualitative research. Handbook of qualitative research 2, 163–194.

Gubler, M., 2011. Protean and boundaryless career orientations - an empirical study of IT professionals in Europe (Ph.D.). Loughborough University.

Gümüş, S., Apak, S., Gümüş, H.G., Kurban, Z., 2013. An Application in Human Resources Management for Meeting Differentiation and Innovativeness Requirements of Business: Talent Management. Procedia - Social and Behavioral Sciences, The Proceedings of 9th International Strategic Management Conference 99, 794–808. doi:10.1016/j.sbspro.2013.10.551

Güngör, Z., Serhadlıoğlu, G., Kesen, S.E., 2009. A fuzzy AHP approach to personnel selection problem. Applied Soft Computing 9, 641–646. doi:10.1016/j.asoc.2008.09.003

Harris, J., 2011. Techniques for managing strategic partnership working arrangements in local government (Ph.D.). University of Reading.

Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., Ullah Khan, S., 2015. The rise of "big data" on cloud computing: Review and open research issues. Information Systems 47, 98–115. doi:10.1016/j.is.2014.07.006

Hatami-Marbini, A., Emrouznejad, A., Tavana, M., 2011. A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making. European Journal of Operational Research 214, 457–472. doi:10.1016/j.ejor.2011.02.001

Hildebrandt, M., Gutwirth, S. (Eds.), 2008. Profiling the European Citizen. Springer Netherlands, Dordrecht.

Hiltrop, J.-M., 1999. The quest for the best: human resource practices to attract and retain talent. European Management Journal 17, 422–430. doi:10.1016/S0263-2373(99)00022-5

Hinrichs, J.R., 1966. High-talent personnel; managing a critical resource. American Management Association.

Hirschheim, R., Klein, H.K., Lyytinen, K., 1995. Information Systems Development and Data Modeling: Conceptual and Philosophical Foundations. Cambridge University Press.

Hor, F.C., Huang, L.-C., Shih, H.-S., Lee, Y.-H., Stanley Lee, E., 2010. Establishing talent management for company's succession planning through analytic network process: Application to an MNC semiconductor company in Taiwan. Computers & Mathematics with Applications 60, 528–540. doi:10.1016/j.camwa.2010.05.001

Hu, Z., Wang, X.T., 2014. Trust or not: Heuristics for making trust-based choices in HR management. Journal of Business Research 67, 1710–1716. doi:10.1016/j.jbusres.2014.02.017

Hughes, J.C., Rog, E., 2008. Talent management: A strategy for improving employee recruitment, retention and engagement within hospitality organizations. International Journal of Contemporary Hospitality Management 20, 743–757. doi:10.1108/09596110810899086

Huselid, M.A., 1995. The Impact Of Human Resource Management Practices On Turnover, Productivity, And Corporate Financial Performance. ACAD MANAGE J 38, 635–672. doi:10.2307/256741

Iles, P., 2013. Commentary on “The meaning of ‘talent’ in the world of work.” Human Resource Management Review 23, 301–304. doi:10.1016/j.hrmr.2013.08.002

Infosecurity Europe, 2014. The 2014 Cyber Security Breaches Survey official launch - All Sessions - Infosecurity Europe [WWW Document]. URL http://www.infosec.co.uk/en/Sessions/4737/The-2014-Cyber-Security-Breaches-Survey-official-launch (accessed 7.24.14).

Ishizaka, A., Labib, A., 2011. Review of the main developments in the analytic hierarchy process. Expert Systems with Applications 38, 14336–14345. doi:10.1016/j.eswa.2011.04.143

ISO/IEC 27000:2009, 2009. ISO - International Organization for Standardization - ISO/IEC 27000:2009 Information technology — Security techniques — Information security management systems — Overview and vocabulary.

ISO/IEC 27032:2012, 2012. ISO - International Organization for Standardization - ISO/IEC 27032:2012 Information technology -- Security techniques -- Guidelines for cybersecurity.

ITU, 2014. ITU - International Telecommunication Union [WWW Document]. ITU. URL http://www.itu.int/en/ITU-T/studygroups/com17/Pages/cybersecurity.aspx (accessed 3.15.14).

Jambari, D.I., 2013. A method for the evaluation of business aligned IT services at the enterprise operation (Ph.D.). University of Reading.

Jeners, N., Prinz, W., 2012. From Groupware to Social Media - A Comparison of Conceptual Models, in: Camarinha-Matos, L.M., Xu, L., Afsarmanesh, H. (Eds.), Collaborative Networks in the Internet of Services, IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, pp. 416–423.

Josang, A., Ismail, R., Boyd, C., 2007. A survey of trust and reputation systems for online service provision. Decision Support Systems, Emerging Issues in Collaborative Commerce 43, 618–644. doi:10.1016/j.dss.2005.05.019

Julisch, K., 2013. Understanding and overcoming cyber security anti-patterns. Computer Networks 57, 2206–2211. doi:10.1016/j.comnet.2012.11.023

Kamal, M.M., 2011. The case of EAI facilitating knowledge management integration in local government domain. International Journal of Information Management 31, 294–300. doi:10.1016/j.ijinfomgt.2011.02.002

Kang, D., Lee, J., Choi, S., Kim, K., 2010. An ontology-based Enterprise Architecture. Expert Systems with Applications 37, 1456–1464. doi:10.1016/j.eswa.2009.06.073

Kanij, T., Grundy, J., Merkel, R., 2014. Performance appraisal of software testers. Information and Software Technology, Performance in Software Development 56, 495–505. doi:10.1016/j.infsof.2013.11.002

Kaplan, A.M., Haenlein, M., 2010. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 53, 59–68. doi:10.1016/j.bushor.2009.09.003

Kelly, K., 1994. Out of control. Perseus Books, Reading, Mass.

Khodabakhshi, M., Aryavash, K., 2012. Ranking all units in data envelopment analysis. Applied Mathematics Letters 25, 2066–2070. doi:10.1016/j.aml.2012.04.019

Klein, H.K., Myers, M.D., 1999. A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. MIS Quarterly 23, 67–93. doi:10.2307/249410

Klimburg, A., 2012. National Cyber Security Framework Manual. NATO CCD COE Publication, Tallinn.

Kotonya, G., Sommerville, I., 1998. Requirements Engineering: Processes and Techniques. Wiley Publishing.

Koutrouli, E., Tsalgatidou, A., 2012. Taxonomy of attacks and defense mechanisms in P2P reputation systems—Lessons for reputation system designers. Computer Science Review 6, 47–70. doi:10.1016/j.cosrev.2012.01.002

Koutrouli, E., Tsalgatidou, A., 2006. Reputation-Based Trust Systems for P2P Applications: Design Issues and Comparison Framework, in: Fischer-Hübner, S., Furnell, S., Lambrinoudakis, C. (Eds.), Trust and Privacy in Digital Business, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 152–161.

Kouttis, S., 2016. Improving security knowledge, skills and safety. Computer Fraud & Security 2016, 12–14. doi:10.1016/S1361-3723(16)30037-9

Kuhn, T.S., 2012. The Structure of Scientific Revolutions: 50th Anniversary Edition. University of Chicago Press.

Kum, H.-C., Duncan, D.F., Stewart, C.J., 2009. Supporting self-evaluation in local government via Knowledge Discovery and Data mining. Government Information Quarterly 26, 295–304. doi:10.1016/j.giq.2008.12.009

Lancaster University, 2014. Cyber Security MSc :: School of Computing and Communications [WWW Document]. URL http://www.scc.lancs.ac.uk/masters/CyberSecurity (accessed 10.15.14).

Land, M.O., Proper, E., Waage, M., 2008. Enterprise Architecture: Creating Value by Informed Governance, 2009th ed. Springer, New York.

Lapalme, J., Gerber, A., Van der Merwe, A., Zachman, J., Vries, M.D., Hinkelmann, K., 2016. Exploring the future of enterprise architecture: A Zachman perspective. Computers in Industry, Special Issue on Future Perspectives On Next Generation Enterprise Information Systems 79, 103–113. doi:10.1016/j.compind.2015.06.010

Lavine, T.Z., Tejera, V., Horowitz, I.L., 2011. History and Anti-History in Philosophy, Second Edition edition. ed. Transaction Publishers, New Brunswick, N.J.

Lee, M.R., Lan, Y.-C., 2011. Toward a unified knowledge management model for SMEs. Expert Systems with Applications 38, 729–735. doi:10.1016/j.eswa.2010.07.025

Lepak, D.P., Snell, S.A., 2002. Examining the Human Resource Architecture: The Relationships Among Human Capital, Employment, and Human Resource Configurations. Journal of Management 28, 517–543. doi:10.1177/014920630202800403

Lewicki, R.J., McAllister, D.J., Bies, R.J., 1998. Trust And Distrust: New Relationships and Realities. ACAD MANAGE REV 23, 438–458. doi:10.5465/AMR.1998.926620

Lewis, J.A., 2005. Aux armes, citoyens: Cyber security and regulation in the United States. Telecommunications Policy 29, 821–830. doi:10.1016/j.telpol.2005.06.009

Lewis, J.A., Timlin, K., 2011. Cybersecurity and cyberwarfare: Preliminary assessment of national doctrine and organization. UNIDIR.

Lewis, R.E., Heckman, R.J., 2006. Talent management: A critical review. Human Resource Management Review 16, 139–154. doi:10.1016/j.hrmr.2006.03.001

Lin, H.T., 2010. Personnel selection using analytic network process and fuzzy data envelopment analysis approaches. Computers & Industrial Engineering 59, 937–944. doi:10.1016/j.cie.2010.09.004

Lin, I.C., Wu, H.J., Li, S.F., Cheng, C.Y., 2015. A fair reputation system for use in online auctions. Journal of Business Research, Special Issue on Global entrepreneurship and innovation in management Special Issue on Global entrepreneurship and innovation in management 68, 878–882. doi:10.1016/j.jbusres.2014.11.045

Lincoln, Y.S., Guba, E.G., 1985. Naturalistic Inquiry. SAGE Publications, Inc, London.

Liu, K., 2000. Semiotics in information systems engineering. Cambridge University Press Cambridge.

Liu, K., Li, W., 2015. Organisational Semiotics for Business Informatics. Routledge.

Liu, K., Sun, L., Barjis, J., Dietz, J.L.G., 2003. Modelling dynamic behaviour of business organisations—extension of DEMO from a semiotic perspective. Knowledge-Based Systems 16, 101–111. doi:10.1016/S0950-7051(02)00077-1

Liu, K., Sun, L., Fu, Y., 2008. Ontological modelling of content management and provision. Information and Software Technology 50, 1155–1164. doi:10.1016/j.infsof.2008.03.008

Liu, K., Sun, L., Tan, S., 2006. Modelling complex systems for project planning: a semiotics motivated method. International Journal of General Systems 35, 313–327. doi:10.1080/03081070600661028

McAfee, A., Brynjolfsson, E., 2012. Big data: the management revolution. Harvard business review 90, 60–68.

McKnight, D.H., Chervany, N.L., 2001. Trust and Distrust Definitions: One Bite at a Time, in: Falcone, R., Singh, M., Tan, Y.-H. (Eds.), Trust in Cyber-Societies, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 27–54.

Merriam, S.B., 2014. Qualitative Research: A Guide to Design and Implementation, 3rd ed. John Wiley & Sons, San Francisco.

Michaels, E., Handfield-Jones, H., Axelrod, B., 2001. The War for Talent. Harvard Business Press.

Moore, J.F., 2006. Business Ecosystems and the View from the Firm. The Antitrust Bulletin 51, 31–75. doi:10.1177/0003603X0605100103

Mushi, C.J., 2012. A method of modelling user requirements for enhancing the quality of government to citizen service provision. University of Reading, Reading.

Myers, M.D., 1997. Qualitative Research in Information Systems. MISQ Discovery 241–242.

Nicholson, A., Webber, S., Dyer, S., Patel, T., Janicke, H., 2012. SCADA security in the light of Cyber-Warfare. Computers & Security 31, 418–436. doi:10.1016/j.cose.2012.02.009

Nonaka, I., Krogh, G. von, Voelpel, S., 2006. Organizational Knowledge Creation Theory: Evolutionary Paths and Future Advances. Organization Studies 27, 1179–1208. doi:10.1177/0170840606066312

NYU Polytechnic School of Engineering, 2014. Cybersecurity, MS: NYU-ePoly | NYU-ePoly Online Graduate Program | NYU Polytechnic School of Engineering [WWW Document]. URL http://engineering.nyu.edu/academics/online/masters/cybersecurity (accessed 11.13.14).

OASIS, 2006. Reference Model for Service Oriented Architecture v1.0 [WWW Document]. URL http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.html (accessed 12.1.14).

Oh, D.-S., Phillips, F., Park, S., Lee, E., 2016. Innovation ecosystems: A critical examination. Technovation 54, 1–6. doi:10.1016/j.technovation.2016.02.004

OMG, 2007. OMG Unified Modeling Language Superstructure, v. 2.1.2.

Orlikowski, W.J., Baroudi, J.J., 1991. Studying Information Technology in Organizations: Research Approaches and Assumptions. Information Systems Research 2, 1–28. doi:10.1287/isre.2.1.1

Ousmanou, K., 2007. A Method for the Articulation of Users' Requirements for Personalised Information Provision (Ph.D.). University of Reading.

Owen, J.M., 1999. Knowledge management and the information professional. Information Services and Use 19, 7–16.

Oxford Dictionaries, 2009. Oxford Paperback Dictionary & Thesaurus, 3 edition. ed. OUP Oxford, Oxford ; New York.

Parmenter, D., 2010. Key performance indicators [electronic resource] : developing, implementing, and using winning KPIs, 2nd ed. John Wiley & Sons, Hoboken, N.J.

Pearson, I.L.G., 2011. Smart grid cyber security for Europe. Energy Policy 39, 5211–5218. doi:10.1016/j.enpol.2011.05.043

Phillips, E., Pugh, D.S., 2000. How to Get a PhD: A Handbook for Students and Their Supervisors, 3rd Revised edition edition. ed. Open University Press, Buckingham England ; Philadelphia.

Pickard, A., 2002. Access to electronic information resources: their role in the provision of learning opportunities for young people. A constructivist inquiry (doctoral). University of Northumbria.

Public Safety Canada, 2014. Canada’s Cyber Security Strategy [WWW Document]. URL http://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/cbr-scrt-strtgy/index-eng.aspx (accessed 3.16.14).

Reece, R.P., Stahl, B.C., 2015. The professionalisation of information security: Perspectives of UK practitioners. Computers & Security 48, 182–195. doi:10.1016/j.cose.2014.10.007

ResearchGate, 2014. RG Score FAQ - ResearchGate [WWW Document]. RG Score A new way to measure scientific reputation. URL https://www.researchgate.net/publicprofile.RGScoreFAQ.html (accessed 8.13.14).

Rich, E., 1983. Users are individuals: individualizing user models. International journal of man-machine studies 18, 199–214.

Rowley, J., 2007. The wisdom hierarchy: representations of the DIKW hierarchy. Journal of Information Science 33, 163–180. doi:10.1177/0165551506070706

Russell, M.A., 2011. Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites 1st, 1 edition. ed. O’Reilly Media.

Saaty, T.L., 2008. Decision making with the analytic hierarchy process. International Journal of Services Sciences 1, 83–98. doi:10.1504/IJSSci.2008.01759

Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15, 234–281. doi:10.1016/0022-2496(77)90033-5

Sani, N.K., 2011. The co-design of business and IT systems based on organisational semiotics (Ph.D.). University of Reading.

Šaša Bastinos, A., Krisper, M., 2013. Multi-criteria decision making in ontologies. Information Sciences, Including Special Section on New Trends in Ambient Intelligence and Bio-inspired Systems 222, 593–610. doi:10.1016/j.ins.2012.07.055

Savvas, I., Bassiliades, N., 2009. A process-oriented ontology-based knowledge management system for facilitating operational procedures in public administration. Expert Systems with Applications 36, 4467–4478. doi:10.1016/j.eswa.2008.05.022

Schalk, R., Timmerman, V., den Heuvel, S. van, 2013. How strategic considerations influence decision making on e-HRM applications. Human Resource Management Review, Emerging Issues in Theory and Research on Electronic Human Resource Management (eHRM) 23, 84–92. doi:10.1016/j.hrmr.2012.06.008

Schiemann, W.A., 2014. From talent management to talent optimization. Journal of World Business 49, 281–288. doi:10.1016/j.jwb.2013.11.012

Silzer, R., Dowell, B.E., 2009. Strategy Driven Talent Management: A Leadership Imperative. John Wiley & Sons, San Francisco.

Sinuany-Stern, Z., Mehrez, A., Hadad, Y., 2000. An AHP/DEA methodology for ranking decision making units. International Transactions in Operational Research 7, 109.

Sourouni, A.-M., Kourlimpinis, G., Mouzakitis, S., Askounis, D., 2010. Towards the government transformation: An ontology-based government knowledge repository. Computer Standards & Interfaces 32, 44–53. doi:10.1016/j.csi.2009.06.002

Staab, S., Studer, R. (Eds.), 2004. Handbook on ontologies. Springer, Berlin.

Stamper, R., Kolkman, M., 1991. Problem Articulation: A Sharper-edged Soft Systems Approach. Journal of Applied Systems Analysis 69–76.

Stamper, R., Liu, K., 1994. Organisational dynamics, social norms and information systems, in: Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994. Presented at the Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994, pp. 645–654. doi:10.1109/HICSS.1994.323455

Stamper, R., Liu, K., Hafkamp, M., Ades, Y., 2000. Understanding the roles of signs and norms in organizations-a semiotic approach to information systems design. Behaviour & Information Technology 19, 15–27.

Strohmeier, S., Piazza, F., 2013. Domain driven data mining in human resource management: A review of current research. Expert Systems with Applications 40, 2410–2420. doi:10.1016/j.eswa.2012.10.059

Sun, L., Mushi, C.J., 2010. Case-based analysis in user requirements modelling for knowledge construction. Information and Software Technology 52, 770–777. doi:10.1016/j.infsof.2010.03.010

Sun, L., Ousmanou, K., Cross, M., 2010. An ontological modelling of user requirements for personalised information provision. Inf Syst Front 12, 337–356. doi:10.1007/s10796-008-9144-x

Sun, L., Yamin, M., Mushi, C., Liu, K., Alsaigh, M., Chen, F., 2014. Information analytics for healthcare service discovery. Journal of Healthcare Engineering 5, 457–478.

Sure, Y., Mädche, A., Staab, S., 2000. Leveraging Corporate Skill Knowledge – From ProPer to OntoProPer, in: Proceedings of the Third International Conference on Practical Aspects of Knowledge Management. pp. 30–31.

Swamynathan, G., Almeroth, K.C., Zhao, B.Y., 2010. The design of a reliable reputation system. Electron Commer Res 10, 239–270. doi:10.1007/s10660-010-9064-y

Tallinn University of Technology, 2014. Cyber Security [WWW Document]. URL http://www.ttu.ee/studying/masters/masters_programmes/cyber-security/?id=84572 (accessed 11.13.14).

Tansley, C., Harris, L., Stewart, J., Turner, P., 2006. Talent management: understanding the dimensions. Change Agenda 1–24.

Tansley, C., Kirk, S., Tietze, S., 2013. The currency of talent management—A reply to “talent management and the relevance of context: Towards a pluralistic approach.” Human Resource Management Review 23, 337–340. doi:10.1016/j.hrmr.2013.08.004

Tarique, I., Schuler, R.S., 2010. Global talent management: Literature review, integrative framework, and suggestions for further research. Journal of World Business, Global Talent Management 45, 122–133. doi:10.1016/j.jwb.2009.09.019

Taylor, C.S., 2013. Validity and Validation. Oxford University Press.

Teo, T.S.H., 2005. Meeting the challenges of knowledge management at the Housing and Development Board. Decision Support Systems 41, 147–159. doi:10.1016/j.dss.2004.05.013

The National Skills Academy, 2014. CESG Certified Professional skills [WWW Document]. The National Skills Academy. URL http://www.itskillsacademy.ac.uk/cyber-security-learning-pathways/cesg/ (accessed 2.12.14).

The Open Group, 2013. ArchiMate® 2.1 Specification [WWW Document]. URL http://pubs.opengroup.org/architecture/archimate2-doc/chap03.html#_Toc371945155 (accessed 10.17.14).

Thirunarayan, K., Anantharam, P., Henson, C., Sheth, A., 2014. Comparative trust management with applications: Bayesian approaches emphasis. Future Generation Computer Systems 31, 182–199. doi:10.1016/j.future.2013.05.006

Thunnissen, M., Boselie, P., Fruytier, B., 2013. Talent management and the relevance of context: Towards a pluralistic approach. Human Resource Management Review 23, 326–336. doi:10.1016/j.hrmr.2013.05.004

Tseng, Y.-F., Lee, T.-Z., 2009. Comparing appropriate decision support of human resource practices on organizational performance with DEA/AHP model. Expert Systems with Applications 36, 6548–6558. doi:10.1016/j.eswa.2008.07.066

UK Cabinet Office, 2013. Protecting and promoting the UK in a digital world: 2 years on - Press releases - GOV.UK [WWW Document]. Gov.UK. URL https://www.gov.uk/government/news/protecting-and-promoting-the-uk-in-a-digital-world-2-years-on (accessed 3.17.14).

UK Cabinet Office, 2011. The UK Cyber Security Strategy and written ministerial statement [WWW Document]. URL https://www.gov.uk/government/publications/cyber-security-strategy (accessed 3.14.14).

UKAIS, 2014. Definition of IS [WWW Document]. UKAIS: UK Academy for Information Systems. URL http://www.ukais.org.uk/about/DefinitionIS.aspx (accessed 8.31.14).

Ulrich, D., Smallwood, N., 2012. What is talent? Leader to Leader 2012, 55–61. doi:10.1002/ltl.20011

University of Dallas, 2014. Center for Cybersecurity Education - Courses [WWW Document]. URL http://www.udallas.edu/cob/centers/cybercourses (accessed 11.13.14).

University of London, 2014. Royal Holloway | Information Security (MSc) | Information Security home [WWW Document]. URL https://www.royalholloway.ac.uk/isg/coursefinder/mscinformationsecurity.aspx (accessed 12.14.14).

University of Oxford, 2014. Software and Systems Security at Oxford | Software and Systems Security [WWW Document]. URL https://www.cs.ox.ac.uk/softeng/security/ (accessed 12.14.14).

University of Surrey, 2014. Information Security [WWW Document]. University of Surrey - Guildford. URL http://www.surrey.ac.uk/postgraduate/information-security (accessed 12.14.14).

University of Warwick, 2014. Full time Masters (MSc) in Cyber Security and Management [WWW Document]. URL http://www2.warwick.ac.uk/fac/sci/wmg/education/wmgmasters/courses/cyber_security_management/ (accessed 11.13.14).

US Government, 2014. National Initiative for Cybersecurity Careers and Studies (NICCS) [WWW Document]. URL http://niccs.us-cert.gov/ (accessed 8.5.14).

US Government, 2009. 2009 Cyberspace Policy Review [WWW Document]. Department of Homeland Security. URL http://www.dhs.gov/publication/2009-cyberspace-policy-review (accessed 3.17.14).

Vaidya, O.S., Kumar, S., 2006. Analytic hierarchy process: An overview of applications. European Journal of Operational Research 169, 1–29. doi:10.1016/j.ejor.2004.04.028

Vaishnavi, V.K., Jr, W.K., 2007. Design Science Research Methods and Patterns: Innovating Information and Communication Technology. Auerbach Publications, Boca Raton.

van Dam, J.-W., van de Velden, M., 2015. Online profiling and clustering of Facebook users. Decision Support Systems 70, 60–72. doi:10.1016/j.dss.2014.12.001

Vavilis, S., Petković, M., Zannone, N., 2014. A reference model for reputation systems. Decision Support Systems 61, 147–154. doi:10.1016/j.dss.2014.02.002

von Solms, R., van Niekerk, J., 2013. From information security to cyber security. Computers & Security 38, 97–102. doi:10.1016/j.cose.2013.04.004

Voskoglou, M.G., 2011. Case-Based Reasoning: History, Methodology and Development Trends, in: Leeland, A.M. (Ed.), Case-Based Reasoning [Electronic Resource] : Processes, Suitability and Applications, Engineering Tools, Techniques and Tables. Nova Science Publishers, New York.

Wang, W., Lu, Z., 2013. Cyber security in the Smart Grid: Survey and challenges. Computer Networks 57, 1344–1371. doi:10.1016/j.comnet.2012.12.017

Wang, Y.-M., Chin, K.-S., Poon, G.K.K., 2008. A data envelopment analysis method with assurance region for weight generation in the analytic hierarchy process. Decision Support Systems, Information Technology and Systems in the Internet-Era 45, 913–921. doi:10.1016/j.dss.2008.03.002

Webber, S.A., Apostolou, B., Hassell, J.M., 1997. The sensitivity of the analytic hierarchy process to alternative scale and cue presentations. European Journal of Operational Research 96, 351–362. doi:10.1016/S0377-2217(96)00081-1

Weber, S., 2010. Design Science Research: Paradigm or Approach?, in: AMCIS 2010 Proceedings. Presented at the Americas Conference on Information Systems (AMCIS).

Wong, K.Y., Aspinwall, E., 2006. Development of a knowledge management initiative and system: A case study. Expert Systems with Applications 30, 633–641. doi:10.1016/j.eswa.2005.07.012

Yates, D., Paquette, S., 2011. Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. International Journal of Information Management 31, 6–13. doi:10.1016/j.ijinfomgt.2010.10.001

Yin, R.K., 2009. Case Study Research: Design and Methods, 4th ed. SAGE.

Yin, R.K. (Robert K.-Z., 2014. Case study research : design and methods / Robert K. Yin., 5th ed. ed, Applied social research methods series ; v. 5. Sage.

Zimmermann, A., Pretz, M., Zimmermann, G., Firesmith, D.G., Petrov, I., 2013. Towards Service-Oriented Enterprise Architectures for Big Data Applications in the Cloud, in: 2013 17th IEEE International Enterprise Distributed Object Computing Conference Workshops. Presented at the 2013 17th IEEE International Enterprise Distributed Object Computing Conference Workshops, pp. 130–135. doi:10.1109/EDOCW.2013.21

Appendix A The List of Abbreviations

ADM	Architecture Development Method
AHP	Analytic hierarchy process
AMO	Ability motivation opportunity
ANP	Analytic Network Process
API	Application programming interface
CBR	Case-based reasoning
CoP	Community of practice
CS	Cyber security
DEA	Data envelopment analysis
DECYSE	Method for Discovering Expertise in Cyber Security communities
DMU	Decision-making unit(s)
DoS	Denial of service
DSR	Design science research
DSS	Decision support system(s)
DW	Data warehouse / warehousing
EA	Enterprise architecture
EBR	Exemplar-based reasoning
END	Brazilian National Strategy of Defence (translated from *Estratégia Nacional de Defesa*)
Eq.	Equation(s)
HR(M)	Human resource (management)
I(C)T	Information (and communication) technology
IS	Information system(s)
KM(S)	Knowledge management (system)
KR	Knowledge representation
(K)PI	(Key) performance indicator(s)
(K)RI	(Key) result indicator(s)
MCDM	Multi-criteria decision-making method(s)
MEASUR	Methods for Eliciting, Analysing and Specifying Users' Requirements
NAM	Norm Analysis Method
NATO CCD COE	NATO Cooperative Cyber Defence Centre of Excellence
NICCS	National Initiative for Cybersecurity Careers and Studies
OS	Organisational semiotics
PPP	Public-private partnership(s)
RDF	Resource description framework
RENASIC	Information Security and Cryptography National Network (translated from *Rede Nacional de Segurança da Informação e Criptografia*)
RS	Reputation system(s)
SAM	Semantic Analysis Method
SCADA	Supervisory control and data acquisition
SNA	Social network analysis
SNS	Social networking site(s)
SOA	Service-oriented architecture
TM	Talent management
TOGAF	The Open Group Architecture Framework

Appendix B The Cyber Security Related Concepts

Concept	Definition
Cyber security	"the collection of tools, policies, security concepts, security safeguards, guidelines, risk management approaches, actions, training, best practices, assurance and technologies that can be used to protect the cyber environment and organization and user's assets" (ITU, 2014)
	"the preservation of confidentiality, integrity and availability of information in the Cyberspace" (ISO/IEC 27032:2012, 2012)
	"methods of using people, process and technology to prevent, detect and recover from damage to confidentiality, integrity and availability of information in cyberspace" (Bayuk, 2012: p. 3)
	"includes strategy, policy, and standards regarding the security of and operations in cyberspace, and encompasses the full range of threat reduction, vulnerability reduction, deterrence, international engagement, incident response, resiliency, and recovery policies and activities, including computer network operations, information assurance, law enforcement, diplomacy, military, and intelligence missions as they relate to the security and stability of the global information and communications infrastructure. The scope does not include other information and communications policy unrelated to national security or securing the infrastructure." (US Government, 2009: p. 2)
	"The desired state of an information system in which it can resist events from cyberspace likely to compromise the availability, integrity or confidentiality of the data stored, processed or transmitted and of the related services that these systems offer or make accessible" (French Government, 2011: p. 21)
	"actions (...) to reduce the risk and secure the benefits of a trusted digital environment for businesses and individuals" (UK Cabinet Office, 2011)
	"the desired objective of the IT security situation, in which the risks of global cyberspace have been reduced to an acceptable minimum. Hence, cyber security in Germany is the desired objective of the IT security situation, in which the risks of the German cyberspace have been reduced to an acceptable minimum. Cyber security (in Germany) is the sum of suitable and appropriate measures. Civilian cyber security focuses on all IT systems for civilian use in German cyber-space. Military cyber security focuses on all IT systems for military use in German cyberspace." (German Federal Ministry of the Interior, 2011)
	"the protection of cyberspace itself, the electronic information, the ICTs that support cyberspace, and the users of cyberspace in their personal, societal and national capacity, including any of their interests, either tangible or intangible, that are vulnerable to attacks originating in cyberspace" (von Solms and van Niekerk, 2013: p. 99)
	"Measures relating to the confidentially, availability and integrity of information that is processed, stored and communicated by electronic or similar means" (Australian Government, 2009)
	"the defense or protection of the integrity, operations and confidentiality of computers and computer networks" (Lewis, 2005: p. 821)
Cyber attacks	"Cyber attacks include the unintentional or unauthorized access, use, manipulation, interruption or destruction (via electronic means) of electronic information and/or the electronic and physical infrastructure used to process, communicate and/or store that information. The severity of the cyber attack determines the appropriate level of response and/or mitigation measures: i.e., cyber security" (Public Safety Canada, 2014)
National cyber	Comprises 3 dimensions of activity (governmental coordination, national cooperation and international collaboration), 5 mandates (military cyber, counter cyber crime,

security	intelligence / counter-intelligence, critical infrastructure protection / national crisis management and cyber diplomacy / internet governance) and 5 dilemmas (Klimburg, 2012: p. 29)
Cyber defence	“The set of all technical and non-technical measures allowing a State to defend in cyberspace information systems that it considers to be critical” (French Government, 2011: p. 21)
Cyber warfare	"(…) actions by a nation-state to penetrate another nation's computers or networks for the purposes of causing damage or disruption.” (Clarke and Knake, 2012: p. 6)
Information security	“preservation of confidentiality, integrity and availability of information. In addition, other properties, such as authenticity, accountability, non-repudiation and reliability can be involved” (ISO/IEC 27000:2009, 2009)
	“the general security objectives comprise the following: availability; integrity, which may include authenticity and non-repudiation; and confidentiality” (ITU, 2014)
	“the protection of information and its critical elements, including the systems and hardware that use, store, and transmit that information” (Whitman and Mattord, 2009, p. 8) apud (von Solms and van Niekerk, 2013: p. 98)
	“without the confidentiality, integrity, availability, non-repudiation, accountability, authenticity and reliability of information resources, information cannot be deemed secure” (von Solms and van Niekerk, 2013: p. 99)
ICT security	“all aspects relating to defining, achieving and maintaining the confidentiality, integrity, availability, non-repudiation, accountability, authenticity, and reliability of information resources” (ISO/IEC 13335-1, 2004, p. 3) apud (von Solms and van Niekerk, 2013: p. 98)
Cyberspace	“the virtual space of all IT systems linked at data level on a global scale. The basis for cyberspace is the Internet as a universal and publicly accessible connection and transport network which can be complemented and further expanded by any number of additional data networks. IT systems in an isolated virtual space are not part of cyberspace” (German Federal Ministry of the Interior, 2011: p. 14)
	“The communication space created by the worldwide interconnection of automated digital data processing equipment” (French Government, 2011: p. 21)
	“the interdependent network of information technology infrastructures, and includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers in critical industries. Common usage of the term also refers to the virtual environment of information and interactions between people” (US Government, 2009)
	“The hypothetical place in which communication over computer networks takes place” (Oxford Dictionaries, 2009: p. 223)

Appendix C The Selected Definitions of Talent and Talent Management

Concept	Definition
Talent	"a unique mix of innate intelligence or brain power, plus a certain degree of creativity or the capacity to go beyond established stereotypes and provide innovative solutions to problems in his everyday world, plus personal skills which make him effective in his relationships with his peers, his superiors, and his subordinates" (Hinrichs, 1966: p. 11)
	"(…) superior mastery of systematically developed abilities (or skills) and knowledge in at least one field of human activity, to a degree that places an individual within the top 10% of age peers who are (or have been) active in that field" (Gagne, 2000: p. 67)
	"(…) the sum of a person's abilities—his or her intrinsic gifts, skills, knowledge, experience, intelligence, judgment, attitude, character and drive. It also includes his or her ability to learn and grow." (Michaels et al., 2001: p. xii)
	"(…) is essentially a euphemism for 'people'" (Lewis and Heckman, 2006: p. 141)
	"Talent can be considered as a complex amalgam of employees' skills, knowledge, cognitive ability and potential. Employees' values and work preferences are also of major importance." (Tansley et al., 2006: p. 2)
	"Essentially, talent means the total of all the experience, knowledge, skills, and behaviours that a person has and brings to work." (Cheese, 2007: p. 46)
	"natural ability or skill", "people possessing natural ability or skill" (Oxford Dictionaries, 2009: p. 948)
	"In groups talent can refer to a pool of employees who are exceptional in their skills and abilities either in a specific technical area (such as software graphics skills) or a competency (such a consumer marketing talent), or a more general area (such as general managers or high-potential talent)." (Silzer and Dowell, 2009: p. 13-14)
	"(…) in some cases "the talent" might refer to the entire employee population." (Silzer and Dowell, 2009: p. 14)
	"Talent = competence [knowledge, skills and values required for todays' and tomorrows' jobs; right skills, right place, right job, right time] × commitment [willing to do the work] × contribution [finding meaning and purpose in their work]" (Ulrich and Smallwood, 2012: p. 60)
	"(…) the collective knowledge, skills, abilities, experiences, values, habits and behaviors of all labor that is brought to bear on the organization's mission" (Schiemann, 2014: p. 282)
Talent management	"defined here as both a philosophy and a practice. It is both an espoused and enacted commitment (…) to implementing an integrated, strategic and technology enabled approach to HRM, with a particular focus on human resource planning, including employee recruitment, retention, development and succession practices." (Hughes and Rog, 2008: p. 746)
	"We define strategic talent management as activities and processes that involve the systematic identification of key positions which differentially contribute to the organisation's sustainable competitive advantage, the development of a talent pool of high potential and high performing incumbents to fill these roles, and the development of a differentiated human resource architecture to facilitate filling these positions with competent incumbents and to ensure their continued commitment to the organisation" (Collings and Mellahi, 2009: p. 304)
	"global talent management is about systematically utilizing IHRM activities (complementary HRM policies and policies) to attract, develop, and retain individuals with high levels of human capital (e.g., competency, personality, motivation) consistent with the strategic directions of the multinational enterprise in a dynamic, highly

competitive, and global environment" (Tarique and Schuler, 2010: p. 124)

"(…) we define talent management as the differential management of employees according to their relative potential to contribute to an organization's competitive advantage." (Gelens et al., 2013: p. 342)

"(…) an HR focused management process that allows organizations to overcome difficulties and to systematically close the gap between the required talents and the existing talents on their way to their targets." (Altınöz et al., 2013: p. 843)

"(…) an organization's ability to attract, select, develop, and retain key employees (in a global context)." (Festing and Schäfer, 2014: p. 263)

"(…) the way in which the talent lifecycle is managed." (Schiemann, 2014: p. 282)

Appendix D The General Service Descriptions

This Appendix presents some of the conceptual views used to describe the services within the service inventory in order to select those concepts to be documented in the DECYSE ontological model. Moreover, it serves the purpose of a repository of ideas for further improvement on the ontological model, since some business concepts were merged, updated or discarded when transformed into ontology elements. For example, Figure D-1 presents how regulation updates and courses or certifications that were accredited are published to raise awareness of participants. These activities are affected by the cyber security strategy, which also establish guidelines for the definition of profiles.

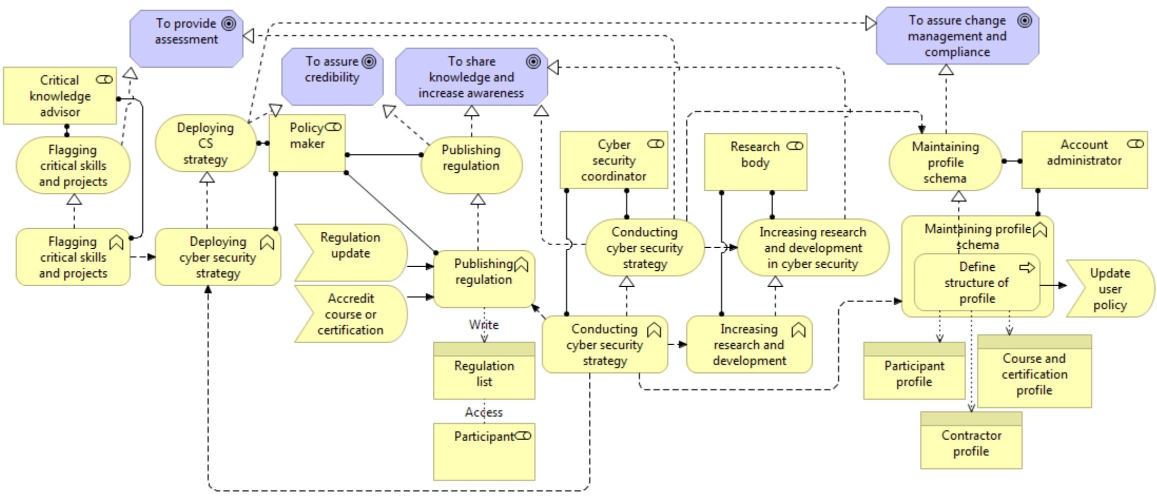


Figure D-1: Publishing regulations for all stakeholders

Managing a user profile throughout its whole lifecycle (Figure D-2) implies in a set of processes performed by multiple roles (e.g. policy maker, participant and account administrator).

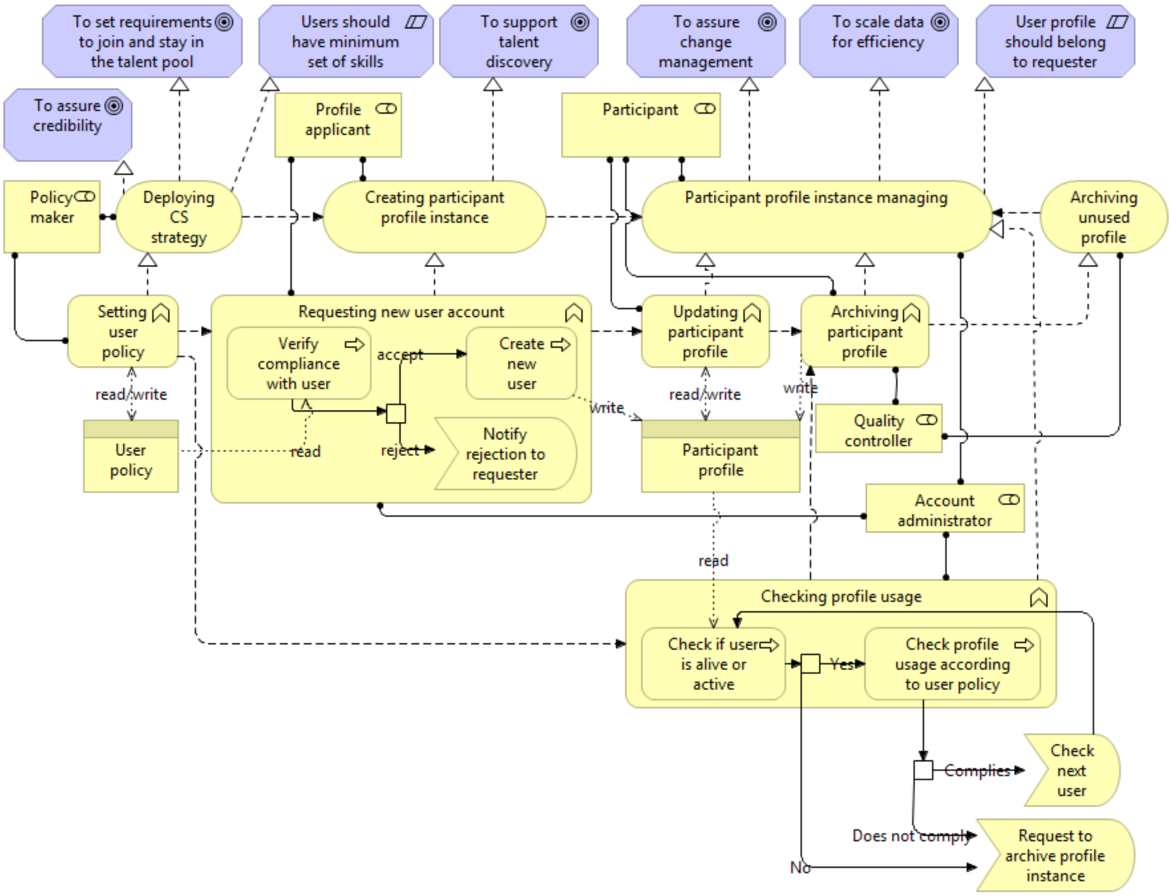


Figure D-2: Managing a user profile

Participant profiles are also updated when receiving evaluations from either peers or contractors (Figure D-3). Since the reputation algorithm is mostly similar in both types of evaluation, the common parts of the algorithms can be embedded in a reputation engine for reuse. A discussion forum service is also used as a platform where peers can interact to share knowledge and evaluate each other. However, modelling this service was beyond the scope of DECYSE.

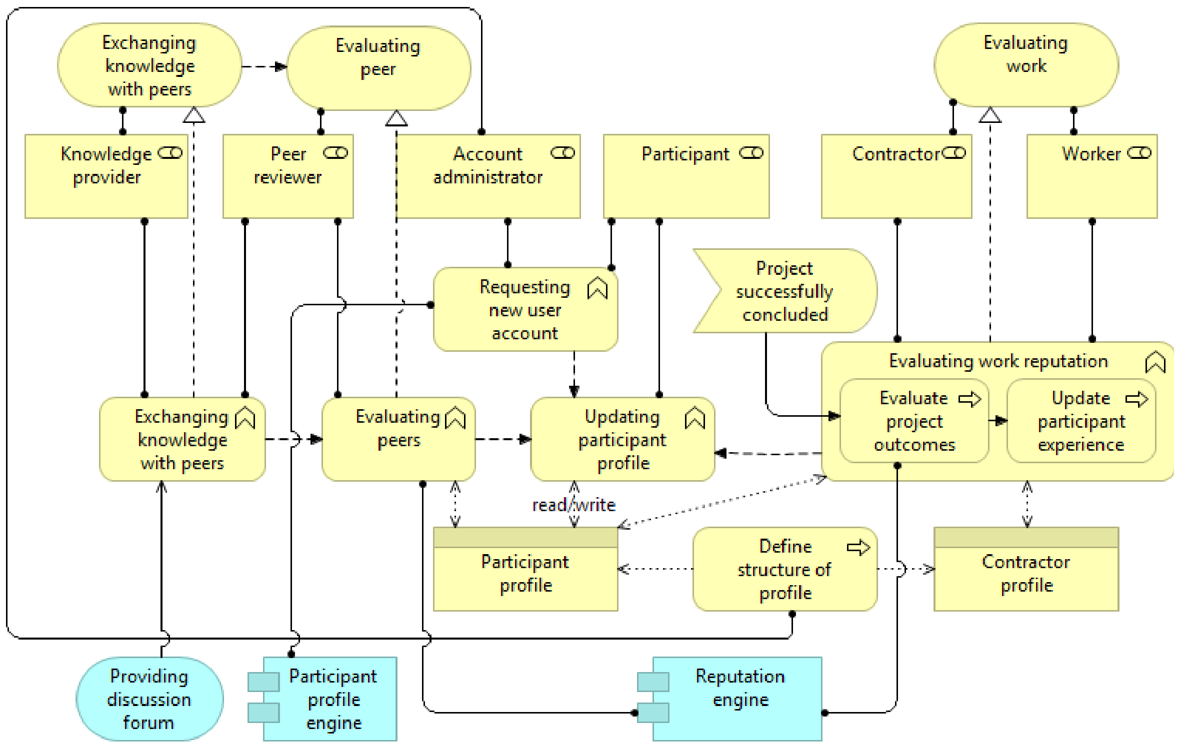


Figure D-3: Updating a participant profile

Selection for expertise comprises some of the processes presented in Figure D-4 (i.e. define project, select candidates and rank candidates). Those processes are associated with components that run the algorithms described by their respective norms. There are three components here presented to illustrate that deployment of the application layer should be guided by the processes' development. However, the DECYSE method relies only on the business layer of the CS community.

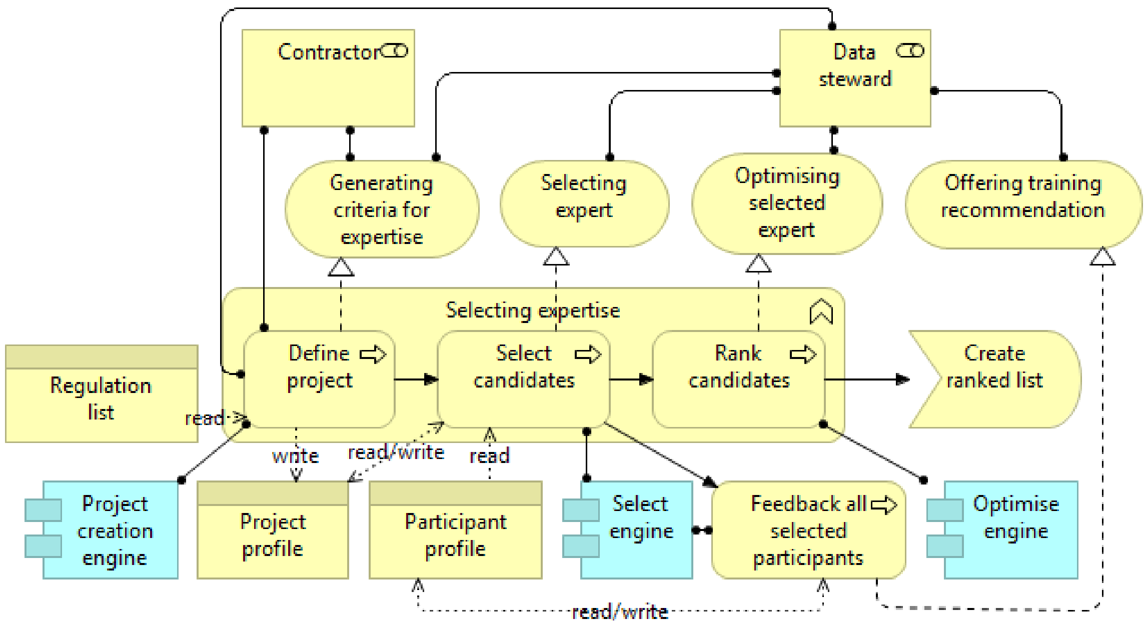


Figure D-4: Creating a project, selecting and ranking candidates for the project and providing feedback for the candidates with project requirements

Aligned with the principles of systems thinking and the collaborative roles in the CS community, every participant should be actively engaged by reporting complaints about improper behaviours from peers (Figure D-5). Although these processes address goals to improve expertise discovery, they are not core to the expertise discovery lifecycle, remaining beyond the scope of DECYSE.

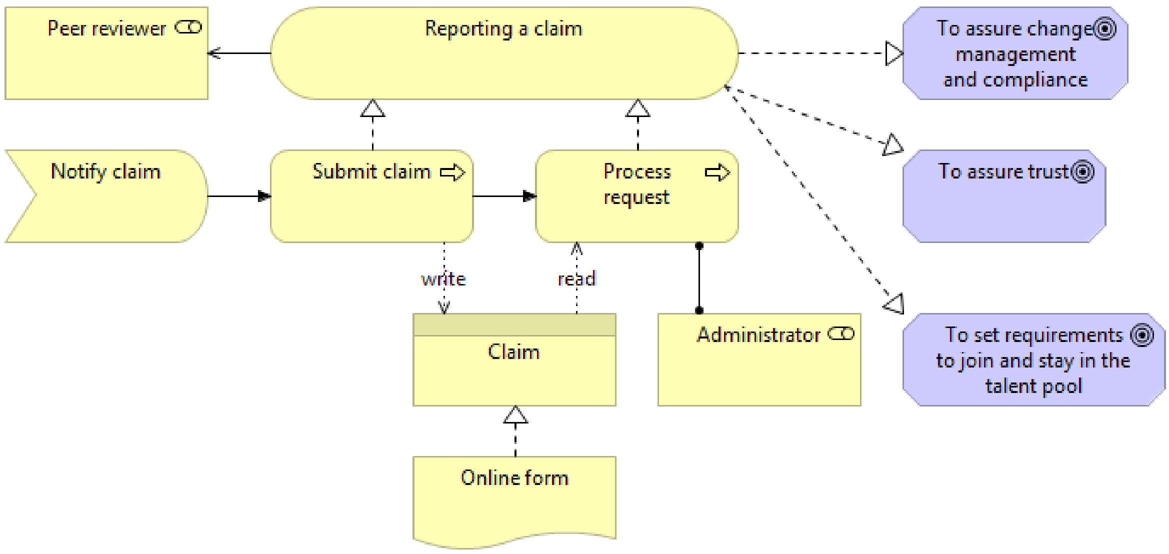


Figure D-5: Reporting problems

"Talent as mastery" (cf. Figure 2.3) is one of the ways in creating new talents, which has been (successfully or not) addressed in CS initiatives. Since the focus of this thesis is "talent as fit", developing capabilities by applying for courses has been reduced to a set of feedbacks delivered to agents according to their roles in the DECYSE method. However, those mapped concepts (Figure D-6) can serve as valuable source to structure course application or even promote personalised information provision in future work (as described in Section 8.2).

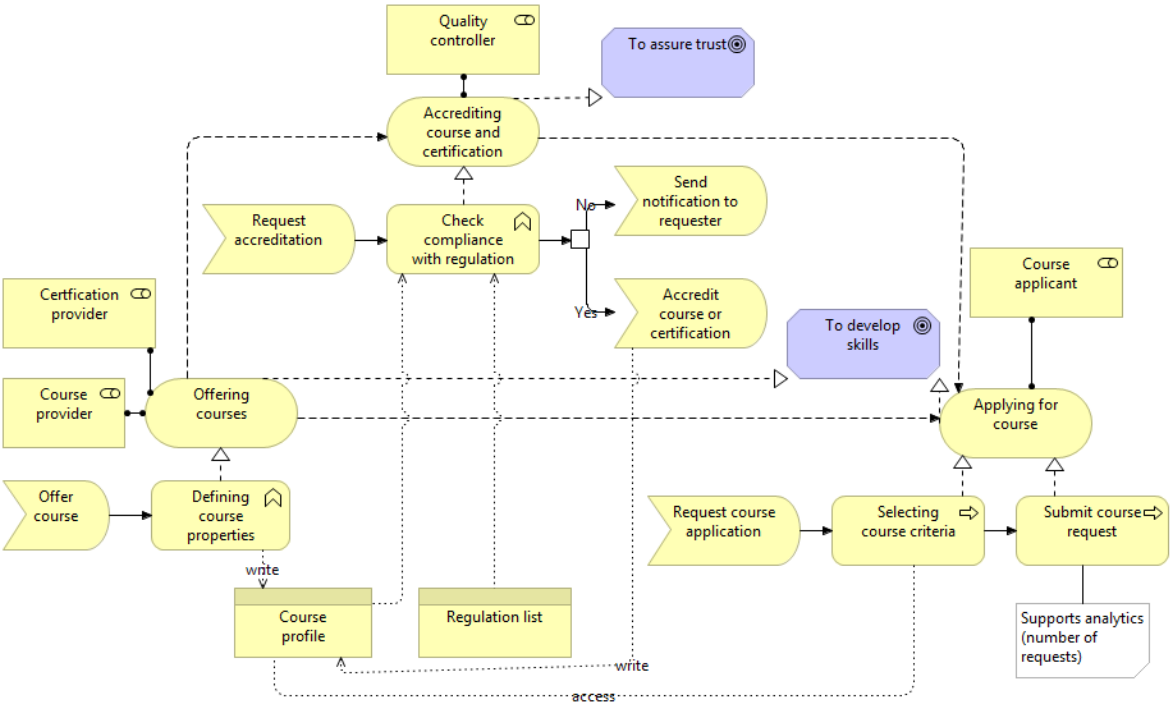


Figure D-6: Offering, accrediting and applying for a course

The DECYSE method postulates that the seven principles within the expertise lifecycle are applied to all actors involved (cf. Section 4.2.4). The reciprocal evaluation of course providers and course applicants (Figure D-7), however, was deemed beyond the scope of expertise discovery in terms of quality of service and performance in courses, respectively. This approach can be useful when focusing on structuring processes in the context of "talent as mastery".

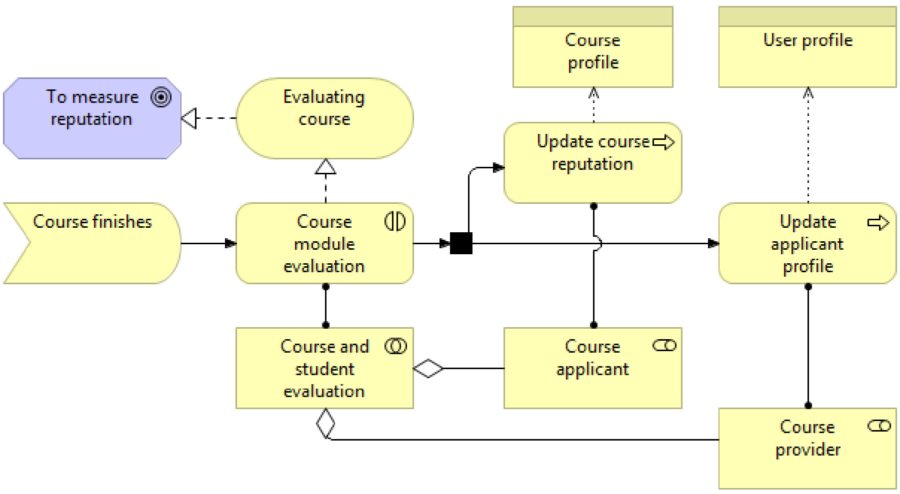
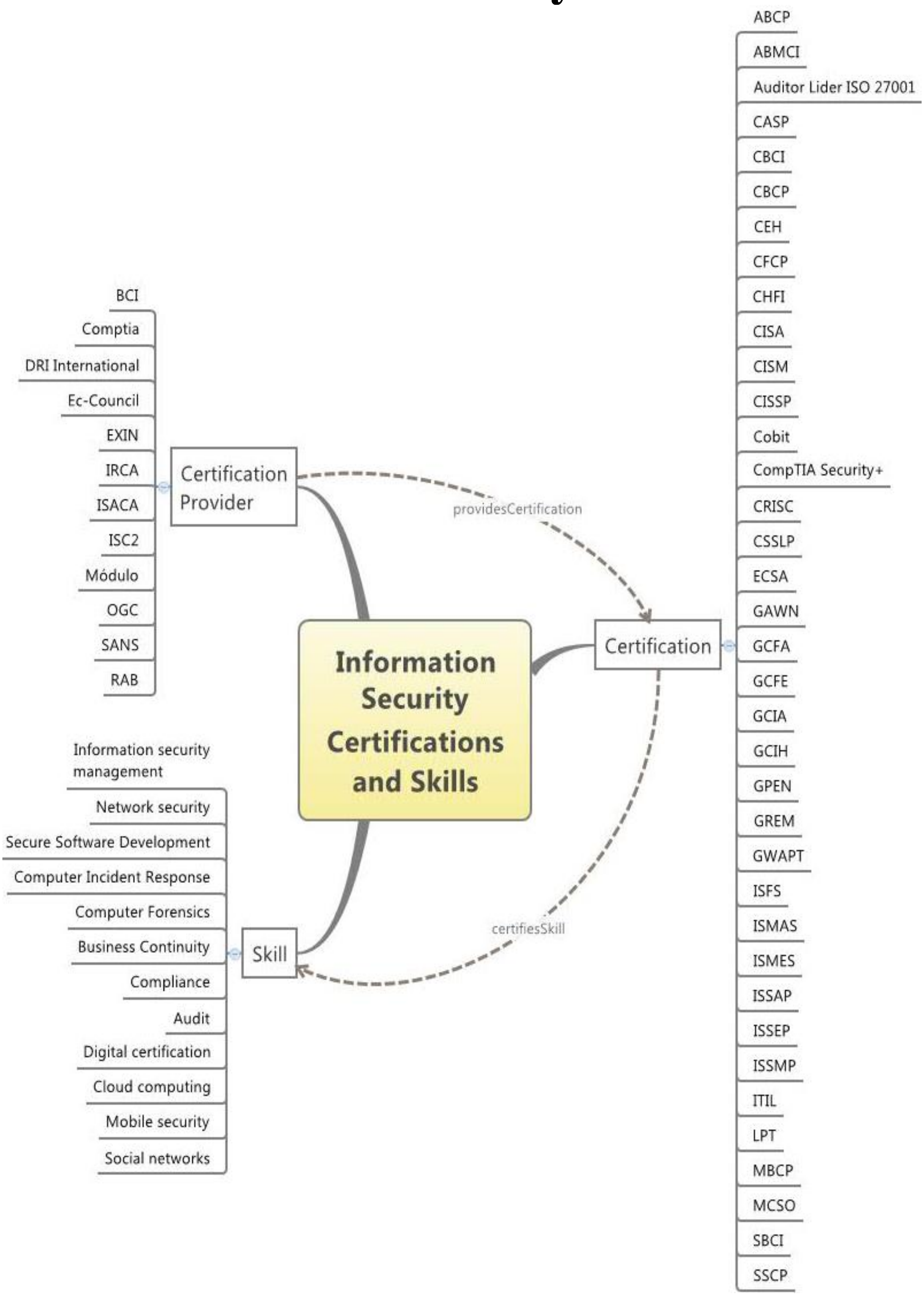


Figure D-7: Evaluating course and course applicant

Appendix E The Information Security Certifications Taxonomy



(adapted from Brazilian Government, 2013a)

Appendix F The Pairwise Comparison Example

	Extremely important	Highly important	Very important	Important	Equal	Important	Very important	Highly important	Extremely important	
Competence	○	○	○	○	○	○	○	○	○	**Certification**
Competence	○	○	○	○	○	○	○	○	○	**Academic area**
Competence level	○	○	○	○	○	○	○	○	○	**Academic degree**
Competence level	○	○	○	○	○	○	○	○	○	**Peer reputation**

Appendix G The Questionnaire for Designing the Context

1. In what country were you born?
2. What is your working area?
3. Which of the following subjects is within your knowledge background? (Please include others that apply).

   (…)

4. Does your agency have a talent identification and retaining system?
5. The following questions involve factors that in your opinion CONTRIBUTE or PREVENT attracting and retaining talents. Select in a scale from 1 (not relevant) to 5 (extremely relevant).
   a. What is the relevance degree of OPENNESS OF INFORMATION ABOUT CORPORATE GOALS, OUTCOMES AND INTENTIONS towards attracting and retaining talents?
   b. What is the relevance degree of an EFFICIENT HR PLANNING towards attracting and retaining talents?
   c. What is the relevance degree of RECOGNITION AND REWARDS towards attracting and retaining talents?
   d. What is the relevance degree of EQUAL OPPORTUNITIES towards attracting and retaining talents?
   e. What is the relevance degree of PARTICIPATION AND TEAMWORK towards attracting and retaining talents?
   f. What is the relevance degree of AUTONOMY AND DECENTRALISATION OF DECISION-MAKING towards attracting and retaining talents?
   g. What is the relevance degree of OPPORTUNITIES FOR SKILL DEVELOPMENT AND COURSES (ACADEMIC OR NOT) towards attracting and retaining talents?
   h. What is the relevance degree of CAREER DEVELOPMENT AND PROMOTIONS towards attracting and retaining talents?
   i. What is the relevance degree of RECRUITMENT FOR SPECIFIC TASKS towards attracting and retaining talents?
   j. What is the relevance degree of LEADERSHIP OF THE COORDINATING AGENCY towards attracting and retaining talents?
   k. What is the relevance degree of PAYMENT towards attracting and retaining talents?
   l. What is the relevance degree of FEEDBACKS towards attracting and retaining talents?
   m. What is the relevance degree of A GOVERNMENTAL AGENCY COORDINATING CYBER DEFENCE towards attracting and retaining

talents?

6. What would be other relevant factors to attract and retain talents in Cyber Defence context?
7. What factors do you consider that PREVENT (i.e. is a NEGATIVE contribution towards) attracting talents?
8. What factors do you consider that PREVENT (i.e. is a NEGATIVE contribution towards) retaining talents?
9. What skills or competencies should be required for acting in a cyber security context?
10. Please write some other issues regarding talent management in cyber security not covered in this questionnaire.

Appendix H The Questionnaire for the Requirements Analysis

1. In what country were you born?
2. In which country do you work?
3. What is your main working area?
4. Do you work or intend to work with cyber security?

   ( ) Yes ( ) No Others:______________

5. Is cyber security career well defined?

   ( ) Yes ( ) No Others:______________

6. Do you have access or knowledge about any solution on a national level aiming to manage talents for cyber security?

   ( ) No (and alternative solutions did not produce satisfactory results)

   ( ) Partially (there are initiatives that are not integrated)

   ( ) Yes (there is an integrated solution already deployed, which is capable to perform such management efficiently)

   Others:______________

7. Integrated talent management is crucial to improve national cyber security.

   ( ) strongly disagree ( ) disagree ( ) agree ( ) strongly agree

8. Documents and scientific production concerning cyber security are scattered or decentralised.

   ( ) strongly disagree ( ) disagree ( ) agree ( ) strongly agree

9. Integration is key to increase cyber security awareness within the population.

   ( ) strongly disagree ( ) disagree ( ) agree ( ) strongly agree

10. Integration is key to provide visibility and talent identification for the cyber security field.

    ( ) strongly disagree ( ) disagree ( ) agree ( ) strongly agree

11. Integration is key to provide job opportunities for the cyber security field.

    ( ) strongly disagree ( ) disagree ( ) agree ( ) strongly agree

12. Who should be the stakeholders for cyber security?

    (…)

13. Which should be the roles that stakeholders could play in terms of talent management for cyber security?

    (…)

14. What should be the services provided by the government to improve talent management for cyber security?

(…)

15. What should be the services provided by individuals to improve talent management for cyber security?

    (…)

16. Is there any other service provided by other stakeholders that should be made available?
17. Searching in open sources (e.g. curriculum databases and social media) contribute to find talents for cyber security.

    ( ) strongly disagree ( ) disagree ( ) agree ( ) strongly agree

18. Which information sources do you use or suggest to find people with key cyber security knowledge? (select all those options that apply and provide other examples)

    (…)

19. Do you have any additional concern or suggestion for talent sources that was not previously addressed?
20. Do you know any official program content to develop the cyber security professional? (please indicate in case you know it)
21. Which would be the types of criteria required to select individuals within a national scale to work with cyber security? (you may select more than one option and suggest others)

    (…)

22. Which would be the best decision-making mechanism to support talent selection for cyber security?

    (…)

23. Should curricula achievements be valued in terms of its recency?
24. Which should be other questions that you would suggest with regard to talent management for cyber security? (The purpose of this question is to support further questionnaires).
25. Use this space for additional comments with regard to talent management for cyber security.

Appendix I The Questions Used in the Semi-Structured Interview

1. What is your main working area?

   (…)

2. What is the level of your current function regarding cyber security?

   (…)

3. What is your agency / institution? (Optional)

4. What would be the possible <u>qualitative criteria</u> you would use for contracting individuals to work with cyber security?

   (…)

5. What would be the possible <u>quantitative criteria</u> you would use for contracting individuals to work with cyber security?

   (…)

6. Please indicate other relevant criteria that was previously mentioned (also indicate if the criterion is quantitative or qualitative).

7. Who would be the stakeholders for selecting talents for cyber security based on the criteria used in the previous questions?

8. Do you have access or knowledge about any list of recommended abilities or competencies aimed at the cyber security professional? Please indicate the source, in case you have any suggestions.

9. Do you have any additional comments on requirements and restrictions to select talents for cyber security?

Appendix J The Questionnaire for Experimenting DECYSE

1. Indicate the areas and levels of your academic background

	BSc	Specialisation	MSc	PhD
Computer science	( )	( )	( )	( )
Information science	( )	( )	( )	( )
Defence	( )	( )	( )	( )
Electric engineering	( )	( )	( )	( )
Mathematics	( )	( )	( )	( )
Microelectronics	( )	( )	( )	( )
Robotics, mechatronics and automation	( )	( )	( )	( )
Cyber security	( )	( )	( )	( )

2. Please indicate your other academic background regarded as relevant for cyber security (i.e. area and degree)
3. Select the competencies you believe to possess, among the following:

   (…)

   Others: ____________________________
4. Indicate the certifications that you currently possess among the following:

   (…)

   Others: ____________________________
5. Indicate your availability to join cyber security projects

   ( ) available ( ) not available (temporarily) ( ) not available (permanently)

   In case you are temporarily unavailable, please indicate the number of days in which you should become available:_______________
6. Do you authorise additional data to be collected from social media? If so, please indicate the URL to your social media webpage.
7. Have you ever defined capability procurement for cyber security projects? If so, please indicate the criteria based on the previous values or others that are relevant for cyber security. Values may be used for a single individual or to build a team.

Appendix K The Questionnaire for Expert Appraisal

1. What is your main working area? (select all that apply)
   ( ) Government - public agency
   ( ) Government - armed forces
   ( ) Academic
   ( ) Private sector
   ( ) Research centre
2. What is the level of your current function regarding cyber security?
   ( ) Decision maker level
   ( ) Management level
   ( ) Technical-operational level
   ( ) Support activities (indirectly related)
3. What is your agency / institution? (Optional)

Approach evaluation

Regarding the presented approach, please evaluate each criterion according to the following scale: 1 - not satisfied, 2 - poorly satisfied, 3 - satisfied, 4 - very satisfied, 5 - highly satisfied

<u>Usability</u>

4. Applicability of the proposal into the scope of managing talents for cyber security
   ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5
   In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____
5. Flexibility to express preferences for talent discovery
   ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5
   In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____
6. The proposed approach enables to expand to comprise further talent practices
   ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5
   In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____

<u>Cost-effectiveness</u>

7. The proposed approach enables to perform talent discovery with low cost in regards to what it proposes

   ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5

   In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____

8. The proposed approach enables to efficiently perform talent discovery within the proposed conditions

   ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5

   In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____

9. The approach enables to use existing physical and data infrastructure

   ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5

   In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____

<u>Acceptance</u>

10. Your control and freedom in the decision process

    ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5

    In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____

11. The approach takes into account the criteria drawn from stakeholders within the proposed scope

    ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5

    In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____

12. The proposed approach contributes with the discovery of talents for cyber security

    ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5

    In case you have evaluated the previous criterion with 1 or 2 values, please point out ways to improve it: ____

13. Use this space for additional comments.