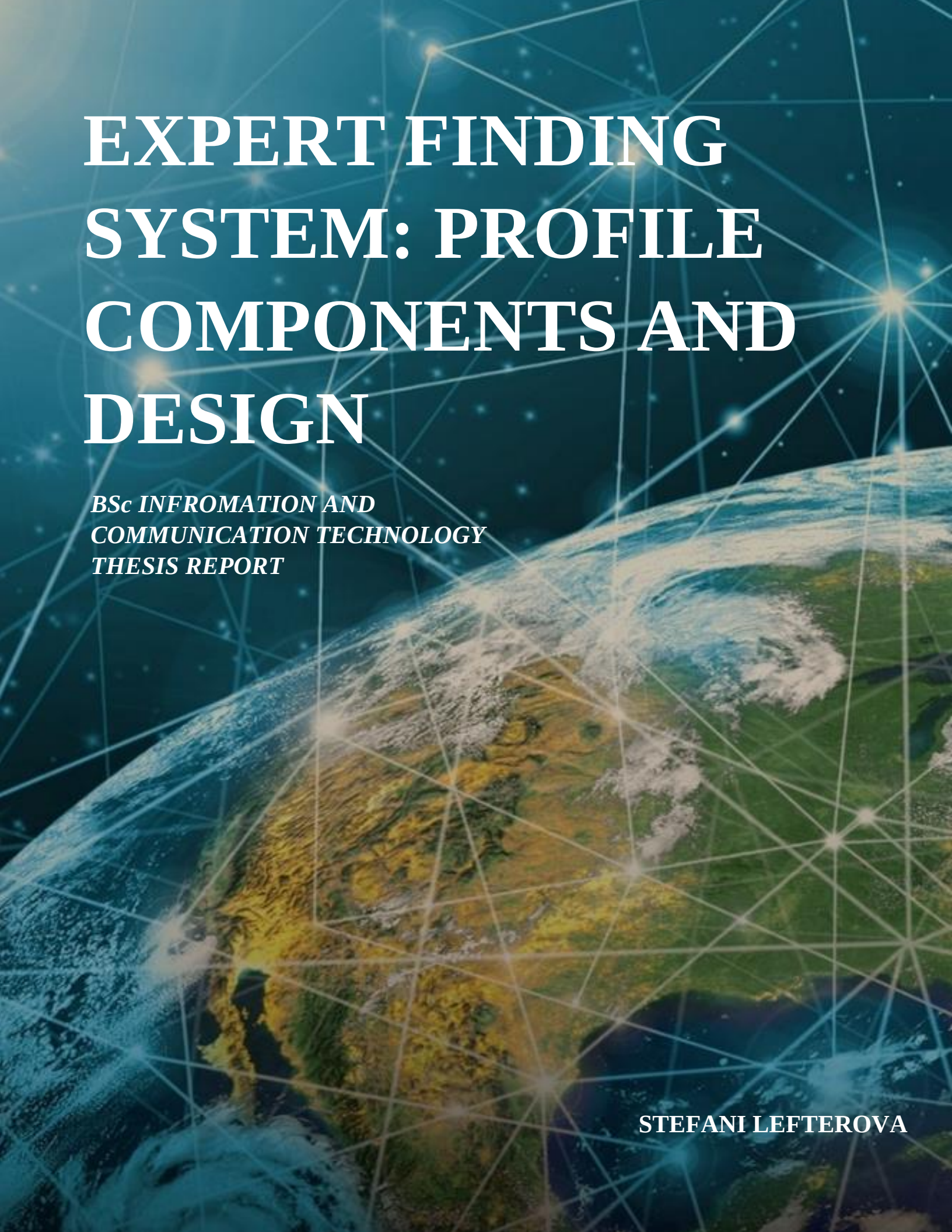


EXPERT FINDING SYSTEM: PROFILE COMPONENTS AND DESIGN

BSc INFROMATION AND COMMUNICATION TECHNOLOGY THESIS REPORT

STEFANI LEFTEROVA

Expert finding system: profile components and design

Author

Name	Stefani Lefterova
SNR	426839
Email	426839@student.saxion.nl
University	Saxion University of Applied Sciences
Faculty	Academy of Creative Technology
Study	BSc Information and Communication technology
Company	Thales Group
Company location	Hengelo, the Netherlands

Thesis supervision

Supervisor university	Ton de Bruyn
Supervisor company	Dr. C.H.M. Nieuwenhuis

Submission

Submission location	Saxion University of Applied Sciences
Submission date	June 17th, 2019
Location	Enschede, the Netherlands

Table of Contents

List of Tables

List of Figures

Preface

Hereby I represent my final thesis report “Expert finding system: profile design and components”. This report has been written as a final deliverable to complete my graduation project in the Information and communication technology program of Saxion University of Applied Sciences. The project was realized in the IT division of Thales Group, located in Hengelo, the Netherlands, and lasted five months as of February till July 2019.

From the start, this project involved a lot of meetings with various employees. This made me feel welcomed and integrated, as everyone was very willing to talk with me. This gave me the professional experience and help to establish a network of connections within the company, which was one of my personal goals in this project. With the supervision and guide from my supervisors Dr. C.H.M. Nieuwenhuis and Ton de Bruyn, I managed to communicate my ideas and brainstorm about the best possible approaches for the study. In addition, during the weekly meetings with Dr. C.H.M. Nieuwenhuis, I gained a lot of knowledge about the company and the corporate culture and mentality. Also with the contact network I established, I explored the company outside the scope of my assignment and that gave me a lot of valuable professional insight.

Therefore, I would like to express my utmost gratitude for the willingness and motivation of my supervisors to guide and support me throughout the project. Furthermore, I am grateful to both Saxion and Thales for giving me the opportunity to gain such an amazing experience. Additionally, I would like to thank my university supervisor Ton de Bruyn for taking the time to provide me with valid feedback and looking out for my best interest.

In addition, I would like to express my utmost appreciation to my parents who always support and believe in me.

To sum up, in this project I have learned and discovered many valuable things that will bring me further in both professional and personal matters. Therefore, with this report, I hope you, the reader, will find as thought-provoking and useful information as I did.

Stefani Lefterova

June 17th, 2019

Executive summary

Expert finding systems (EFS) is a method of finding experts by typing in a search query that results in multiple expert profiles. An expert profile is a profile of a user that has certain knowledge and expertise. After a matching process between the search query and the available profiles in the system, several profiles are suggested that would be able to answer the query. With this technique, the person – to – person communication is motivated rather than a person – to – machine. The possible benefits of an EFS in the corporate environment is a decrease in cost and increase of know-how of employees aka flat knowledge base. Some additional benefits are the establishments of a knowledge network between employees and an open work environment that would boost productivity.

With the current rapid technological developments, expert finding systems have gained more and more popularity in the world of the industries. Till now, expert profiles were mostly applied in the academics field, where one can use his/her education background and publications as evidence of your expertise. However, when it comes to defining the knowledge and expertise in the industries, a challenge emerges. This is because of several reasons: (1) products are not owned by one single person but by corporates, (2) there is not one clear definition of knowledge and expertise, and (3) there is not one determined method to capture and measure expertise and knowledge.

To tackle this, this report deals with the capturing of knowledge and expertise of software engineers into an expert profile. Throughout the research, it became evident that expertise is a very perceptual matter and there are three ways to perceive expertise: technical expertise, soft skills expertise and a combination of technical and soft skills expertise - hybrid. To define the right structure for the profile the software engineering body of knowledge (SWEBOK) and the content of the curriculum for Computer science program was used (Saxion University of Applied Sciences & UTwente, 2019). This resulted in several sections and parts of the profile such as Software systems, Network systems, Operating systems and etc.

In addition, the profile contains a section for proficiency level on the topics mentioned in it. The different levels to choose from are basic, limited working proficiency, intermediate, full professional proficiency. There is a section that addresses the hobbies, interests, and communities

of the user. This an important section as one's hobby might be a "hidden gem" for someone else's problem.

There is more research needed into the possible reliable sources of knowledge and expertise. Also, a test run of the system is essential, because the user engagement will give the most valuable insight.

Additionally, several profile designs need to be developed and evaluated by means of case studies. It is necessary to involve people as much as possible in the future project because the system is highly dependent on people. Also, this is a way that the news of a new expert finding tool will spread across the workplace, which is also very important for the success of an EFS.

1. Introduction

1.1 Problem analysis

Knowledge management still is a weak spot in large technology-driven companies, such as Thales. A few decades ago, the problem was perceived to be manageable because the product cycles were widely spaced in time and thus there was enough time to educate and train on the job and at the same time look around in the world.

Today, the situation has changed. Technological development emerges at a rapid pace and it is often challenging for big companies to keep up with each and every one. Another problem with large international companies is that technology is applied and developed in multiple entities and for a variety of product lines and markets. Because of that no transverse organization and simplistic 'paper shuffling – based' budgeting system can compensate the lack of communication and dissemination over large (physical, social and organizational) distances and barriers (Riege, 2005). Therefore, without any efficient way to communicate effectively in this rapidly changing world of IT, companies are finding challenges. Thus this problem needs to be dealt with and one way of doing that is engaging in an Expert Finding System (EFS). Such a management technique can not only break down the communication barrier but it has other positive results that will be discussed further in the report (Balog & Rijke, 2007).

In literature, EFS is one of the ways of improving the flow of knowledge in a company. However, such a system depends on the registration and parsing of so-called expert profiles of employees, which need to be complete, machine-searchable, maintainable and usable.

The advances in information technology motivate many organizations to place more and more emphasis on capitalizing the increasing mass of knowledge that is accumulated during the course of the business. However, attempting to gather all that knowledge into one server in the "Wiki-style" is inefficient and often leads to failure (Thales, 2018).

Therefore, in order to be able to fully exploit the accumulation of knowledge in a corporation, it is needed to not only have access to documented knowledge but to a tacit knowledge held by individuals. A knowledge management solution that is able to present the valuable information

contained in both documents and in people's heads. This is the so-called expert seeking that is prompted by the need for information and expertise (Yimam-Seid & Kobsa, 2009).

Nowadays, there are two main drives for seeking an expert: searching for a source of information and searching for someone that can perform a certain function or task. When it comes to people searching for experts as a source of information there are several types of needs that they could be seeking to be fulfilled with this (Rus, Lindvall, & Sinha, 2001):

A) Need to access undocumented information.
   As people, we naturally understand that receiving information through a textual format does not necessarily provide us with the full package of information that we have required. As detailed as a document could be, there are certain experiences and knowledge that can only be transferred via a person – to – person interaction. In some cases, information is not being documented or published deliberately because of political, social or economic reasons.

B) Need for a specification.
   Every so often when people do research they go through a phase of exploration. This is a stage in which the person tries to formulate the problem into a specific and clear search statement. Naturally, this process is accompanied by feelings of frustration and stress. This is because people often cannot specify their questions so that their search query will display the right answer. Therefore, users seek experts that are knowledgeable in the problem topic, which could help in specifying the problem and provide guidance on how to formulate the right search statement.

C) Leverage from others' expertise.
   Users try to minimize their effort and time to find a specific piece of information. Therefore, instead of spending a lot of time going through huge amounts of data to find one single paragraph of interest, it is useful to find an expert in the field that can filter the information for you.

D) Need for interpretation.
   Sometimes information needs to be interpreted for it to be useful and applicable to the given situation. However, users often are not able to accurately read between the lines or

even understand what the documents say. Consequently, this is where experts are called upon to decipher the information.

E) Need for socialization.
People are social beings. As a result, some would prefer the social interaction that is involved in seeking an expert's opinion rather than the person – to – machine interaction with computers.

The other drive for expert finding is seeking someone that has the knowledge and expertise to perform a specific function or task. Some cases that represent this drive are:

A) Searching for a consultant.
B) Searching for a collaborator for a project.
C) Searching for a speaker, a presenter, a researcher or other for media representation purposes.

Nevertheless, whatever motives expert seekers have, there is a need for a platform that contains a range of people with their knowledge and expertise. This calls for the development of an expert database aka knowledge directories, in which expertise data is entered manually or automatically. Some organizations already took this initiative such as Hewlett-Packard's CONNEX knowledge management system (Becerra-Fernandez & Sabherwal, 2014) or SkillView, which is very common in human resource management domains (Centro Universitario Internazionale, 2018).

Nevertheless, there are several shortcomings that come with the development and maintenance of a knowledge directory.

1) In the case that the expertise information is entered manually, that requires extra time that comes at a cost.
2) The database would depend on the experts' willingness to spare the extra time to provide a detailed description of their skills and expertise.
3) After completing the profile, the expert is responsible for keeping it up-to-date according to his/ her latest expertise and skills. As these change constantly and updating the profile is time-consuming, most often the profiles become outdated.
4) The expert profiles are very general because it is very difficult to determine the best way to identify one's specific area of expertise. This is explicitly a challenge for the industries.

In the academic field, it is simpler to determine the expertise of an individual because the academic topics and subtopics are clearly defined.

However, when it comes to expert search in the industrial world only a few of the above-discussed needs and drives apply.

The fundamental aspect that causes this phenomenon is the cost-benefit relationship of voluntarily investing time and effort to deal with another person's problem.

In the industries, employees have been hired in their respective positions because they have certain knowledge and capabilities. In the event of a lack of that knowledge and skill, they become a bottleneck for the business which results in being removed from the company. Therefore, employees need to have a certain level of expertise to be able to do their jobs. Thus the previously mentioned motives for expert search in points A – E could not directly apply in that type of workplace because that would be evidence of being unqualified. It should be noted that the term "not directly" is used to describe the link to the motives for expert search.

Comparing to the academic field, the industrial workplace has a different mentality and culture. This is why motives A – E need to be reformulated to apply to the industries.

To do that there are several factors to be considered:

1. No one will openly admit their lack of knowledge in an area.
2. Without a personal gain or a monetary incentive, there aren’t many volunteers that would add extra work hours to their day to give assistance on an issue. This would not only be an additional expense but it would add a feeling of annoyance and distress between the expert and the user.
3. In the case of lack of knowledge and no assistance provided, there is a high risk of wrongly interpreted data, which leads to more person-hours invested. Thus, an increase in costs.

Therefore, there are two main drives for expert search in the industries (Employee 5., 2019):

A) The need to decrease cost.

In the case that an employee is not fully familiar with a topic, they often take it upon themselves to fill in the missing information. This is because in the industries knowledge is power, and people use it to move up in the organizational hierarchy. However, while

investing the extra work hours in figuring out what is not known, the person can go in the wrong direction or get lost in all the new information. This results in a lot of unnecessary time spent on the wrong subject with an additional feeling of stress and dissatisfaction. Therefore, an EFS can give the opportunity to find someone who has more experience in the subject, who the employee can consult with on how to get the right information. Consequently, he\she will be able to conduct more efficient research in an optimal time span.

B) The need to increase the know-how of employees.
On a regular basis, employees deal with a project on the same or similar topics. If initially, the project topic was unfamiliar, the employee has asked an expert where to find information. Thus, when encountered with the same or similar topic again, the employee would already know how to approach the subject and has the ability to give guidance to someone else if needed. Consequently, this gradually increases the skill set of employees across a company. This results in a flat knowledge base for all employees that enhances their capabilities and saves costs for the business.

All in all, to find the right expert, his/her knowledge and expertise areas have to be presented in a clear and understandable way. This is where the concept of the expert profile comes in. In the academics field, such profiles are encountered often and can be easily accessed from the website of the academic institution.

The academic expert profiles provide an overview of the fields and topics the person has knowledge of. Additionally, they can give information about the related works of the academic, current and previous occupation, and contact information.

However, when it comes to the world of the industries it is not so clear what the profile needs to be composed of and how to present the relevant knowledge and expertise areas of the employee. This issue arises from the fact that once employed people start to develop a certain set of skills and competencies. Thus, some of the topics they have learned during their academic times fade away and some are strengthened. It is unreliable to consider the academic background as a valid source of expertise and knowledge for an industrial expert profile. Hence, comes the problem of how to define that and what is its source. There are multiple questions that arise from this issue and are discussed in the next subchapter.

1.2 Research questions

The aim of this research is to find a solution for one of the critical elements of an EFS, namely the way in which the knowledge of a person can be recorded (and thus structured and captured in a machine-readable way). In addition to the way the knowledge of a person can initially be defined, it needs to be presented in a standard way that is clear for all employees regardless of their background.

The assignment comes from Thales who is the client in this project. Thales realizes that the number of knowledge areas that are relevant for Thales is extensive. It is therefore allowed to restrict the research to the area of software engineering and to construct a sample profile design for an expert in that field – software engineer.

To sum it up, the objective is to establish requirements for a Meta profile that is recognized by employees of Thales Group and other partner industries.

This lead to the formulation of the main research question that is:

In what way should a knowledge profile be arranged to be useful for employees in the industry, who have a question, to be able to understand if that knowledge could be of help to them?

To answer the main research question, several sub-questions were developed. This tactic provided a more detailed and structured approach to the research which delivered better insight into the topic:

1. How to classify knowledge areas and skills?
2. What definition to use to accurately describe expertise?
3. How to make an expertise description that is recognized across companies?
4. What is a source of expertise evidence?

1.3 Research methodology

For this project, qualitative data research is applied as the main methods were interviews, surveys and documentary analysis. This helped for a thorough examination of the expert profile phenomena. This chapter presents an overview of the methods applied in the research. For more detailed information please refer to chapter 3 “Methodology”.

Interviews and surveys

The main goal of an EFS is to motivate knowledge sharing between people and help to establish and efficient communication network. For that reason, an expert profile needs to be clear and readable for everyone. Therefore, it is very important to understand the perception of employees regarding a “good” expertise description. With this insight, the answer to subquestions 1 to 4 will be given.

For starters, it was needed to gather information about how where job descriptions and competencies are defined. It was useful to get to know what makes the "perfect persona" according to the human resources department. A “perfect persona” is the person that holds the characteristics that add the most value to a project. Once made, the profile is considered to be a general description of the best fitting employee. Thus, it was considered a good starting point for understanding how expertise and knowledge are described. Understanding the know-how by which HRM operates gave valuable insight into the possible components of a reliable expert profile. In addition, interviews were initiated with non – Thales employees because one of the goals for the expert profiles is to make them generally applicable. Thus, narrowing the research field only to the internal environment of the client would be counterproductive. In the case of this project, the more insight was gathered from different companies the better. Due to that, interviews with academics in HRM were included because they are able to provide guidelines for how the job codes and competencies are created. This information was used as a starting point for the components of an expert profile.

Further insights were taken from interviews with numerous software engineers. As the report focuses specifically on this profession it was crucial to discuss the components of an expert profile with experts in the field of software engineering.

After conducting several interviews, a survey was created to widen the scope of participation. The results from the interviews were the base for the creation of the survey questions.

The implementation of interviews and surveys as a research and data gathering method in this project was chosen due to several reasons. Interviews deliver the personal opinion of the interviewee and give the opportunity for sudden questions that would provide even more information. This input is extremely valuable for this project because personal perception plays an important role in the overall acceptance of the expert profile. Also, it was important to reach out to a specific amount of people to gain enough data for a reliable conclusion. It is known that surveys are an interviewing method that has a wide reach. Therefore, a questionnaire was created to reach out to more software engineers. This method lacks the personal approach but after a certain amount of F2F interviews, a level of understanding of the necessary information was gained and the interviews reduced their added value to the research.

Documentary analysis

Every project starts with primary desk research. For this particular assignment, a primary documentary analysis was of utmost necessity because the research topic was unfamiliar to the researcher. Thus, it was mandatory to gain a full understanding of the topic to conduct efficient research that will give a sufficient result.

To begin with, literature regarding the topic of Expert Finding Systems (EFS) was provided by the previous work of the student who started this assignment. His desk research was the starting point of the documentary analysis for this report.

Nevertheless, individual analysis of the available articles online was conducted. As establishing a definition of EFS and an expert profile is part of the thesis report, it is vital to research this concept independently to gain a full understanding. However, to avoid repetition and overlap with the previous student, the main focus was on the expert profiles and not on the EFS definitions. The theoretical information necessary to understand the results and conclusions is presented in chapter 2 “Theoretical concepts”.

1.4 Reading guide

The report starts with an introductory chapter that describes in detail the problem and research questions. It also presents an overview of the research methodology used in this project.

In chapter 2, an overview of the theory applied in this report is presented. This review presents all the theoretical information needed to understand the main concepts of the report and the final product of the project. It is recommended that the reader goes through that chapter if the topic of the report is not familiar. Nevertheless, it is always useful to go over the theory before diving into the report to ensure proper understanding of the material.

Then, chapter 3 describes in detail the methodology of this research and discusses the research tools, methods, and approach. In addition, this chapter presents the reason for choosing certain methods in comparison to others.

Chapter 4 “Results and analysis” refers to all the outcomes of the research activities and their relevance to the project. The chapter aims to define and outline the knowledge and expertise for software engineers. In addition, the chapter presents the components and functional design of an expert profile for a software engineer. The last part of this chapter is an analysis subchapter that interpretation of the findings.

Chapter 5, summarizes all information presented thus far and provides a recommendation for future actions to ensure the correct development of an expert finding system. After, in chapter 6, the results and their overall quality and relevance are discussed.

2. Theoretical concepts

This chapter deals with the findings from the conducted review of the literature. It covers several concepts concerning Expert profiles. The information presented here is important to the reader, especially if he/she does not have any background information on the topic of EFS. It will help in understanding the thought and decision - making process described in this report. By getting acquainted with the topic of EFS, it will be easier to understand the concept of an industrial expert profile. To ease the load of reading, there are figures and tables that contain overviews of the information in each subchapter.

2.1 Knowledge and knowledge management

In this section, knowledge will be discussed from a general and software engineering point of view. This section will discuss how knowledge is managed and what the perception of knowledge sharing is in the IT community.

2.1.1 Knowledge

There is not one clear and conclusive definition of knowledge. Looking at a dictionary definition, knowledge is:

“The fact or condition of knowing something with familiarity gained through experience or association; acquaintance with or understanding of a science, art, or technique; the sum of what is known: the body of truth, information, and principles acquired by humankind.” (Merriam-Webster, 2019)

There is another common definition that describes knowledge as "a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices, and norms." (Davenport & Prusak, 1998).

Defining knowledge has been a topic of discussion for ages, but still, people have a general understanding of what knowledge is. The definition from Davenport and Prusak addresses some

very important aspects of knowledge. Knowledge is multi-layered and a mix of various concepts that are difficult to determine. Secondly, there is a strictly dependent owner – knowledge relationship as the two cannot exist as separate entities. The simplest example is if we look at teachers and students. A teacher has proficient knowledge and skills in a certain subject f.e. mathematics. However, after completing the course of mathematics, it cannot be expected that the students will be as proficient in it as the teacher.

In addition, there are different items related to knowledge such as data, information, and experience. Data is the discrete, objective facts about events. It is the raw material without any relevance or importance stated and it can be qualitative or quantitative. It serves as a base for creating information. Consequently, information is an organized set of data that is made relevant and useful for an end – user. Therefore, if we refer back to knowledge, it is the understanding of data and information and the relationships between them. Finally, there is experience, which is knowledge applied. It is the human factor in knowledge as it cannot be clearly defined or stored (Rus, Lindvall, & Sinha, 2001).

Conclusively, data, information, knowledge, and experience are interconnected terms. The more one dives deeper into characterizing them, the less clear it gets when one ends and another one starts. Following is a discussion about the types and classes of knowledge that will be used further in the report.

Knowledge is either tacit or explicit (Nonaka, 1995). According to this statement, tacit knowledge is impossible to document and it can only be transferred through human – to – human interaction. However, one - time contact does not suffice for a successful transfer from one individual to another. It is necessary to engage in prolonged periods of intensive contact and shared experiences. To further describe tacit knowledge, another statement adds that it is highly influenced by one's beliefs, perspectives, and values that are embedded in an individual. It also adds that awareness of knowledge or the lack of it is important (Agresti, 2000). Thus, it makes sense why it is necessary to invest time in communication when trying to acquire tacit knowledge.

On the other hand, explicit knowledge is the easily verbalized information, often captured in a written form. It corresponds to the information and skills that are easily communicated and

codified such as processes, templates and media data. Explicit knowledge is easier to exploit and reuse across organizations (Rus et al., 2001).

In organizations, however, specifically software engineering companies, there can be different types of knowledge:

- Organizational knowledge, which is being aware of how to run the company, its business objectives, human resources aspects, etc.
- Managerial knowledge, which relates to planning, staffing, tracking and leading a project
- Technical (engineering/development) knowledge refers to the software engineering body of knowledge which will be discussed in the next subchapter 2.1.2.
- Domain knowledge that relates to the specific product or system an employee is a part of.

An overview of knowledge characteristics is presented in Figure 1 below.

Figure 1 Types and elements of knowledge overview

2.1.2 Knowledge areas in software engineering

Fortunately, the IEEE Computer Society has come up with a textual guide that discusses the basic body of knowledge an employee in the field of software engineering should possess. This document is called SWEBOK: A guide to the software engineering body of knowledge (IEEE Computer Society , 2014).

In the following section, the various knowledge areas will be presented and described. What needs to be pointed out is that not all of the areas presented here would be directly applicable when creating knowledge areas for a profile of a software engineering expert. This will be discussed further in the report in chapter 4 and 5. Below each area of SWEBOK is described briefly. To see the additional topics that belong to each area please refer to Table 1.

Software requirements

As the name implies, this knowledge area deals with software requirements. This involves an understanding of the way of creating, analyzing, specifying and validating requirements. It involves knowledge about how to manage those requirements throughout a project. This type of information is very crucial in the software industry because it is widely acknowledged that poor requirements lead to poor project performance.

Software design

Design can be defined as both a process and a result. As a process, it is the "defining of the architecture, components, interfaces and other characteristics of a system or component". As a result, it is simply the outcome of the designing process.

In the field of software engineering, this area is about knowing how to analyze software requirements to produce a description of the interface that will serve as a base for construction. The result of this process is a software architecture with a detailed description of its components.

Software construction

This knowledge area refers to the detailed construction, verification, unit testing, integration testing and debugging of working software. In the construction process, there is a lot of design and testing involved. Thus, those knowledge areas are tightly related to software construction.

Software testing

This area consists of the dynamic verification that a program runs in acceptable behavior on a specific set of tests. The reason why the verification is considered dynamic is that the testing is done by executing the program in certain inputs.

Software maintenance

After the release of every product, there is a hyper – care stage. When it comes to software this means that once live, defects emerge, new user requirements might surface and the operating environment could change. Maintenance is needed to keep the software operating as long as possible and bring more return on investment.

Software configuration management

This discipline is defined by activities that aim to administrate and surveil any technical or administrative changes in a product, throughout its life cycle. This includes any hardware, firmware or software characteristics of the end product and their related versions. This knowledge area keeps track of the applied changes and activities, which are documented and make sure the compliance with specific requirements and standards is verified.

Software engineering management

As the name implies, this area is presented by the various activities involved in managing the software engineering process. These activities aim to ensure that software products and services are delivered in an efficient and effective way that is also beneficial to the stakeholder.

Software engineering process

Generally, engineering is concerned with interrelated activities that transform inputs into outputs while consuming resources to achieve the transformation. In software engineering, this is referred to as activities, performed by software engineers, to develop, maintain and operate the software.

Software engineering tools and methods

This area aims to give structure to the software engineering process by making it systematic, repeatable, and more success – oriented. The relevant models help in problem – solving and the

methods make sure the different stages of the product life cycle (specification, design, construction, test, verification) are approached correctly.

Software quality

This term has received various definitions throughout the years. The most recent one defines software quality (SQ) as the capability and degree to which a product is satisfying stated and implied needs under specific conditions. Quality is dependent upon requirements and software requirements are considered to be a constraint of functional requirements.

Software engineering professional practice

The knowledge, skills, and attitudes that engineers in the software industry should have to be able to practice the profession in a professional, responsible, and ethical manner.

Software engineering economics

The subject that teaches how to make decisions in the software engineering field in a business context. This relates to the software product, service, and solution, which depend on good business management.

Foundations of mathematics, computing, and engineering

Knowledge about the computer is essential for a software engineer as its principles serve as a framework for the field. Additionally, a mathematical background aids in understanding the logic of programming, which is then translated into programming language code. Also, an engineering foundation is essential because it teaches all engineers to apply a structured, systematic, disciplined, and quantifiable approach to machine, products, systems or processes.

Table 1 Overview of the software engineering body of knowledge (SWEBOK)

SOFTWARE REQUIREMENTS
• *SOFTWARE REQUIREMENTS FUNDAMENTALS* • *REQUIREMENTS PROCESS* • *REQUIREMENTS ELICITATION* • *REQUIREMENTS ANALYSIS* • *REQUIREMENTS SPECIFICATION* • *REQUIREMENTS VALIDATION*

- *PRACTICAL CONSIDERATIONS*
- *TOOLS*

SOFTWARE DESIGN

- *SOFTWARE DESIGN FUNDAMENTALS*
- *KEY ISSUES IN SOFTWARE DESIGN*
- *SOFTWARE STRUCTURE AND ARCHITECTURE*
- *USER INTERFACE DESIGN*
- *SOFTWARE DESIGN QUALITY ANALYSIS AND EVALUATION*
- *SOFTWARE DESIGN NOTATIONS*
- *SOFTWARE DESIGN STRATEGIES AND METHODS*
- *TOOLS*

SOFTWARE CONSTRUCTION

- *SOFTWARE CONSTRUCTION FUNDAMENTALS*
- *MANAGING CONSTRUCTION*
- *PRACTICAL CONSIDERATIONS SUCH AS:*
  - *CONSTRUCTION DESIGN*
  - *LANGUAGES*
  - *CODING*
  - *REUSE*
  - *QUALITY*
  - *INTEGRATION*
- *CONSTRUCTION TECHNOLOGIES*
- *TOOLS*

SOFTWARE TESTING

- *SOFTWARE TESTING FUNDAMENTALS*
- *TEST LEVELS*
- *TEST TECHNIQUES*
- *TEST – RELATED MEASURES*
- *TEST PROCESS*
- *TOOLS*

SOFTWARE MAINTENANCE

- *SOFTWARE MAINTENANCE FUNDAMENTALS*
- *KEYS ISSUES IN MAINTENANCE*
- *MAINTENANCE PROCESS*
- *TECHNIQUES FOR MAINTENANCE*
- *TOOLS*

SOFTWARE CONFIGURATION MANAGEMENT

- *MANAGEMENT OF THE SCM PROCESS*
- *SOFTWARE CONFIGURATION IDENTIFICATION*
- *SOFTWARE CONFIGURATION CONTROL*
- *SOFTWARE CONFIGURATION STATUS ACCOUNTING*
- *SOFTWARE RELEASES MANAGEMENT AND DELIVERY*
- *SOFTWARE CONFIGURATION MANAGEMENT TOOLS*
- *TOOLS*

SOFTWARE ENGINEERING MANAGEMENT

- *INITIATION AND SCOPE DEFINITION*
- *SOFTWARE PROJECT PLANNING*
- *SOFTWARE PROJECT ENACTMENT*
- *REVIEW AND EVALUATION*
- *CLOSURE*
- *SOFTWARE ENGINEERING MEASUREMENT*
- *TOOLS*

SOFTWARE ENGINEERING PROCESS

- *SOFTWARE PROCESS DEFINITION*
- *SOFTWARE LIFE CYCLES*
- *SOFTWARE PROCESS ASSESSMENT AND IMPROVEMENT*
- *SOFTWARE MEASUREMENT*
- *TOOLS*

SOFTWARE ENGINEERING TOOLS AND METHODS

- *MODELING*
- *TYPES OF MODELS*
- *ANALYSIS OF MODEL*
- *METHODS*

SOFTWARE QUALITY

- *SOFTWARE QUALITY FUNDAMENTALS*
- *SQ MANAGEMENT PROCESSES*
- *PRACTICAL CONSIDERATIONS: SQ REQUIREMENTS, DEFECTS, SQ MANAGEMENT TECHNIQUES, SQ MEASUREMENT*
- *TOOLS*

SOFTWARE ENGINEERING PROFESSIONAL PRACTICE

- *PROFESSIONALISM*
- *GROUP DYNAMICS AND PSYCHOLOGY*

- *COMMUNICATION SKILLS*

SOFTWARE ENGINEERING ECONOMICS

- *FUNDAMENTALS: FINANCE, ACCOUNTING, CONTROLLING, TAXATION, DEPRECIATION, ETC.*
- *LIFE CYCLE ECONOMICS*
- *RISK AND UNCERTAINTY*
- *ECONOMIC ANALYSIS METHOD*
- *PRACTICAL CONSIDERATIONS*

FOUNDATIONS OF MATHEMATICS, COMPUTING, AND ENGINEERING

2.1.3 Knowledge management and systems

The concept of knowledge management and knowledge databases has been around for decades. However, not until recently the concept of EFS emerged. Unlike other goods, knowledge is enriched when being shared and it is not diminished by its use.

Ultimately, the individual is the one who is performing a task to achieve the organizational goals. Nevertheless, that individual is constantly learning and improving his/her skills and is working in teams of people where knowledge is shared to solve a problem or perform a task. Knowledge sharing is done in such a way that it resembles a knowledge spiral (Appendix 1), where knowledge is being transformed into information and then back to knowledge (Nonaka & Takeuchi, 1995). There are four means distinguished for that:

- *Socialization*: this helps to bring out and transfer the tacit knowledge that resides in the brains of employees, within the community.
- *Externalization*: is when the information and knowledge are captured in the means of conversation, written documents, figure, presentation or teaching.
- *Combination*: adding new knowledge to an already existing one. This happens whenever one's new experience or insight is combined with previous knowledge.
- *Internalization:* gaining an understanding of the acquired information and combining it with one's existing knowledge. This process transforms information into knowledge.

Following this insight, there are numerous knowledge management frameworks that have been developed. From an extensive study of 160 knowledge management frameworks, it was concluded, that even though the processes have different names, they all refer to similar activities

(Heisig, 2009). This comes from the fact that there is no one definition of knowledge and it is very much based on perception and values. And in this case organizational culture and values.

However, in the past, knowledge management had been focused only on codification and knowledge repositories. Yet, recent studies reported that the success of a knowledge management system depends on the technological factors, human, social and organizational nature of a company (Heisig, 2009). Since the establishment of these aspects, the new knowledge management systems have started to focus more on people rather than only on knowledge and its code.

2.1.4 Knowledge sharing

Implementing a new type of knowledge management system in any organization would be a challenge because of the time and effort required for implementation and the time needed for it to bring a return on investment.

However, when it comes to software engineers, this implementation might come easier due to the fact that all of the sources of information in that industry are already in a digital format. This makes them very easy for distribution and sharing. The most encouraging fact is that the sharing of knowledge is a daily practice with software engineers (Rus et al., 2001).

Software engineering (SE) involves a multitude of knowledge-intensive tasks:

- Analyzing user requirements for new software systems
- Identifying and applying best software development practices
- Collecting experience in project planning and risk management and etc.

SE is document – oriented and during a project, the following documents are produced:

- Contracts
- Project plans
- Requirements and design specifications
- Source code
- Test plans and related documents
- Decisions

In addition, there are numerous community question answering (CQA) forums and communities that have emerged for the sole purpose of sharing knowledge (Baltadzhieva, 2015):

1. The Maryland software industry consortium (SWIC)
2. The software experience consortium (SEC)
3. The software program managers network (SPMN)
4. The world wide web consortium (W3C)
5. Software process improvement network (SPIN)
6. Special interest groups of the IEEE or ACM

These CQA forums could be also considered as another example of the knowledge spiral. In such environments, people share their knowledge in a visual or textual way (Socialization & Externalization), which is then topped up with more knowledge (Combination) by other users. In the end, each user can make his/her own conclusion (Internalization).

2.2 Expert finding system

An expert finding system (EFS) is a computer system that accumulates all knowledge, documented and undocumented, in an organization or worldwide, for the sole purpose of finding an expert for a given problem. This problem can be of two natures: the need for information or need of expertise (Yimam-Seid & Kobsa, 2009).

The process of finding an expert is addressed in two main domains – enterprise and online community (Figure 2) (Wang, Jiao, Abrahams, Fan, & Zhang, 2013).

An enterprise domain is where the hierarchy of knowledge is well defined and the quality of the information is surely high. The aim of this type of domain is to have a clearly defined expertise description of each employee, which would help managers find the right people for a particular job or task. As a source of evidence of expertise self – disclosed information, documents, and social networks are used. There are some issues associated with this type of EF source. Most often the expertise of employees is not fully documented in the organization. Document-based information cannot be used to determine the level of influence the employee has in the organization and the relevant social network. Additionally, self – disclosed information is often outdated and sometimes biased.

An online community domain is the popular community question answering (CQA) platforms. Some of the most commonly known platforms are Quora (quora.com), StackOverflow (stackoverflow.com), and Yahoo! Answers (answers.yahoo.com). These are types of

crowdsourcing knowledge services. They make use of human knowledge to solve complex problems by asking and answering questions. However, as much as a user can leverage from the "wisdom of the crowd" and get answers from multiple people simultaneously, this type of web application does not guarantee good quality information.

Overall there are three main sources of expertise information for an EFS: Meta databases, document collections, and referral networks.

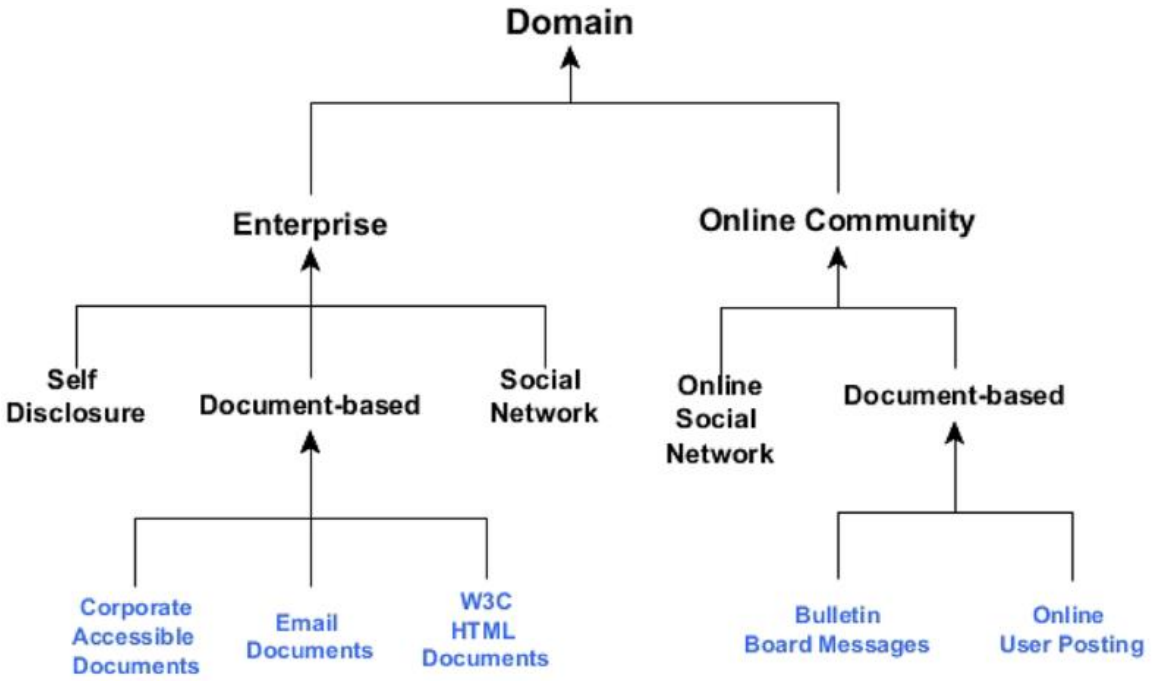


Figure 2 Domain classification of EF systems (Al-Taie, 2018)

2.2.1 Techniques

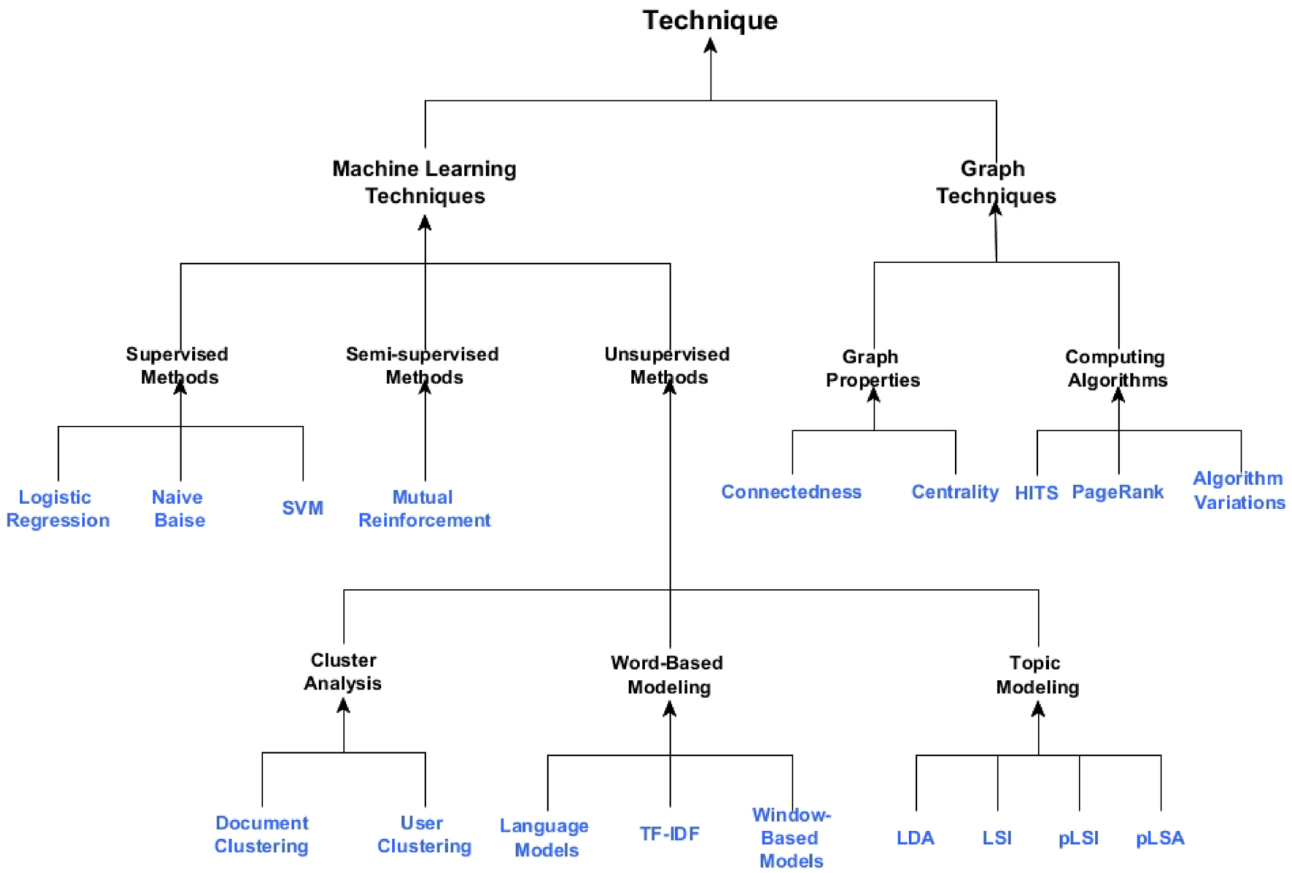
The methods for expertise retrieval usually fall into two search questions: “Who is the expert on topic X?” or “What does expert X know?”. In the first question, the main interest of the user is to find experts in a specified knowledge domain. For the second question, a user wants to figure out a specific expert's knowledge and his/her information. The most common query, however, is the first one, which is also the focus of most expert finding algorithms (Lin, Hong, Wang, & Li, 2017).

There are three types of expert finding methods: graph-based, machine learning and hybrid (Figure 3).

Graph-based models aka network – based are very common means for finding information about people. This is the type of technique social networks and referral webs are using. With this type of model, expertise retrieval can be done in two ways. The first way is to apply graph properties in the means of connectedness and centrality. This way takes documents and candidate experts as nodes and views their relationship as edges. You can imagine the nodes as entities and the edges as their relations. The second way uses algorithms such as HITS and PageRank. These algorithms view the candidate experts and documents as pages and the candidate - candidate or candidate – document associations are like a hyperlink on a web page on the Internet (Lin et al. 2017).

Figure 3 Techniques classification in EF systems (Al-Taie, 2018)

The machine – learning technique is presented by the generative probabilistic and voting models. Such models include linear regression, k – means cluster and ranking such as RR data fusion techniques (Al-Taie, 2018). Data fusion techniques usually apply minimum, maximum, median



and average of relevance scores to evaluate the score of a candidate expert when a query is submitted (Lin et al. 2017).

To be able to choose or combine (in the case of a hybrid) an algorithm there are three components to be considered: candidate, document and topic. A candidate is a person that might hold a certain level of expertise in a topic; a document is a resource that holds textual information related to the expert such as publications, reports, emails, web pages; and a topic is a specific domain (Lin et al. 2017).

2.2.2 Expertise profile

Knowledge profiling also is known as a candidate or expert profiling is the answer to the second most important search question in expertise retrieval - "What does expert X know?".

A knowledge profile is the profile of an individual X that shows an overview of records of the types and areas of skills and knowledge of that individual ("topical profile") with an additional description of his/her collaborative network ("social profile") (Balog & Rijke, 2007).

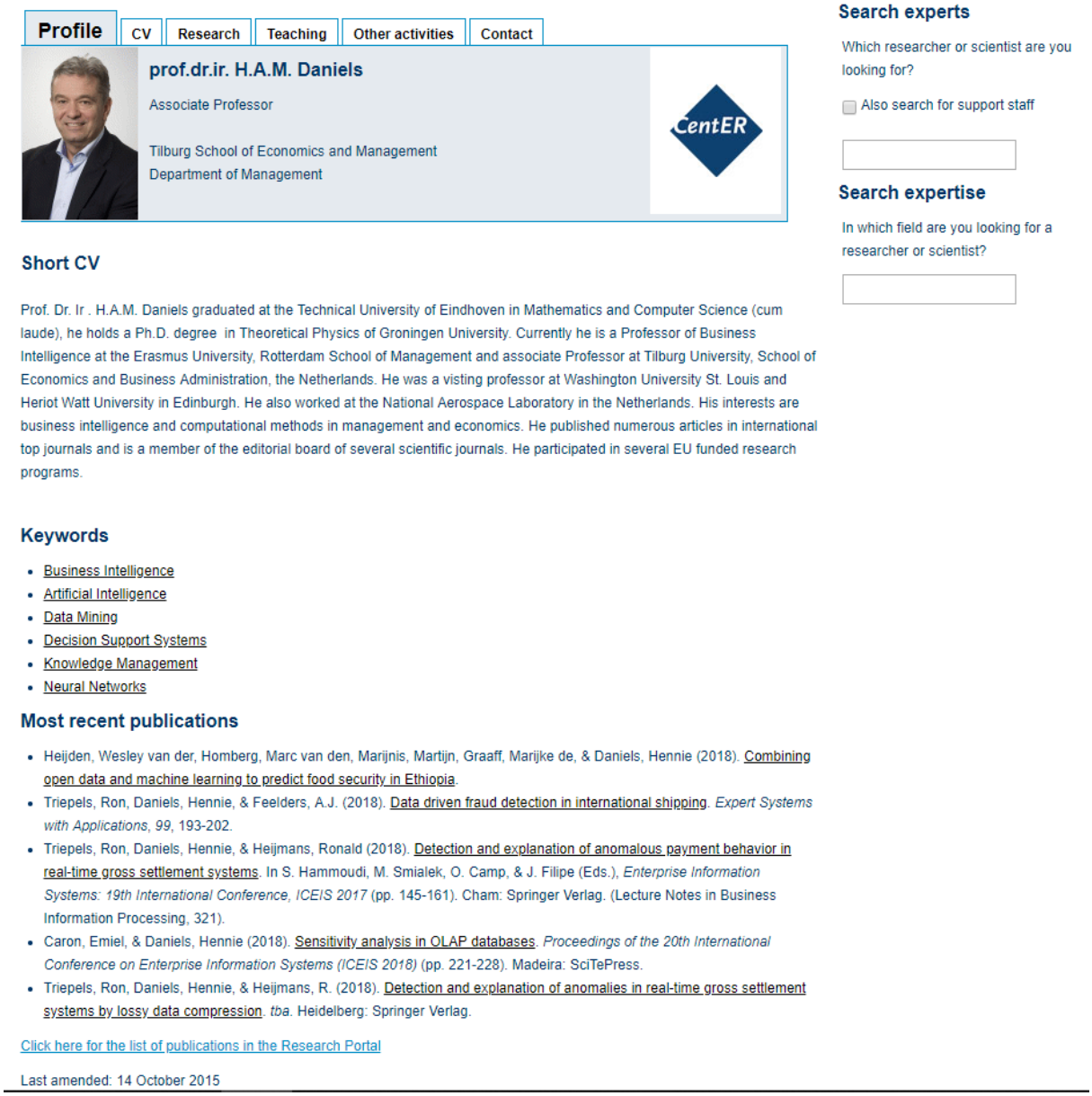
Many studies that discuss the topic of Expert-Finding systems put their focus on the aspect of expert finding methods and techniques that actually are based on the result of expert profiling, but without diving deeper into the specific methods for expert profiling (Lin et al. 2017). This is because expert profiling proves to be a very challenging task that is still being investigated. Also, it is not possible to reverse expert finding algorithms to deal with expert profiling issues (Balog & Rijke, 2007).

Expertise identification poses such a challenge because the expertise pool is very large. Also, expert qualities are multidimensional which contradicts with the very specific and finely grained expertise needs of users. This is why the existing EFSs provide generalized expert findings accompanied by information assumptions, which manages to deal with the expert profiling issue to a certain extent (Yimam-Seid & Kobsa, 2009).

A sample of an expert profile is shown in Figure 4. This is an academic profile, taken from publicly available data from the website information platform of Tilburg University called "Experts and Expertise" (Tilburg University , 2019).

Knowledge profiles can be manually built or extracted automatically from relevant information sources which were discussed in the previous chapter. As mentioned manually inserted self – disclosure is time and resource consuming. Additionally, candidate – documents associated profiles also meet some issues related to finding such association and name disambiguation.

In a study, there were two defined classes in the profiling task – topical and social profile (Balog & Rijke, 2007). A topical profile answers the question "what does expert X know?" by focusing on the type and areas of knowledge and skills of the candidate expert. A social profile, however, aims to provide the user with information about the social influence of the candidate by presenting his/her collaboration network. Social profiling answers the question "who is related to the expert?" and it contributes to the network – graph retrieval model discussed in the previous chapter (Ehrlich, Lin, & Griffiths-Fisher, 2007). Sources for building the social profile are considered to be chat logs, co-authorships, and group member relationships from projects (Zhang, Tang, & Li, 2007).



Profile | CV | Research | Teaching | Other activities | Contact

prof.dr.ir. H.A.M. Daniels

Associate Professor

Tilburg School of Economics and Management
Department of Management

CentER

Search experts

Which researcher or scientist are you looking for?

Also search for support staff

Search expertise

In which field are you looking for a researcher or scientist?

Short CV

Prof. Dr. Ir . H.A.M. Daniels graduated at the Technical University of Eindhoven in Mathematics and Computer Science (cum laude), he holds a Ph.D. degree in Theoretical Physics of Groningen University. Currently he is a Professor of Business Intelligence at the Erasmus University, Rotterdam School of Management and associate Professor at Tilburg University, School of Economics and Business Administration, the Netherlands. He was a visting professor at Washington University St. Louis and Heriot Watt University in Edinburgh. He also worked at the National Aerospace Laboratory in the Netherlands. His interests are business intelligence and computational methods in management and economics. He published numerous articles in international top journals and is a member of the editorial board of several scientific journals. He participated in several EU funded research programs.

Keywords

- Business Intelligence
- Artificial Intelligence
- Data Mining
- Decision Support Systems
- Knowledge Management
- Neural Networks

Most recent publications

- Heijden, Wesley van der, Homberg, Marc van den, Marijnis, Martijn, Graaff, Marijke de, & Daniels, Hennie (2018). Combining open data and machine learning to predict food security in Ethiopia.
- Triepels, Ron, Daniels, Hennie, & Feelders, A.J. (2018). Data driven fraud detection in international shipping. *Expert Systems with Applications, 99*, 193-202.
- Triepels, Ron, Daniels, Hennie, & Heijmans, Ronald (2018). Detection and explanation of anomalous payment behavior in real-time gross settlement systems. In S. Hammoudi, M. Smialek, O. Camp, & J. Filipe (Eds.), *Enterprise Information Systems: 19th International Conference, ICEIS 2017* (pp. 145-161). Cham: Springer Verlag. (Lecture Notes in Business Information Processing, 321).
- Caron, Emiel, & Daniels, Hennie (2018). Sensitivity analysis in OLAP databases. *Proceedings of the 20th International Conference on Enterprise Information Systems (ICEIS 2018)* (pp. 221-228). Madeira: SciTePress.
- Triepels, Ron, Daniels, Hennie, & Heijmans, R. (2018). Detection and explanation of anomalies in real-time gross settlement systems by lossy data compression. *tba*. Heidelberg: Springer Verlag.

Click here for the list of publications in the Research Portal

Last amended: 14 October 2015

Figure 4 Sample of an expert profile (Tilburg University , 2019)

2.2.3 Competencies

To begin with, a competency is a characteristic of an employee that contributes to successful job performance and the achievement of organizational results. A person is competent when he/she has sufficient measurable of assessable knowledge, skills, and abilities in combination with values, motivation, initiative, and self – control that distinguish a superior from an average performer (J.S Shippman et al., 2000; L.M. Spencer et al., 1994).

When it comes to the industrial world there are three main types of competences:

3. *Organizational competency*, which is also known as a core competency (Prahalad & Hamel, 1990). Over the years this term has been defined in many different ways, which now causes confusion when used by different people. Nevertheless, core competence is a design component of an organization's competitive strategy. Some examples of such competencies are value pricing, customer service, reliability, and quick service.
4. *Foundational competency (rear wheel competency)*, is the set of skills, knowledge, and attitude required of an individual regardless of their area of expertise or role. This type of competency usually aligns with the organizational core competencies and culture. Such competencies are teamwork, initiative, adaptability and professional attitude.
5. *Functional competency (front wheel competency),* is the specific set of skills that are required by a candidate to perform his job successfully. Such as a financial specialist should be proficient in financial analysis and accounting.

An overview of the competencies is shown in the figure below (Figure 5).

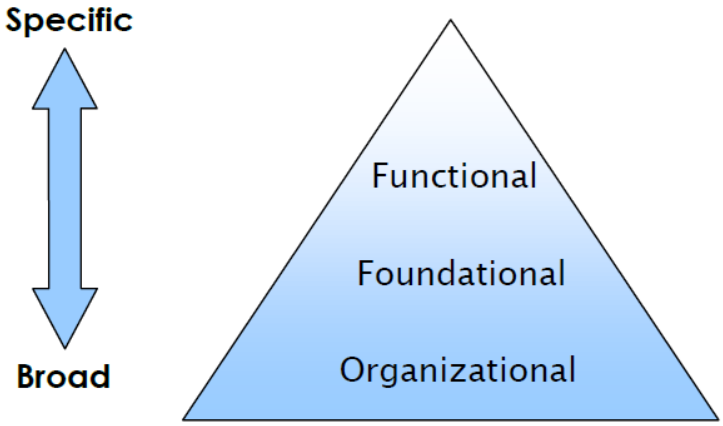


Figure 5 Types of competencies

3. Methodology

This chapter will discuss the research design of this project. This will include information about the chosen method to create the design, the different stages in the research and the relevant research tools applied.

3.1 Research design

For this research, a research design canvas was used (Appendix 2). The canvas consists of nine building blocks in two groups – "T"(green section) and "U"(blue section), which help to align the "DNA" of the study to deliver sufficient insight. The T part consists of four blocks – the problem, the purpose of the study, the research questions and the conceptual framework. It is the foundation of the research. The "U" is the methodology that helps to answer the research questions. It consists of a literature review, overall approach, data collection, data analysis and conclusion (Latham, 2016).

This canvas presents a very structured approach to designing research. By gathering the various activities involved in a research project in two groups (T and U), J. Latham offers an easy to grasp framework, with the help of which the research design can be constructed. As this is a solely conducted project it is very necessary to create a clear structure and design of the research from the beginning to avoid being misled and/or deviate from the topic.

To begin with, a conceptual framework aka concept map is created (Figure 6). In this way, the main problem topics in relation to the research questions are diagramed. A diagram engages visual perception, which is very helpful when analyzing and determining the right research methods. In addition, a diagram presents the main things to be studied such as key factors and their relationship.

“A diagram of the topic is literally worth more than 10 000 words” is said (Latham, 2019). In Figure 6 the concepts and their key factors are presented in a hierarchical order. The hierarchy is displayed in the means of the bigger and smaller sizes of circles and their two-level positioning. In addition, the relationships between the concepts are presented by two types of arrows - a wider and thinner one, pointing out the main and sub-research questions (Novak & Cañas, 2008). The

a concept map is considered to be the very first step of research design. Thus the reason for it to be discussed in the first section of this chapter. This concludes the establishment of the foundation aka "T" of this study.

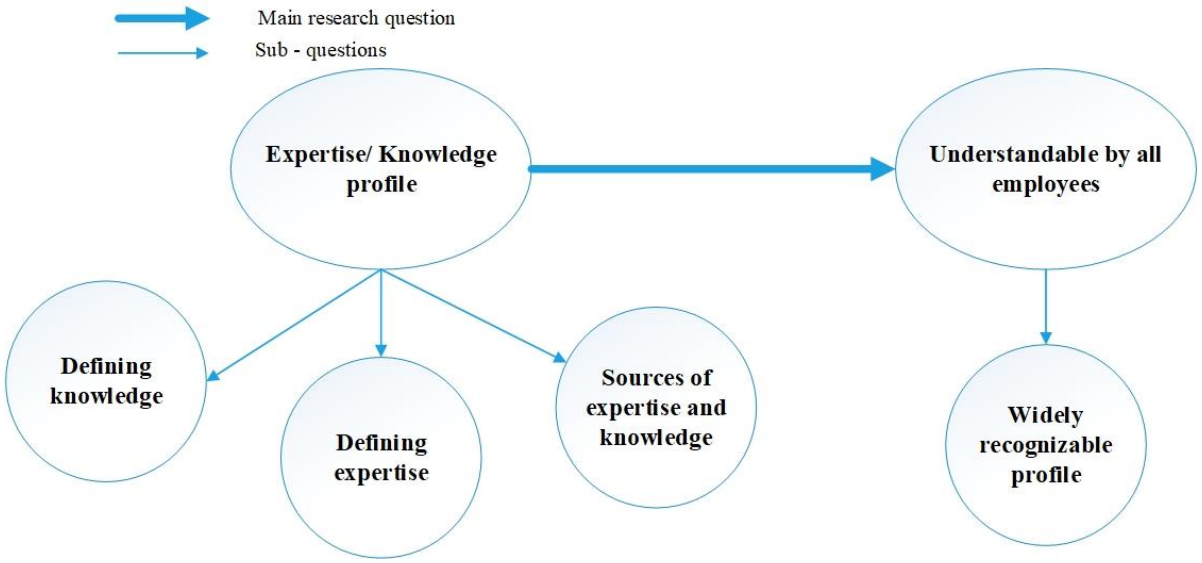


Figure 6 Concept map

3.2 Data collection

This section will describe some of the blocks from the "U" part of the research design (Appendix 2). In the previous subchapter, a concept map was introduced. After the documentary analysis and the field research was conducted, an extended map was made (Figure 7).

To begin with, it was discovered that the research questions are interdependent on one another. This interconnection is visualized with a dotted arrow. On Figure 7 the three sub-questions concerning the definition of knowledge and expertise and their sources are related to the widely recognizable profile sub-question. This is due to the fact that there cannot be a widely recognizable profile without a defined knowledge and expertise, with valid sources. In addition, if a profile is understandable by all employees, then it has a very high potential to be widely recognizable.

On the extended concept map in Figure 7 in some of the bubbles, the number of references is displayed. One reference is considered to be 1 response (from interview or survey) or 1 literature document such as a book or an article.

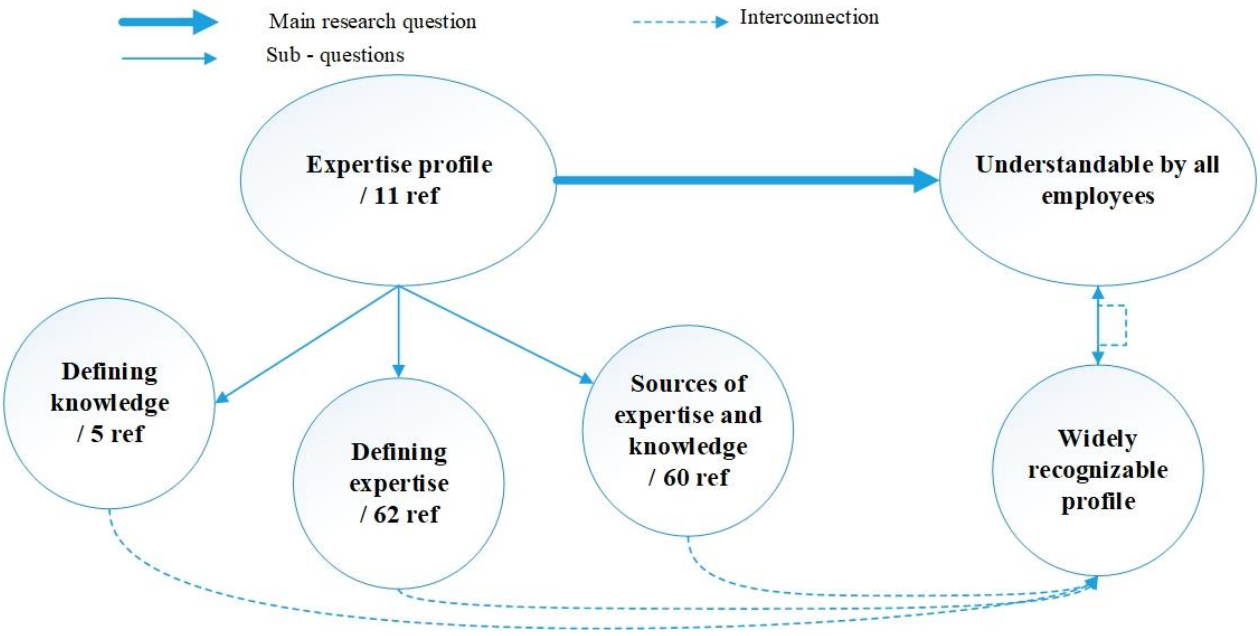


Figure 7 Concept map extended

The next section of this chapter will focus on the data collection and data analysis building blocks from the research design canvas (Appendix 2).

Interviews

When it comes to data collection the first aspect that was considered was the means for gathering information. The selected means needed to sufficiently cover the key factors from the concept map and provide enough insight to answer the main research question.

Since the research topic was directly connected with the employees of Thales, the first chosen research tool was interviewing. In this way, the interviewees would be able to tell their own story in their own terms, which directly links to the end goal to find out how software engineers define their knowledge and expertise. Engaging in a person – to – person interaction was preferred by the researcher because it gives the opportunity to not only develop her communication skills but get to know the personalities of the software engineers. This type of impression aided in determining the components and design of the profile.

For selecting the right interviewees the following criteria were applied:

- The employee must be working in Thales for 1+ years.
- It is preferred the employee belongs to the Application development department.
- It is preferred that an employee resides at Thales Hengelo.

It needs to be pointed out that this criterion was applied when selecting professionals in the software engineering field.

To make sure the most appropriate tool is chosen, the following alternatives were considered: documentary review, observations and unobtrusive measures.

A documentary review is done in the primary stage of the project. It is an utmost necessity to get acquainted with the definitions and specifications of expert finding systems and profiles. However, the usability of the documentary review is mostly applied during the stage of defining these concepts. In the next stages of the project, the documentary review was applied only in small portions throughout the research, which will be discussed later in the chapter.

Observations and unobtrusive measures are not applicable to this research. To get an adequate impression of the knowledge and expertise of software engineers it is necessary to engage in a direct conversation and be intrusive.

In addition, the interviews were taking approximately 20 minutes and they were recorded upon the consent of the interviewee (which was given in all cases). There are several pros and cons of recording interviews, which are as follows:

PROS

- It does not overload the working memory of the interviewer.
- It gives the opportunity to give your undivided attention to the interviewee.
- It allows for a more thorough examination of what people say.
- It allows for the data to be reused and re-examined as many times as needed.

CONS

- There is a chance the interviewee will not give his/her consent to be recorded.
- It can introduce a feeling of discomfort.
- Creates a more formal atmosphere, which can be off-putting.
- The follow – up transcribing process is time – consuming.

Upon development of the interview questions, the three sub-questions were taken into account from the concept map (Figure 6). This consideration divided the question set in three parts: (1) the individual's expertise, (2) the individual's knowledge, (3) sources of knowledge and expertise.

The ending part of the questions was focusing on the knowledge sharing mentality and experience of the individual. This aspect did not directly relate to the research questions, but it was considered useful to discuss because it gives an impression of the current knowledge sharing situation within Thales. Knowing this helped estimate what kind of points attract the attention of a knowledge seeker and what are some preferences for knowledge sharing. The questions set for software engineers is presented in Appendix 3.

Surveys

Another data collection means that was selected was a survey. It is one of the most used data collection means because of its ease to create and wide reach. However, the decision to conduct a survey was taken later in the research. Initially, it was decided to conduct F2F interviews with all participants. This limited the number of participants in the research but guaranteed the personal story to be received. As the interviews went on, it was noticed that there has been a certain information level reached and the meetings started to seem more and more alike.

Therefore, the decision to conduct a survey was taken (Appendix 4). This allowed to reach a wider number of participants and get the most important information needed for the research.

The tool used to create the surveys is called LimeSurvey. It is the method that Thales uses for its internal e-communication. Since Thales has strict security protocols, it was necessary to use the already established tool LimeSurvey. It gives the opportunity to create various types of questions such as array questions with different choices, multiple choice questions, single choice questions, and text questions.

From the initial F2F interviews it was clear what questions and aspects needed to be addressed to get the same results as from the interviews. Using that information the survey was also structured in three parts:

1) Description of the roles and responsibilities of the participant – this was an important question for the survey. Since there is no in-person interaction, it is important that the specifications of the function of the participant are addressed. This is because even though people would hold the title software engineer, the tasks each of them is responsible for

2) could differ based on their department and product focus. This is why there was also a question that specifically asks for mentioning the product focus.
3) Software engineering knowledge and expertise – this section asked the participant to select from a number of options regarding the languages, tools, frameworks and internal tools and frameworks that he\she is familiar with. In addition, the user was required to indicate the level of proficiency of the selection. This section asks for the education background and any other knowledge and expertise that the participant may hold (outside of the sector of Thales).
4) Personal opinion – in the last third part of the survey, the participant was asked about his/her opinion about the topic(s) of their expertise. This gives an indication of the personal perception of the individual.

The survey had a total of 10 questions as each part of it had its own question types, which were as follows (Appendix 4):

- Type 1 makes use of long free text and questions with multiple options with a comment section.
- Type 2 questions were array by column, multiple options and Yes/No question with a comment section.
- Type 3 was presented by a Yes/No question, long free text and Yes/No with comment section question.

Nevertheless, there are several pros and cons of surveys, which are as follows:

PROS

- Easy to create
- Has a wide reach
- Delivers a large amount of data

CONS

- The threat of survey fatigue
- Wrongly formulated questions can lead to inaccurate data
- Risk of biased answers

To estimate the number of participants needed to conduct sufficient research, a statistical calculation for determining sample size was used (Figure 8) (Select Statistical Services Limited, 2019).

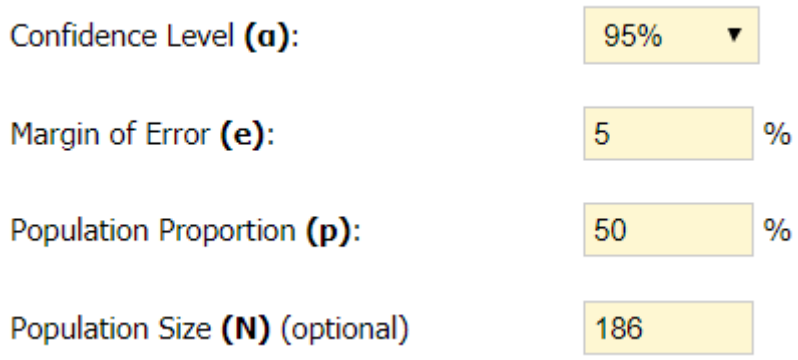


Figure 8 Determining sample size

The sample size (n) is calculated according to the formula:

$$n = \frac{[z2 * p * (1 - p) / e2]}{[1 + (z2 * p * (1 - p) / e2 * N)]}$$

Where: z = 1.96 for a confidence level (α) of 95%, p = proportion (expressed as a decimal), N = population size, e = margin of error.

z = 1.96, p = 0.5, N = 186, e = 0.05

$$n = \frac{[1.96^2 * 0.5 * (1 - 0.5) / 0.05^2]}{[1 + (1.96^2 * 0.5 * (1 - 0.5) / 0.05^2 * 186)]}$$

$$n = \frac{384.16}{3.0654} = 125.322$$

$$n \approx 126$$

The sample size (with finite population correction) is equal to 126.

The 50% sample proportion is due to the goal to get a sufficient amount of input for this project. Even though the time frame of the project was only 5 months the best possible scenario was to get input from 126 software engineers.

Additionally, the value for the population size is taken from Thales' human resources department database. The amount is the total number of software engineers in the naval domain, from the job

family 06 Software engineering. This type of domain composes the majority of the workforce at Thales Hengelo (Employee 1. , 2019).

However, it should be noted that this number is not taken as definite. It is used as an indication of the best sample size for a sufficient result. This needs to be considered because employees from other companies were involved as well. This factor makes it more difficult to stick to an exact number. In addition, the level of responsiveness played a very important role in the final number of participants in the research. Luckily, employees at Thales were very welcoming and willing to spare some time to come to an interview or fill in the survey form.

3.3 Data analysis and conclusions

As mentioned in the previous subchapter each F2F interview was recorded and transcribed. The transcription was made as a summary of the most important and relevant information mentioned during the interview and not the usual word by word writing down that comes with transcribing. The interviews were 20 in total so their manual transcribing was not that time-consuming. Afterward, to analyze the summaries, a list was made. This list consisted of several points: (1) role of the interviewee, (2) technical knowledge such as languages, frameworks, and etc., and (3) product focus. This list was made manually.

Additionally, for the survey results, the Lime Survey tool was used. This software has the option to store and provide a statistical overview of the responses. The results can be exported in a word or pdf sheet that will contain cross-tabulated data and pie chart summaries of the responses (example can be seen in Appendix 5).

The survey was sent out the Application engineering group of Thales Hengelo. There were a total of 154 invitations sent. However, it needs to be pointed out that not all of the employees included in the Application engineering group were software engineers. This is due to the structures of the contact groups at Thales. Therefore, it was expected that the response rate would be lower than 154.

After the survey was conducted, the total number of respondents was 40. This lead to a survey response rate of 25.97 % and a total of 60 responses in the whole research (Table 2) (Typeform, 2019).

Table 2 Response analysis in numbers and percentage

SAMPLE SIZE - 126

RESPONSES		
INTERVIEW	*20*	*15.87%*
SURVEY	*40*	*31.75%*
TOTAL	*60*	*47.62%*

The results will be used to see compare the expertise of people (Appendix 4 q. 1, 2, 3, 5, 8) with the perception of their expertise (Appendix 4 q. 4). The addressed information will give an insight into whether there is a correlation between the two variables and if it can be used as a reliable source of expertise for the expert profile.

With these insights, the right components of the expert profile can be estimated.

Nevertheless, the reliability and validity of the results are very important issues in this research. This assignment is focused on people, their knowledge, expertise, and perception. These factors cannot be fully represented by numbers, thus a numerical dataset cannot be created to provide explicit and objective marks. Marks that are to be used to support the development of the final design of a profile.

Also, when people talk about themselves, they are often biased and untruthful. It is hard to be fully objective when one opens up about his/her knowledge and expertise, especially in a professional environment. The threat to appear unqualified and incapable of successfully performing the job is quite present in such cases.

To deal with this matter, a lot of focus was put on the interviewing technique. The questions were putting attention on what does the person know and feel confident in. This emphasized the person's strengths, which set a positive mood for the talk. When such an environment and motivation are present, people tend to be more honest because they do not feel threatened. In addition, it has to be taken into account that the position of the researcher was not threatening. Being an intern and asking employees for their input does not intimidate them thus they do not find any reason to be untruthful about their knowledge and capabilities. In reality, some even find

it complimenting that they have been contacted. The case might have been different if the talk was between a CEO and an employee.

In the end, on a larger scale, the findings will be used to develop an expert finding system within Thales Nederland. With time the system would broaden its scope to Thales Group and other companies. The version of the design of an expert profile will be used as a starting point in this development. This research gives insight into how do people profile their knowledge and expertise. More importantly, it shows where they consider themselves to be experts in. This information cannot be taken from social media and curriculum vitae.

Most importantly, a knowledge profile will fill in the knowledge gap of the new coming employees of Thales. It would aid the transfer of knowledge from senior to junior workers and create a network where people would be aware of who knows what and the different knowledge groups.

4. Results

This chapter focuses on presenting the information that resulted from the research activities including documentary analysis, interviews, and surveys. The structure of the chapter and the way the results are presented are motivated by the end goal of this project – creating an expert profile. Thus, the chapter is divided into three main subchapters. The first subchapter focuses on the results of determining software engineering knowledge areas and expertise and the second subchapter presents the information about the components, design, and sources for the expert profile. The third subchapter will deal with the interpretation of the results. Relevant references are made to clarify where the data resulted from.

4.1 Determining knowledge areas and skills: Software engineering

4.1.1 Classification of knowledge areas and skills

Firstly, a recap from chapter 2 about the software engineering body of knowledge - SWEBOK. According to this guide, there are 15 knowledge areas that are composing the knowledge of a professional software engineer in the industries (IEEE Computer Society , 2014). These areas are:

- Software requirements
- Software design
- Software construction
- Software testing
- Software maintenance
- Software configuration management
- Software engineering management
- Software engineering process
- Software engineering models and methods
- Software quality
- Software engineering professional practice
- Software engineering economics
- Computing foundation
- Mathematical foundations

- Engineering foundations

These areas are used as a classification for the knowledge areas for the expert profile. This was the starting point in this research project. In addition, the curriculums of the educational institutions Saxion and UTwente were taken into account (Table 3). When looking at the course Computer Science (the degree held by the majority of the interviewees; Appendix 6), it was visible that the knowledge is divided into several parts that also compliment the SWEBOK knowledge areas (Saxion, 2019). The topics in the program were focused on computer systems, software systems, network systems, data and information (which was presented by data management and relational database), mathematical algorithms, programming paradigms and project management, communicational skills and personal development (UTwente, 2019).

Table 3 Overview of the curriculum for a Computer science program

Computer Science program curriculum overview	
Hard skills	**Soft skills**
Computer systems Software systems Network systems Data and information Mathematical algorithms Programming paradigms	Project management Communication skills Personal development

Regarding skills, the interviews showed that they are perceived as the personal qualities that are applied on the job and distinguish one person from another. Some examples of skills are:

- Logical thinking and problem solving
- Analytical skills
- Good time and task management
- Quick learning ability
- Attention to detail
- Focus on continuous skills development

From the field research conducted in the means of interviews and a questionnaire, it became evident that there are 3 types of perception of expertise that are concluded: technical expertise, soft skills expertise and hybrid expertise (Table 4).

Technical expertise is presented by the hard technical knowledge on topics such as Java, Testing, Data visualization, Software architecture design, and Network technology.

Soft skills expertise is presented by the skills such as being able to bridge the gap between system engineering and software engineering, management of foreign projects, being able to establish a common ground between people with different points of view, being able to motivate and guide a team towards a collective goal and analytical skills.

The third type of perception of expertise – a hybrid, is the combination of the two mentioned above such as being able to reduce the complexity of a task to a pragmatic design, reversed thinking related to root – cause analysis. The latter requires both an understanding of the topic and strong analytical skill.

The table below presents a summary of these types accompanied by the number of people that hold that perception, based on interview and survey response.

Table 4 Views on the types of perception of expertise

EXPERTISE PERCEPTION TYPE	N OF PEOPLE WITH THAT PERCEPTION[1]	
TECHNICAL EXPERTISE	*32*	*53.3 %*
SOFT SKILLS EXPERTISE	*8*	*13.3 %*
HYBRID EXPERTISE	*13*	*21.7 %*

4.1.2 Outline of knowledge areas and skills

The research resulted in numerous knowledge topics and skills that are concluded in the following groups (UTwente, Saxion, 2019):

- Software systems

[1] 11.7 % (number of 7) of participants' responses were not applicable to this table

    - Programming: C, C++, ADA, Java, Python, Bash, HTML, CSS, Groovi, etc.
    - Design: UML, BPMN, AML, SysML, etc.
    - Testing: Cucumber, SMARTesting, etc.
- Network systems and protocols
    - Ethernet, GSM/UMTS
- Operating systems
    - Linux, Windows, Android, iOS
- Software management tools
    - Active MQ, Camel, Jira, Maven, InteliJ, Junit, Storm, etc.
- Individual's expertise trait
    - Communicative, analytical, the good root causes analysis skills, able to master new content fast, etc.

4.2 Expert profile: Software engineer

The expert profile is to be applied within the Thales network. However, it is aimed that this system will expand beyond the borders of Thales. Therefore, this has an influence on the way the profile is designed. The components and structures discussed in this chapter aim to give a general overview of the individual's knowledge, expertise, professional experience, and personality. Each of these variables will play a role in the process of selecting an expert by the user. Conversely, since the users would not be only Thales – internal, these variables need to be presented in a very general and clear way.

4.2.1 Structure

To begin with, the profile is structured in 3 main parts:

- General information
- Expertise
- Hobbies and communities

General information comprises the information about the expert's name, profession, department, team and detailed contact information.

The expertise section is the main body of the profile. It will present all the information about the body of knowledge and expertise the person holds. This section will not contain only the topics that the "expert" is fully proficient in. To be able to provide a full picture of the body of knowledge of the individual all topics that he\she is familiar with will be included. To avoid confusion and misunderstanding, the proficiency level of each topic will be indicated. This way it will still be clear where the expertise of the individual lies.

The third section concerns a softer aspect of expertise. It deals with the topics that the person considers as a leisure activity and the communities he/she is a part of. The communities section is specifically applicable for Thales – internal.

4.2.2 Components

Section 1: General information

As mentioned in the previous subchapter, this is a must-have part of any profile. In the case of the expert profile what is included is:

- Name
- Profession
- Department
- Current team
- Contact details and profile photo

The name section is a free text bar, which will allow the individual to type in his/her name. This approach was taken because it would be more efficient to let people type in their names rather than having a drop-down list of selection. Another option, to automatically identify the person by making a link with the account he/she is in. This will allow for an automatic insert of the name.

The profession is a drop-down list, from which the user will have to select the one that is closest to their profession (Figure 9).

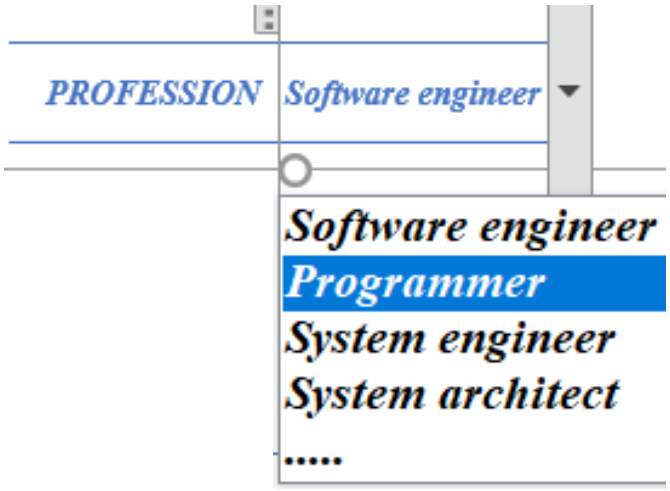


Figure 9 Design of profession component

As there are numerous professions, a single drop-down list would be endless. Therefore, it is advisable to combine the drop – down with a text box that will serve as a filter and cut down the options by typing in the first letters of the desired title.

In addition, it is taken into account that there are many different titles in the industrial world and many are presenting the same thing. Therefore, a drop - down list will engage the user's perception of his/her position and then will choose the one position that best fits his/her activities. This way a general database of titles, understood by all, will be maintained.

There are professions that are overlapping in knowledge and expertise such as a programmer and a software engineer. The knowledge and expertise of a programmer are much like the one of a software engineer but it is more focused on coding (Stowe, 2018). Programmers are experts in programming languages that are used to implement the design and architecture of a product. However, a software engineer is mainly responsible for troubleshooting and managing the project from start to finish. This profession concerns all stages of software development such as requirements setting, risk management, aftercare and etc. Therefore, if the title programmer is selected, the expertise section will present certain parts as highlighted and other as not – highlighted to give the user an opportunity to decide for himself, which is applicable to his/her job.

After selecting the profession, the rest of the profile will automatically adjust to present options that are in relation to the position selected. The department component will be a drop-down list (Figure 10). The variables considered here are the profession selection and the company

environment the profile is created in. Some departments' names are only meaningful for the company they are in. Thus, for this profile, the names are to be generalized.

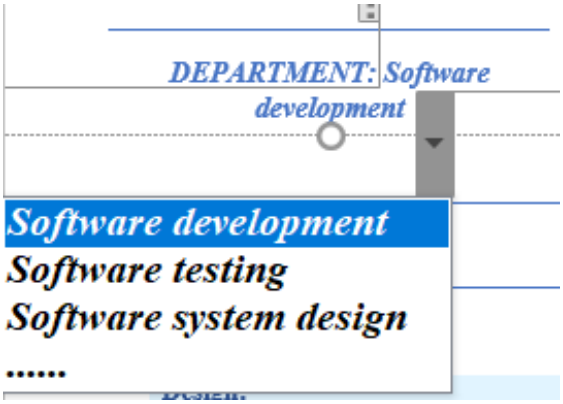


Figure 10 Design of department component

Following is the current team box, which is a free text box. The last component of this section is the profile picture and contact details, which entail phone number, email address, and location. Since this is explicitly private information a warning message will be shown stating: 'this information is privacy sensitive and will not be displayed, unless explicit consent from the owner of the information is obtained, to share this information with another person'

Section 2: Expertise

For the components of the expertise section, there were several variables taken into account: the software engineering body of knowledge, relevance to the industries and personal traits (Figure 11).

For the structure of this section, the content of the Computer science course had great influence (Saxion, 2019). This was useful in presenting the different topics of expertise in a clear and general way.

As it is visible in Figure 11 the left side of the section is dedicated to the hard technical knowledge and expertise of the user. As a software engineer, the first topic of expertise that is addressed is Software systems, with subtopics of Design, Programming skills, and Testing. They are only a part of the process of developing software. A professional in that field should first have an understanding of modeling, which is where the design of the software is discussed. After that, comes the specific programming language that is to be used to develop the software. Consequently, the developed software needs to be tested. However, depending on the engineer, testing plays different roles in a person's position. Some software engineers are more responsible

for testing than others (Employee 4. , 2019). Nevertheless, it is an essential part of the body of knowledge.

The next component is Network systems and protocols. This focuses on computer networking for infrastructures such as GSM/UMTS and embedded networks. Following are Operating systems such as Linux, Windows, and MacOS. Then is Data and information which is about database management and discussed SQL/PLSQL, HTTP, and XML/Jason.

It is a section for Software management tools, which as the name entails are the different tools that the user is familiar with regarding software project management.

All sections, except Software management tools, required to indicate a proficiency level. The user can choose from four levels:

- Basic: can understand and is familiar with the concept of the topic; can discuss it in simple terms; has no practical experience in it
- Limited working proficiency: can describe details of the topic and discuss it using complex terms; has limited practical experience (applied in a project or in their free time)
- Intermediate: has extensive knowledge on the topic and can provide advice and recommendations; has sufficient practical experience (applied in several projects) and can provide references for other people with similar knowledge and expertise
- Full professional proficiency: has extensive knowledge on the topic and can analyze it; can provide advice and recommendation and come up with own ideas for improvement of the area; has extensive experience (applied in 4+ projects) and has an established network of people with similar knowledge and expertise

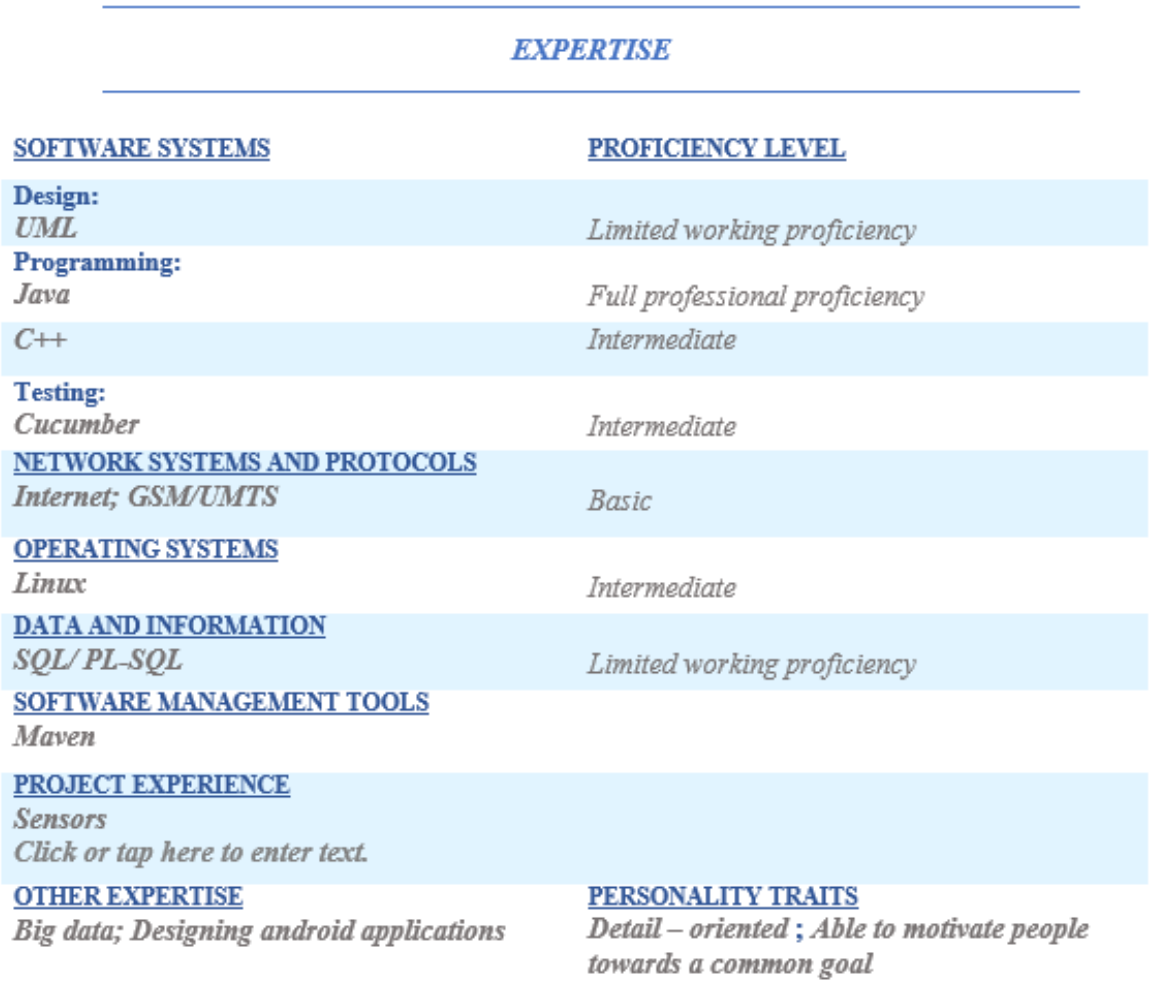
The motivation for the development of these levels was taken from the Common European network of reference for languages (COE, 2019). Going through the commonly accepted definitions of proficiency resulted in the expertise level definitions above. The aspects that differentiate one level from another are knowledge of the topic, practical experience and knowledge network. Practical experience is expressed by the amount of project – related work the topic has been used in. In addition, the knowledge network is the number of people that the expert is acquainted with, that have the same or similar knowledge and expertise. This is included because one part of expertise is that the expert can provide a reference to other people that can help solve the problem aka social profile.

Last but not least, are the sections Project experience and Other experience. Project experience refers to the specific product or topic of the projects the expert participated in. For example, in the case of Thales, this could be Sensors or Tacticos. It is a section that addresses the projects in connection with the company, the user is currently employed in. On the other hand, Other experience addresses the previous or non – current employment-related experience. This can be any expertise or knowledge that the person has but does not apply in his/her day – to – day tasks in the workplace (Figure 11).

To sum up the expertise section, there is a Personality trait component. It is an important part of expertise. As seen in Table 4, many people perceive expertise to be a personality trait that they exceed in and not necessarily a hard technical skill. Therefore, it is important to add this component to the profile because it could also help in better matching a search query and an expert profile also stated in subchapter 4.1.1.

Section 3: Hobbies and communities

In Figure 12, there is a representation of the hobbies and communities section. There is a box that addresses the communities the expert is a part of. This component is very specific for the Thales Group employees and would only be visible within the Thales network. Much like the hobbies and interests component, it is important to include it to ensure the best matching with a query. It can be the case, that the user uses the system to find information about the software engineering communities within Thales Group. Thus having this component on the profile, will aid in the matching process and the user's issues to be resolved.



EXPERTISE

SOFTWARE SYSTEMS	PROFICIENCY LEVEL
Design: *UML*	*Limited working proficiency*
Programming: *Java*	*Full professional proficiency*
C++	*Intermediate*
Testing: *Cucumber*	*Intermediate*
NETWORK SYSTEMS AND PROTOCOLS *Internet; GSM/UMTS*	*Basic*
OPERATING SYSTEMS *Linux*	*Intermediate*
DATA AND INFORMATION *SQL/ PL-SQL*	*Limited working proficiency*
SOFTWARE MANAGEMENT TOOLS *Maven*	
PROJECT EXPERIENCE *Sensors* *Click or tap here to enter text.*	
OTHER EXPERTISE *Big data; Designing android applications*	PERSONALITY TRAITS *Detail – oriented ; Able to motivate people towards a common goal*

Figure 11 Design of Expertise section

HOBBIES AND INTERESTS

Beginner course in ADA

Karting

THALES ENGINEERING COMMUNITY MEMBER OF

Apache, Celix, ASF, Kubernetes, ISTIO

Figure 12 Design of hobbies and communities section

4.2.3 Sources

For the information presented on the profile to be valid, a reliable source of information is needed. Throughout the research several possible types of sources were discussed:

- Project documentation
- Project's final product
- Others' feedback and opinion
- Educational background

However, it was discovered that these are not considered reliable sources of knowledge and expertise. As all work is project – based, this means that each member of the team is equally involved in the development of the documentation and the final products. When it comes to software, it is possible to point out the value of the code written by the means of comments but that cannot be used as a source for the expert profile as these comments are revised and updated constantly.

In addition, other's evaluation of your skills and expertise could be a valid source of information if it wasn't for the high risk of biased opinions. As we are human beings, we often form an opinion of others based on feelings. Thus, this could cloud the judgment of the person's actual performance in the workplace. Nevertheless, how people make others feel is an important factor in their way of working. If a person likes to keep to him/herself and does not communicate, it would be very challenging to work in a team with. On the contrary, if someone is more expressive and willing to converse openly that will create an easy work environment that could have an influence on the quality of work.

The fourth factor, the educational background, can be considered relevant but its value is still arguable. Based on the survey results (Figure 13) educational background is applicable to a software engineer's responsibilities. However, it should be considered that most of the survey respondents have either computer science or software engineering background (Appendix 6). With time some of the knowledge acquired during academic times fade away and is replaced with newer one, more applicable to the current responsibilities.

Additionally, there are cases where people, with a background for example in psychology, become software engineers. There is a lot of training offered by companies, which is an opportunity to learn and develop further on the job.

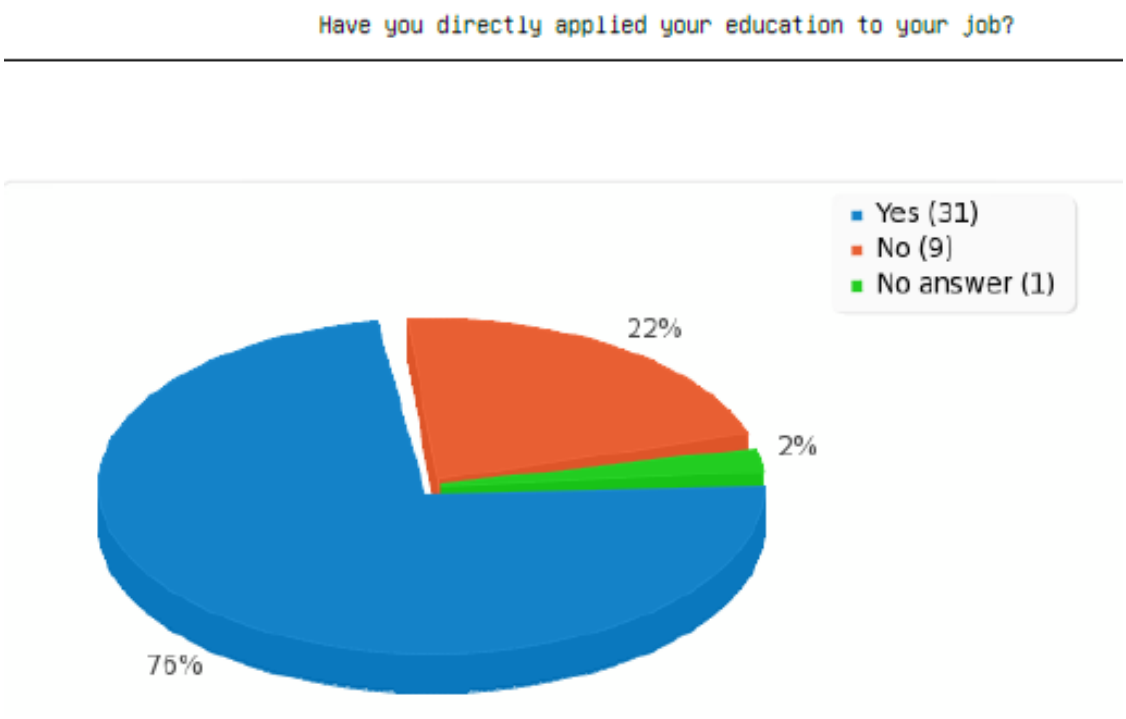


Figure 13 Survey results for the application of educational background on the job

4.2.4 Functional design

Considering the variables and components discussed so far in chapter 4, the functional design was created for an expert profile for a software engineer (Figure 14). Since the goal is to make this profile general, a definition of terms is needed (Figure 15). The definition will be displayed, when the user clicks on the term. In this way, all users will have the same understanding of the terms mentioned on the profile. That will aid in selecting the right topics of expertise by experts and the right profiles by users.

A term is considered to be a word that is displayed on the profile in regards to the person's expertise such as the name of the department, name of the profession the topic of expertise, etc. Each term will have a definition available.

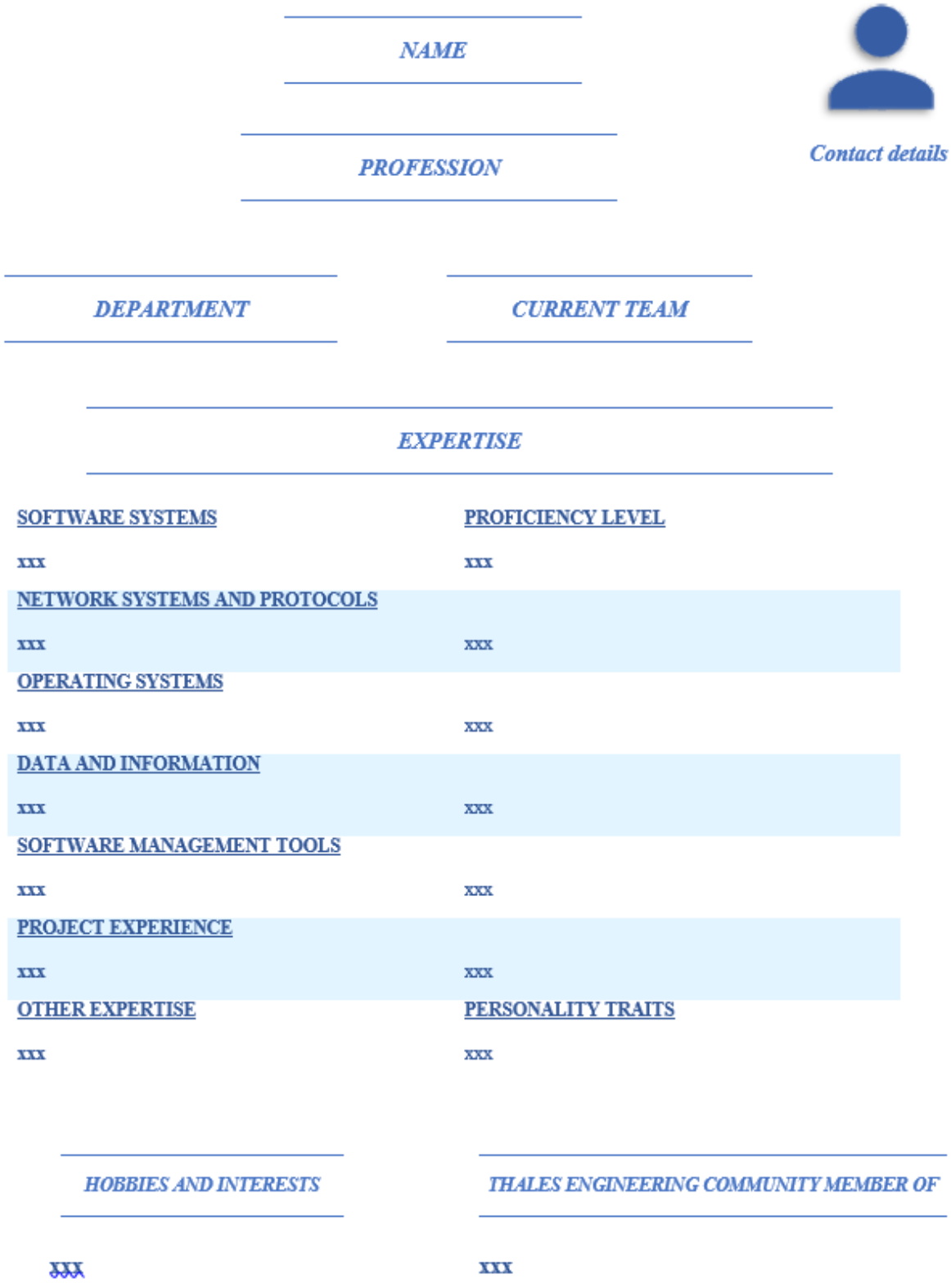


Figure 14 Functional design of an expert profile of a software engineer

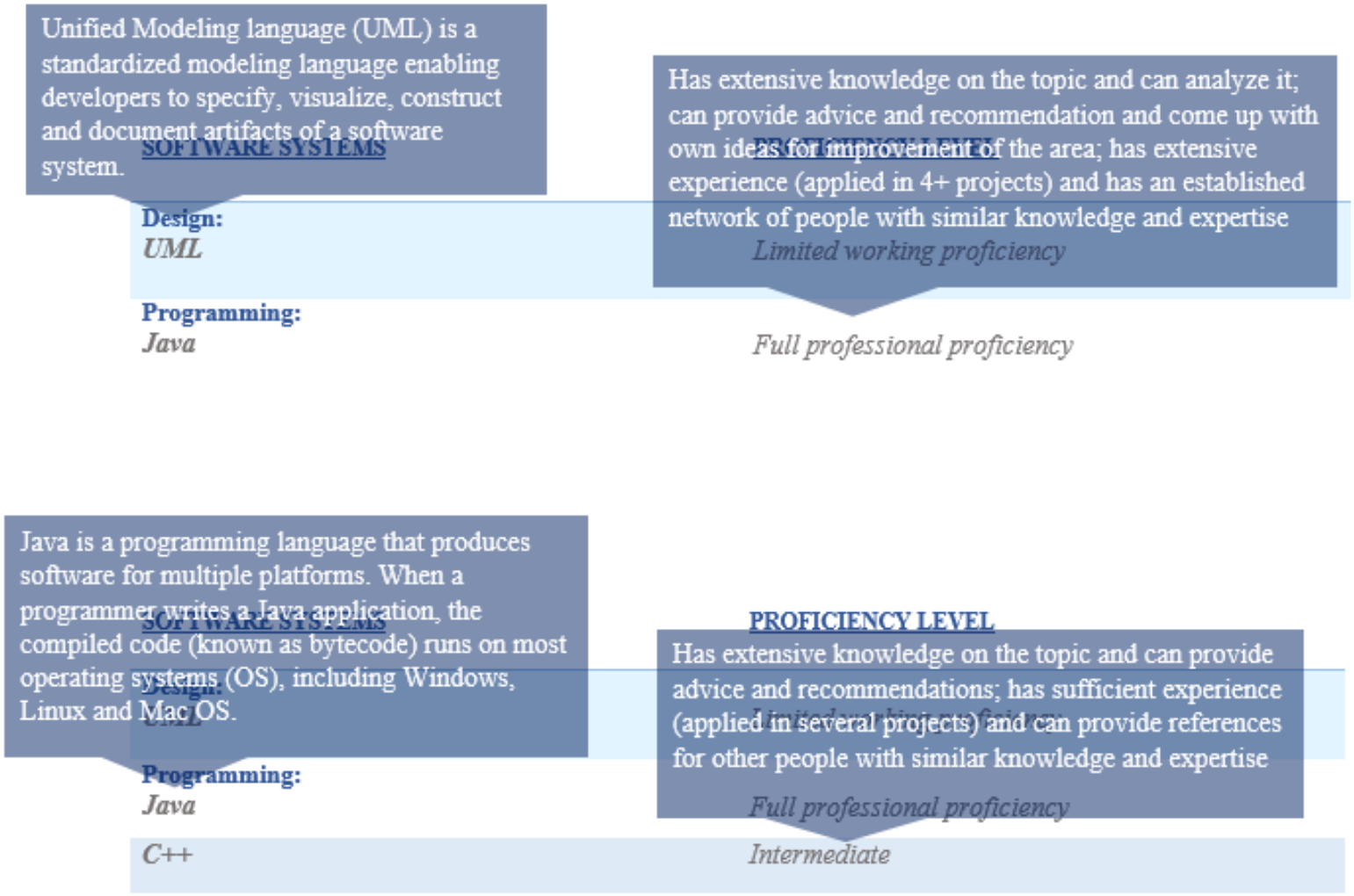


Figure 15 Example of terms definitions on expert profile

4.3 Results interpretation

4.3.1 Interpretation of knowledge areas and skills result

Chapter 4.1 presented a set of determined knowledge areas and skills that directly connect to the knowledge and expertise of a software engineer. However, in practice, these areas are not clearly defined and it is challenging to try to define them. Software engineers are applying most of these skills and knowledge to perform their daily tasks but they are so interconnected that it has become an unconscious process. For example, the areas of computing, mathematical and engineering foundations are more of a second nature than a conscious application of knowledge on the field. During programming, the engineer does not actively think about how the computer works. In addition, rarely there is a need for a mathematical calculation.

On the other hand, areas such as software requirements, software quality, and software engineering professional practice receive a lot more attention and focus. In addition to these hard skills, a soft skill – project management expertise (f.e. PRINCE, Kanban, Six Sigma, etc.) also plays an important role. This soft skill expertise also connects to several SWEBOK topics – software engineering management and software engineering professional practice.

The above mentioned are knowledge areas that are actively applied in the day – to – day activities of a software engineer. They are considered to be of high importance for a person to be able to perform successfully on the job. Also, such areas are emphasized in job descriptions, often as a must, and are containing the core skills and competencies desired for a position (Employee 2. , 2019).

The personal traits such as logical thinking and analytical skills, are important to mention in an expert profile because it gives an impression of how a person is approaching his knowledge and expertise and what is his/her method of working (Employee 3. , 2019). This helps to formulate a certain expectation about the answer a user would receive from the expert on the profile. Upon examining the profile results, the user already has a predefined expectation of the type of answer he/she wants to receive. People do not go into research with a blank mind without formulating certain requirements for the results (Bondarouk, 2019). Thus, if a user expects a detailed explanation, mentioning “detail – oriented” in the profile might determine the final match.

From the data in Table 4, it became evident that expertise is a very perceptual matter. Just as in the case of knowledge, there is not one clear way to define expertise. The most common perception of expertise according to Table 4 is the technical expertise. However, it is important to note that the research was conducted in a technical environment, where technical expertise is highly required to perform well on the job. With that said, if the same research was to be conducted in a different department, f.e. HRM, the results would differ.

4.3.2 Interpretation of expert profile result

The components of the expert profile were motivated by the general structure of a social profile. Every profile needs a general information section and this is mandatory. A profile without a name or contact details would not be complete. However, that section could also engage in personal preferences and activate a biased selection process. Therefore, this section would be visible after the user has selected the profile based on the expertise.

In the case of departmental and profession titles, generalization is of utmost importance. The expert profile in this research aims to be understood by everyone. Since such titles heavily differ from company to company, it is crucial to create a set of terms that each user needs to select for him/herself. In addition to the definition of terms, this will achieve the desired simplification and understanding of the information in an expert profile.

To achieve this goal, the levels of proficiency are also of great value. This is just another way towards creating a generally understandable profile. When a common definition is presented to the public, each individual would adjust his/her own perception and select the definition that is closer to it. This is why the definitions of proficiency levels are described in such a way to address multiple aspects of knowledge and expertise.

Addressing the hobbies and community memberships of a person can also be a valid expertise trait. As many relate their expertise to what they do as a profession, this section can be considered irrelevant to an expert profile. More often than not, people have hobbies that differ drastically from their profession such as a leisure activity like painting, horseback riding, sailing, etc. However, it is not certain that this will always be the case. There are individuals, who enjoy programming as a leisure activity. That programming may not be part of their profession but could prove to have a connection with the search query of the user. Therefore, this component is

included in the profile as a reassurance that the right expert will be discovered and matched by the system. Also, in the case of this report, the communities' component is Thales – specific. However, it is a voluntary component that can be included if a company makes use of online communities.

4.3.3 Interpretation of sources result

The possible sources presented in 4.2.3 are concepts that were shortly discussed during interviews. The topic was included in the interviewing sessions (Appendix 3). However, due to time limitations, it was not possible to further dive into examining the validity of the mentioned options.

4.3.4 Interpretation of functional design result

The design presented in chapter 4.2.4 is the very first step towards an expert profile design. Therefore, it is very simple and aims to serve as a guide for future developments. This kind of the first step is very important when it comes to research projects such as this one. In such cases, the final product might not be 100% satisfactory but its main goal is to give insight for future actions. Thus, here the main focus is content rather than design and aesthetics.

In the case of Figure 14 and the sample profile in Appendix 7, there are several components that should grab the attention of an interested party. Profession and department are an obvious component for such a profile. However, here the interesting part is the terminology chosen for the drop-down list. Also, the fact that each term displayed on the profile will be provided with a definition is a very important aspect that guarantees the generalization of the profile.

Another goal – oriented component in this profile is the proficiency level column. By addressing both topical and social profiles in the descriptions of the levels, it's aimed to reach the desired generality.

4.3.5 Further interpretation

The results from this research raise some questions regarding the future of an EFS. Below are some suggestions, which are not limited to but including the following research topics:

- What are valid sources of expertise?
- How to best match a search query to a profile?

- How to best present the new system, so that all employees are aware of it?
- How to stimulate the use of an EFS?

In addition, the research methods applied in this project were restricted to the depth of insight they delivered. EFS's main center is people and its success highly depends on people's understanding and involvement. Thus, future research methods should focus on high people input.

It would be advised to consider the setting up of a focus group and conducting a case study. The case can go on as follows:

1. Gather 4 – 5 people, who already know each other.
2. Create an expert profile of each of the people in the group.
3. Anonymize the profile by hiding the name, department, and profession
4. Distribute the anonymous profiles within the group and ask them to point to which person does it belong

The outcome of this case will show the level of understandability of the profile. The results can be used for further adjustments in its content and terminology.

5. Conclusion and recommendations

This chapter will present the final conclusion of the project by considering all the information discussed in the report. No new information is presented as far as the conclusion is considered. The recommendation will follow as a next subchapter, which deals with advice on future actions in concern with the successful development of an Expert Finding system.

5.1 Conclusion

To begin with, according to the results from the interviews and survey, expertise is a very perceptual matter. Just like the definition of knowledge, there is not one generally accepted and applied definition and it can change from person to person. According to one definition, knowledge is the fact or condition of knowing something with a familiarity that is gained through experience; the sum of what is known: the body of truth, information, and principles acquired by humankind (Merriam-Webster, 2019). Another definition states that knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It can be found in the mind of knowers and in organizations, it is embedded in documents, routines, processes, practices, and norms (Davenport & Prusak, 1998).

From these two definitions, it can be concluded that knowledge is composed of data, information, and experience. It is a multi-layered mix of concepts, which makes it very difficult to define and capture in one general statement. The aspect that makes knowledge that challenging to outline is the human factor of experience. Until now, there is no established way to define and measure experience in a valid and reliable way.

Therefore, this project aimed to establish a common understanding of knowledge and expertise by creating a sample design of an expert profile for a software engineer.

SWEBOK and the curriculum for a Computer science program from Saxion and UTwente set the base for understanding the knowledge points in SE. But in practice, these areas are intertwined and cannot be fully distinguished with the use of a single tool or method.

Personal perception of the individual's body of knowledge mainly lies within the day to day activities, languages, tools, and frameworks applied. The time spent using these affects the

perception of the level of proficiency. The more time spent, the higher the level of expertise the person perceives to have. In addition, there is another factor that determines the proficiency level on a topic – people's network. The more one person holds expertise on a topic, the more he/she knows other people that have similar or overlapping knowledge and can be used as a referral aka social profile.

Nevertheless, it proves very challenging to determine the knowledge and expertise that one holds. It was noticed that during the F2F interviews people gave different types of answers when their personal expertise was addressed in comparison to the expertise of a software engineer. Throughout the conversation, it was detected that people have the same perception of software engineering knowledge and expertise. However, when their personal expertise was addressed they often turned to personality traits that distinguished them on the job from the rest. As a result, there are three types of expertise perception: technical expertise, soft skills expertise and hybrid expertise. This result motivated to create two different types of components in Section 2 of the profile – one for technical skills (e.g. Software systems) and one for personality traits (e.g. Personality traits). The fact that many people consider soft skills as expertise should not be ignored even though there is no clear way to objectify them yet. These personal traits give a feeling of importance to the person stating them.

In an expertise profile, it is very important to add a section for hobbies and communities. As mentioned before, the hobbies of a person might be a "hidden gem" that would be a perfect match for the user's search query. The same applies to the communities' component, even though it will not be visible for non – insiders. Both sections aid in determining the social profile and the proficiency level of an expert.

5.2 Recommendation

The results presented in this report are one of the first steps on the road towards a successful EFS. The research discovered several topics of interest that need further investigation. Also, there are some concepts that need to be tested in order to discover their usefulness.

The first point of attention is the sample design. In section 1, general information, containing the name, photo and contact details, should not be visible right away. This should be displayed only

after the selection is made to ensure an objective, non – biased selection process. This is a way to guarantee that the most fitting expert is chosen to answer the query.

In addition, since there is no method to evaluate the experience, it is challenging to find a reliable source of knowledge and expertise. Therefore, more research is needed on this topic. The best possible way to gain more insight into this is to conduct case studies with real people.

Another way to determine whether or not the person is a real expert in the field he/she mentions is to add a free text box in the expertise section. By describing their knowledge and expertise in their own words, it can become evident, who really knows what they are talking about. The downside of this method is that it can be considered time-consuming. Therefore, it is advisable to make it an option for the user and not a necessity to complete the profile.

To determine which component and design work best for the profile tests are highly advised. EFS is a tool for people that aims to help them connect and establish an expert network. Therefore, a study such as this has to hold a high degree of people involvement by the means of case studies and focus groups. The idea presented in chapter 4.3.5 is one of many that can be conducted to test hypothesizes. However, for such a method a lot of time is needed.

The same applies to the case of expertise sources and evidence. This topic has already been started by another student and as a starting point the current design of an expert profile, presented in this report, was used. By presenting the design, insight into its generalization and understandability will be collected. That is also a type of testing of the profile.

To sum up, the recommended future actions to be taken towards the development of an EFS are (Figure 16):

- Develop several profile designs.
- Conduct case studies with the designs.
- Gather insight.
- Analyze insight.
- Make corrections to the design.
- Set up on 1 – 2 new designs.
- Conduct testing again.

Figure 16 Mapping of recommended future steps

It could be a challenge to decide when to stop the case studies and set up the final design. It should be kept in mind that the design could never be perfect and there would always be room for improvement. Nevertheless, there are some points that can be taken into account in the decision – making process:

- Stop testing when the deadline has been reached.
- Stop testing when the budget comes to an end.

Now, it is advisable to have an "acceptance criteria" to know if a design has passed or failed. Chapter 4.3 of this report presents some points for improvement and questions that can be used as a starting point for the first case study.

Another source for creating acceptance criteria could be the feedback from interviewees in the ongoing research into expertise sourcing.

6. Discussion

This is the final chapter of this research thesis report. This chapter aims to show how the research results contributed to answering the research questions. It also will state some effects of the findings and make suggestions for future research.

The topic of knowledge management has been around for decades, but not until recently the matter of Expert Finding systems emerged and grabbed the attention of companies. It is still an investigated topic, which leaves a lot of room for research and discoveries. So every research results provide useful insight and raise more questions.

The main question and its supporting sub-questions in the research discussed in this report are:

In what way should a knowledge profile be arranged to be useful for employees in the industry, who have a question, to be able to understand if that knowledge could be of help to them?

1. How to classify knowledge areas and skills?
2. What definition to use to accurately describe expertise?
3. How to make an expertise description that is recognized across companies?
4. What is a source of expertise evidence?

The results for each question are presented below.

6.1 Sub-question 1 “How to classify knowledge areas and skills?”

To answer this question, both the literature review and interviews were applied. This resulted in two main deliverables. A layout for a knowledge base of a software engineer and an example set of skills. The knowledge base was developed based on the insights of SWEBOK and the Computer science university curriculums (Saxion & UTwente, 2019) (IEEE Computer Society , 2014). The skills set resulted from numerous interviews with software engineers. The outcomes are as follows:

Knowledgebase

SWEBOK	
Software requirements	Software engineering process
Software design	Software engineering models and methods
Software construction	Software quality
Software testing	Software engineering professional practice
Software maintenance	Software engineering economics
Software configuration management	Computing foundation
Software engineering management	Mathematical and Engineering foundations

Computer Science program curriculum overview	
Hard skills	**Soft skills**
Computer systems Software systems Network systems Data and information Mathematical algorithms Programming paradigms	Project management Communication skills Personal development

Table 3 Overview of the curriculum for the Computer Science program (from chapter 4.1)

Skills

- Logical thinking and problem solving
- Analytical skills
- Good time and task management
- Quick learning ability
- Attention to detail
- Focus on continuous skills development

This result gives a guideline for the classification of knowledge areas and skills. After knowing that skills are perceived as the personal qualities that considered useful on the job, it is clearer how one could structure a skills component on an expert profile.

The application of this insight is apparent in the left side of the expertise section on the profile in Appendix 7. What should be noticed is the order and terminology of the knowledge and skills of a software engineer. The terminology of the Computer science curriculum was found to be generally understandable by software engineers. And the "Project experience" and "Other experience" were influenced by the information gotten from SWEBOK. More particularly, SWEBOK showed the importance of projects and that it has an influence on one's expertise.

6.2 Sub-question 2 "What definition to use to accurately describe expertise?"

When it comes to question number 2 the most valuable result is shown in the table below:

EXPERTISE PERCEPTION TYPE	N OF PEOPLE WITH THAT PERCEPTION[2]	
TECHNICAL EXPERTISE	*32*	*53.3 %*
SOFT SKILLS EXPERTISE	*8*	*13.3 %*
HYBRID EXPERTISE	*13*	*21.7 %*

Table 4 Views on the types of perception of expertise (from chapter 4.1)

It was revealed that people in the software engineering industry have 3 different types of perception when it comes to their expertise. As it can be expected the most common one was technical expertise. However, the occurrence of soft skills and hybrid expertise was a valuable insight when deciding on components for the profile. Also, this result raises the possibility that most people would fall into one of the three types. Thus, this finding could be considered generally applicable. However, more research into different departments is needed to fully support this statement.

This highly influenced the design of the profile (Appendix 7). Because of this insight the component "Personality traits" was added. It should be noted that this question is tightly connected with question 1. Knowledge and expertise are the backbones of an expert profile and an EFS. Thus, after setting up the knowledge components, with the help of this insight the expertise component was established. This was done by including components that address both technical and soft skills knowledge and expertise.

[2] 11.7 % (number of 7) of participants' responses were not applicable to this table

6.3 Sub-question 3 "How to make an expertise description that is recognized across companies?"

This question could be considered the most challenging in the whole research because it concerns the perceptions of many people across companies. Due to the limitation of time, it was not possible to gather sufficient field data from people of different companies.

However, a different non - intrusive approach was taken. It was decided that in order to create a generalized environment that could be understood by most, the following should be initiated:

- Definitions of terms
- Proficiency levels
- Names of department and professions should be generalized

The definitions would make sure that each user is presented with information that should aid in understanding the expertise presented in the profile. Thus, help to make the right choice of an expert that will solve the user's issue. A term is considered each topic mentioned in the profile such as Java, UML, Scrum, etc. And upon mouse click, a definition of that term will be displayed (Figure 15).

In addition, the proficiency levels are also considered a term, that will have a definition. The levels that were established are Basic, Limited working proficiency, Intermediate, Full professional proficiency. The development of the proficiency levels was stimulated by the theory that an expert profile is composed of a topical and social profile (Balog & Rijke, 2007). Therefore, the proficiency levels addressed the experience of the expert and the number of references he/she could make in regards to the topic of expertise. The more experience and the bigger the network of an expert, the higher is the level of proficiency.

The next activity towards profile generalization was the simplification of job titles and departments. Considering the number of job titles and the individuality of the departmental names across companies, it is only wise to create a general list. That way users can decide for themselves which title and name apply to them the most and will be understood by other users. The visualization of this activity can be seen in Figure 9 and 10.

The implication of these methods of generalization needs to be tested by the means of case studies and focus groups. This is the only way to determine if the profiles are understandable by everyone.

6.4 Sub-question 4 "What is a source of expertise evidence?"

When it comes to this question, the research discovered what unreliable sources of expertise are. The sources that were investigated were:

- Project documentation
- Project's final product
- Other's feedback and opinion
- Educational background

There was not enough valid proof that these sources could be considered adequately valid to be included in the EFS. The deliverables during projects are a team – effort and even though sometimes there are comments on the programming code, that kind of feedback is insufficient for sourcing expertise. Also, the feedback from fellow employees is too biased to be considered reliable and the educational background validity fades away with time.

Therefore, it was decided that a separate project will be conducted that specifically focuses on the sourcing of expertise by conducting field research in other companies.

Another method of expertise verification that could be considered is scoring. The user could score the expert after he/she has received an answer. However, such an approach is also subjective and it is very likely that a user could score based on emotion and not on the objective evaluation of the received answer.

6.5 Final thoughts and main research question

The data that was gathered for the types of perceptions of expertise was an eye-opener. The result of the three types of perception was not completely unexpected, but the data creates a strong foundation that the development of the profile can only benefit from.

Therefore, the final design presented in Figure 14 and Appendix 7 would serve as a good start for the projects to come. The sections presented in this report: General information, Expertise, Hobbies, and communities provide a base for future projects in the topic of EFS. One being the

sourcing of expertise project that is already ago. The results of which will definitely affect and alternate the current design. This is certain because the current design will be presented to the interviewees in the next projects and their feedback and advice on the design will be valuable input on the EFS project.

To sum up, the input from this research into knowledge and expertise provided useful information for the generalization of an expert profile. Defining knowledge and expertise was a crucial step, which helps to set a solid base to build on and develop an Expert finding system.

It is important to keep the two main incentives for companies to adopt an EFS: (A) the need to decrease cost; (B) the need to increase the know-how of employee in mind. Especially when the EFS system is being presented in a company. Inefficient communication channels are causing a lot of useful products to stay unknown and unused. Therefore, to ensure the success of an EFS, it is crucial to make sure as many employees as possible know about its development and launch. Because of the system's focus on connecting people. Thus, without people's involvement, the system could not a success.

From here on, the following research in expertise sourcing and the testing of expert profile designs will continue to give material for improvement, which eventually will lead to the realization of an Expert finding system. With that system, a network of expertise within and across companies could be established, which in return could drastically improve the communication channels on the workplace.

Bibliography

Agresti, W. W. (2000). Knowledge Management. *Advances in Computers*, 171-283.

Al-Taie, M. K. (2018, August 22). Understanding expert finding systems: domains and techniques. *Social Network Analysis and Mining*, 57.

Balog, K., & Rijke, M. d. (2007). Determining Expert Profiles (With an Application to Expert Finding). *IJCAI*, 2657-2662.

Baltadzhieva, A. &. (2015). Question Quality in Community Question Answering Forums: a survey. *SIGKDD Explorations 17*, 8-13.

Becerra-Fernandez, I., & Sabherwal, R. (2014). *Knowledge Management: Systems and Processes.* Routledge.

Centro Universitario Internazionale. (2018, February 22). *SKILL VIEW TEST*. Retrieved from originalskills.com: https://originalskills.com/en/skill-view/

COE. (2019, May 8). *Common European Framework of Reference for Languages*. Retrieved from www.coe.int: https://www.coe.int/en/web/common-european-framework-reference-languages

Davenport, T. H., & Prusak, L. (1998). *Working Knowledge: How Organizations Manage What They Know.* Cambridge: Harvard Business School Press.

Ehrlich, K., Lin, C.-Y., & Griffiths-Fisher, V. (2007). Searching for experts in the enterprise: combining text and social network analysis. *GROUP '07 Proceedings of the 2007 international ACM conference on Supporting group work* (pp. 117-126). Florida: ACM New York.

Employee, 1. (2019, March 18). Private communication. (S. Lefterova, Interviewer)

Employee, 2. (2019, March 26). Private communication. (S. Lefterova, Interviewer)

Employee, 3. (2019, March 20). Private communication. (S. Lefterova, Interviewer)

Employee, 4. (2019, March 26). Private communication. (S. Lefterova, Interviewer)

Employee, 5. (2019, February 20). Private communication. (S. Lefterova, Interviewer)

Good Calculators. (2019, March 15). *Sample size calculator*. Retrieved from goodcalculators.com: https://goodcalculators.com/sample-size-calculator/

Heisig, P. (2009). Harmonization of knowledge management - comparing 160 KM frameworks around the globe. *Journal of Knowledge Management*, 4-31.

IEEE Computer Society . (2014). *Guide to the Software Engineering Body of Knowledge.* Piscataway: IEEE.

Latham, J. R. (2016). *Research design canvas: A framework for designing and aligning the "DNA" of your research study (Version 2.0 ed.).* Colorado Springs: Colorado: Organization Design Studio™ Ltd.

Latham, J. R. (2019, April 3). *Conceptual framework*. Retrieved from www.drjohnlatham.com: https://www.drjohnlatham.com/frameworks/research-methods-framework/conceptual-framework/

Lin, S., Hong, W., Wang, D., & Li, T. (2017). A survey of expert finding techniques. *Intell Information Systems*, 49:255–279.

Merriam-Webster. (2019, March 5). *Dictionary Merriam-Webster*. Retrieved from https://www.merriam-webster.com: https://www.merriam-webster.com/dictionary/knowledge

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create The Dynamics Of Innovation.* Oxford: Oxford University Press, Inc.

Novak, J. D., & Cañas, A. J. (2008). *The Theory Underlying Concept Maps and How to Construct and Use Them.* Florida: Institute for Human and Machine Cognition.

Prahalad, C. K., & Hamel, G. (1990). The Core Competence of the Corporation. *Harvard Business Review, Vol. 68, Issue 3*, p. 79-91.

Riege, A. (2005). Three-dozen knowledge-sharing barriers. *Journal of knowledge management*, 18-35.

Rus, I., Lindvall, M., & Sinha, S. S. (2001). *Knowledge management in Software engineering.* Rome: DoD Data & Analysis Center for Software.

Saxion University of Applied Sciences . (2019, May 8). *Applied Computer Science: Course content*. Retrieved from www.saxion.edu: https://www.saxion.edu/programmes/bachelor/applied-computer-science/course-content

Select Statistical Services Limited. (2019, March 15). *Sample size calculator*. Retrieved from select-statistics.co.uk: https://select-statistics.co.uk/calculators/sample-size-calculator-population-proportion/

Shipman, J., R.A. Ash, M. B., Carr, L., Eyde, L., Kehoe, B., Pearlman, J., . . . Sanchez, J. (2000). The practice of competency modeling. *Personnel Psychology*, 703 -740.

Spencer, L., McClland, D., & Spencer, S. (1994). *Competency assessment methods: History and state of the art.* Boston: Hay -McBer Research Press.

Stowe, L. (2018, September 8). *Programmer vs Developer*. Retrieved from devskiller.com: https://devskiller.com/programmer-vs-developer/

Thales. (2018, December 5). *Discover Thales*. Retrieved from www.thalesgroup.com: https://www.thalesgroup.com/en/career/discover-thales

Tilburg University. (2019, February 14). *Experts & Expertise*. Retrieved from www.tilburguniversity.edu: https://www.tilburguniversity.edu/webwijs/

Typeform. (2019, May 2). *Survey data analysis*. Retrieved from www.typeform.com: https://www.typeform.com/surveys/survey-data-analysis/

UTwente. (2019, May 8). *Technical computer science study program*. Retrieved from www.utwente.nl: https://www.utwente.nl/en/education/bachelor/programmes/technical-computer-science/study-programme/

UTwente. (2019, May 2). *Technical computer Science: Study program*. Retrieved from www.utwente.nl: https://www.utwente.nl/en/education/bachelor/programmes/technical-computer-science/study-programme/second-year/

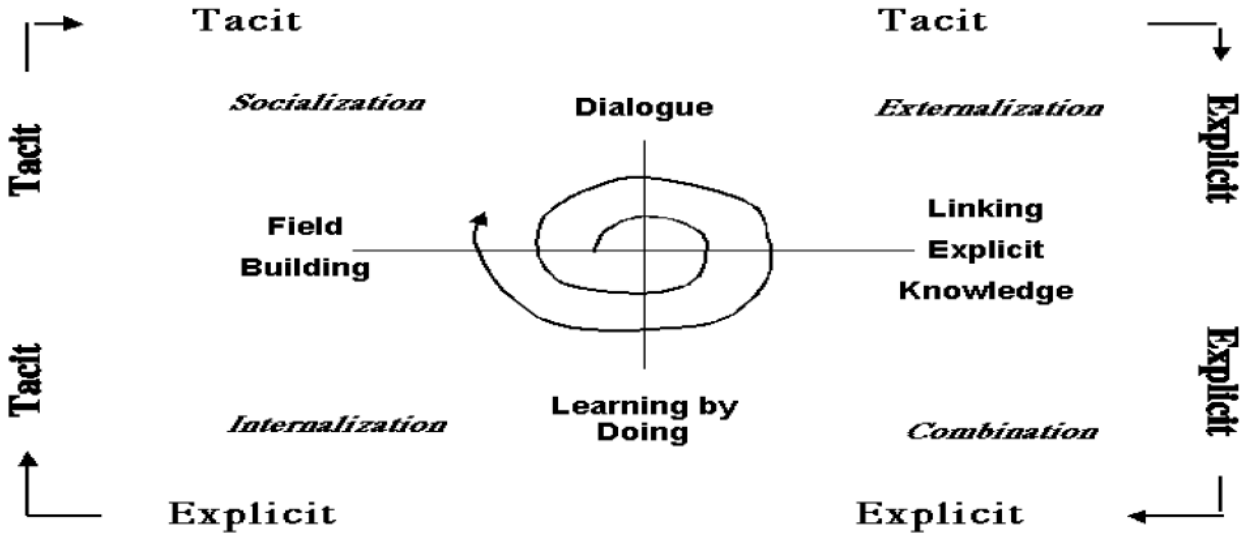
Wang, G. A., Jiao, J., Abrahams, A. S., Fan, W., & Zhang, Z. (2013). ExpertRank: A topic-aware expert finding algorithm for online knowledge communities. *Decision Support Systems*, 54(3):1442–1451.

Yimam-Seid, D., & Kobsa, A. (2009). Expert-Finding Systems for Organizations. *Journal of Organizational Computing and Electronic*, 1-24.

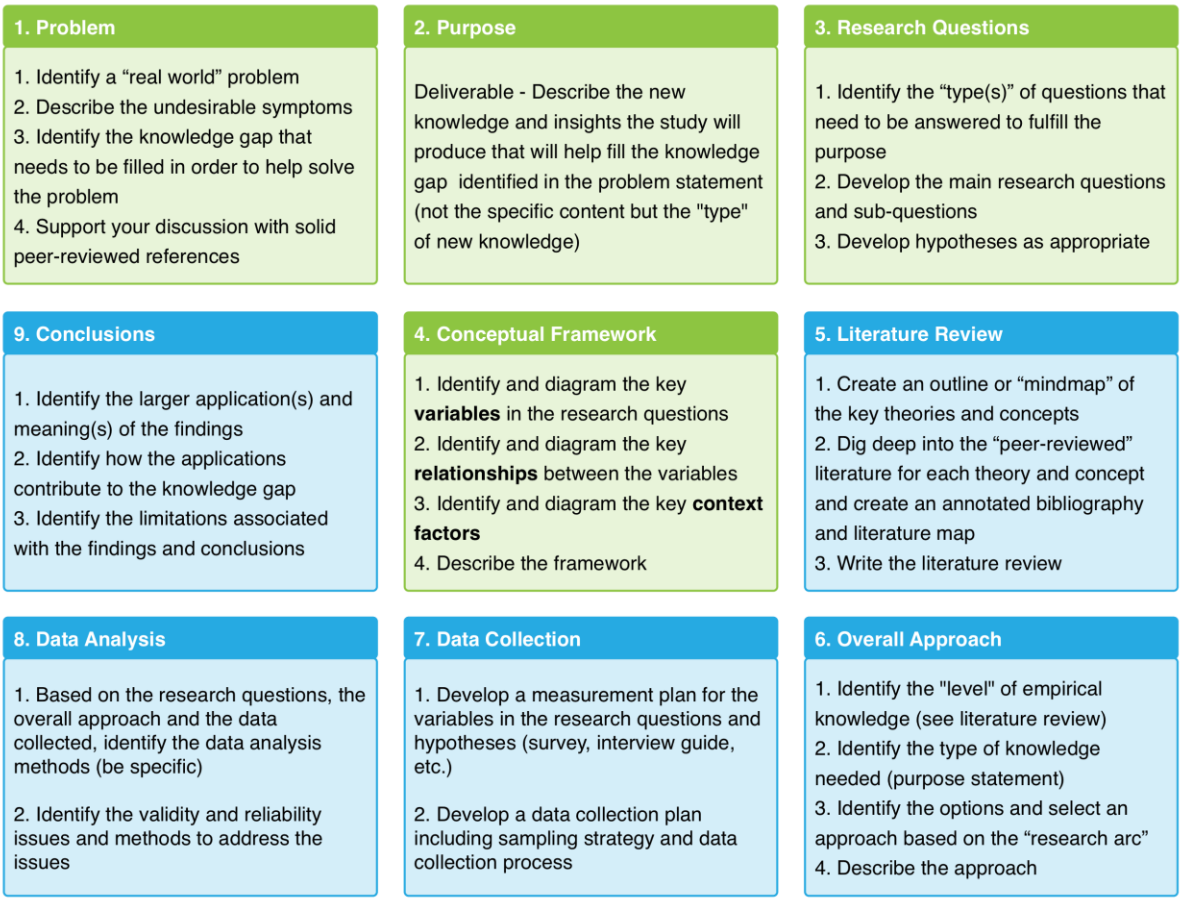
Zhang, J., Tang, J., & Li, J. (2007). Expert Finding in a Social Network. *DASFAA.*

Appendices

Appendix 1 Knowledge spiral by Nonaka and Takeuchi (1995)



Appendix 2 Research canvas by J.R. Latham



Latham, J. R. (2014) *The Research Canvas: A Framework for Designing and Aligning the "DNA" of Your Study.* Leadership Plus Design, Ltd.

Appendix 3 Interview questions

Primary goal: Understand the knowledge and expertise SEs apply in the workplace

Secondary goal: Gain perspective on the knowledge sharing mentality of SEs

1. Would you describe the expertise you apply to your daily tasks within the organization?

2. If you were to describe your knowledge with five words, what would you say?

   2.1 Did you consider any of the knowledge areas you acquired from your previous education?

   2.2 Are there any SE knowledge areas that are the most popular in the SE community?

3. If you were to describe your skills with five words, what would you say?

4. What is some new knowledge that you have acquired from your time at Thales?

5. What kind of documents are you producing during projects?

   5.1 Which of those documents would you refer to when starting a new project?

   5.2 Which of those documents do you consider to be evidence of expertise?

6. What kind of documents do you produce as an individual?

7. What kind of teams have you worked with? Only SE or interdisciplinary

8. Would you say software engineers are willing to share knowledge with each other?

   8.1 Do you consider knowledge sharing to be beneficial to the SE community within Thales?

   8.2 What is the current state of knowledge sharing b/w SE within Thales?

9. Have you ever received 360-degree feedback?

10. Do you have any advice on where can I find information about the knowledge areas and skills in SE?

11. Do you know anyone who I can reach out to for more information?

Appendix 4 Survey questions

Expert profiling

Welcome! Thank you for your participation in this survey. Your input is much appreciated. The questions here aim to address the nature of your expertise and knowledge. This information will be used in creating a sample expert profile that will be part of a future expert finding system within Thales.

There are 10 questions in this survey

Knowledge profiling

1 Please describe your day - to - day responsibilities in short (specify department) *

Please write your answer here:

2 To what product are your responsibilites connected to: *

Please choose all that apply and provide a comment:

- ☐ Radar
- ☐ Sensors
- ☐ Tacticos
- ☐ IFF systems
- ☐ INFRA/Platform
- ☐ Other

If needed please specify the area in the comment section

3 What knowledge do you have of software engineering languages ? *

Please choose the appropriate response for each item:

	Begginer	Intermediate	Expert	Not applicable
Java	○	○	○	○
C++	○	○	○	○
C	○	○	○	○
ADA	○	○	○	○
Python	○	○	○	○
JavaScript	○	○	○	○
Bash	○	○	○	○
Csharp	○	○	○	○
HTML	○	○	○	○
CSS	○	○	○	○

Please indicate your level of proficiency

4 What about tools and frameworks? *

Please choose the appropriate response for each item:

	Begginer	Intermediate	Expert	Not applicable
Modbus	○	○	○	○
Jenkins	○	○	○	○
MatLab	○	○	○	○
Camel	○	○	○	○
Spring	○	○	○	○
Active MQ	○	○	○	○
Maven	○	○	○	○
Tefares	○	○	○	○
Akeria	○	○	○	○
O2	○	○	○	○
O2N	○	○	○	○
Devon	○	○	○	○
IDE Environment	○	○	○	○

5 And internal tools? *

Please choose the appropriate response for each item:

	Begginer	Intermediate	Expert	Not applicable
Jira	○	○	○	○
Twiki	○	○	○	○
Engwiki	○	○	○	○

6 What is your educational background? *

Please choose all that apply:

- ☐ Computer Science
- ☐ Electrical engineering
- ☐ Software engineering
- ☐ ITSM
- ☐ Other:

7 Have you directly applied your education to your job? *

Please choose only one of the following:

- ○ Yes
- ○ No

8 Do you have any other knowledge and/or expertise that you haven't mentioned thus far? Not necessarily connected to Thales *

Please choose only one of the following:

- ○ Yes
- ○ No

Make a comment on your choice here:

If Yes please specify in the comment

9 What do you consider yourself to be an expert in? *

Please write your answer here:

10 Are you a member of any of the Thales Online Software engineering communities? *

Please choose only one of the following:

- ○ Yes
- ○ No

Make a comment on your choice here:

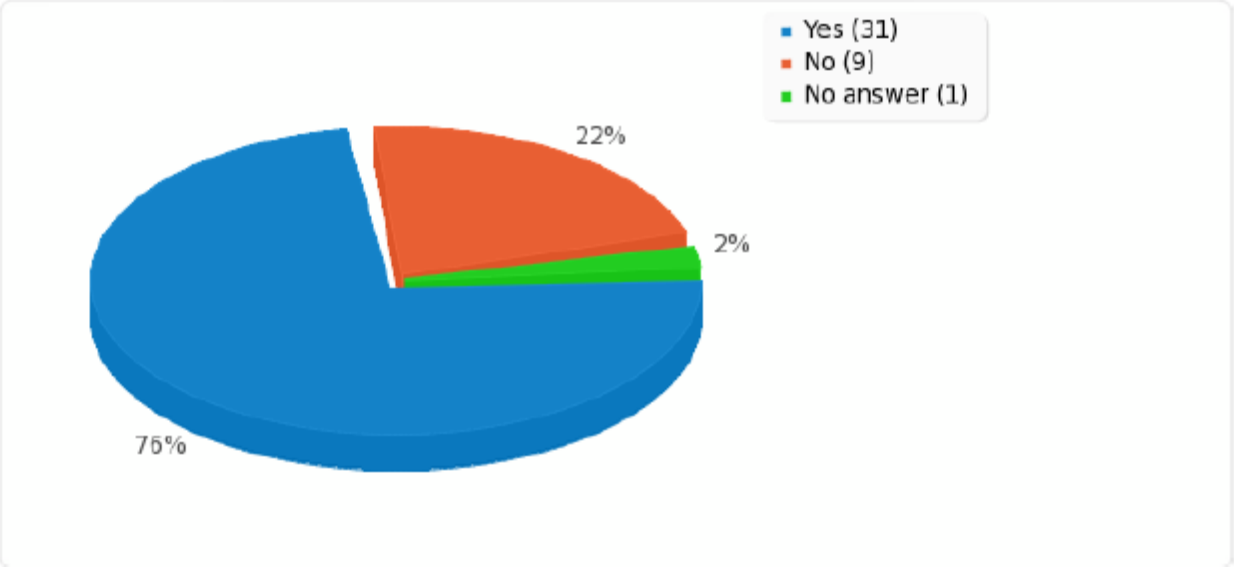
If Yes please specify in the comment

Appendix 5 Example of a summary of responses from the survey

Field summary for 07

Have you directly applied your education to your job?

Answer	Count	Percentage
Yes (Y)	31	75.61%
No (N)	9	21.95%
No answer	1	2.44%

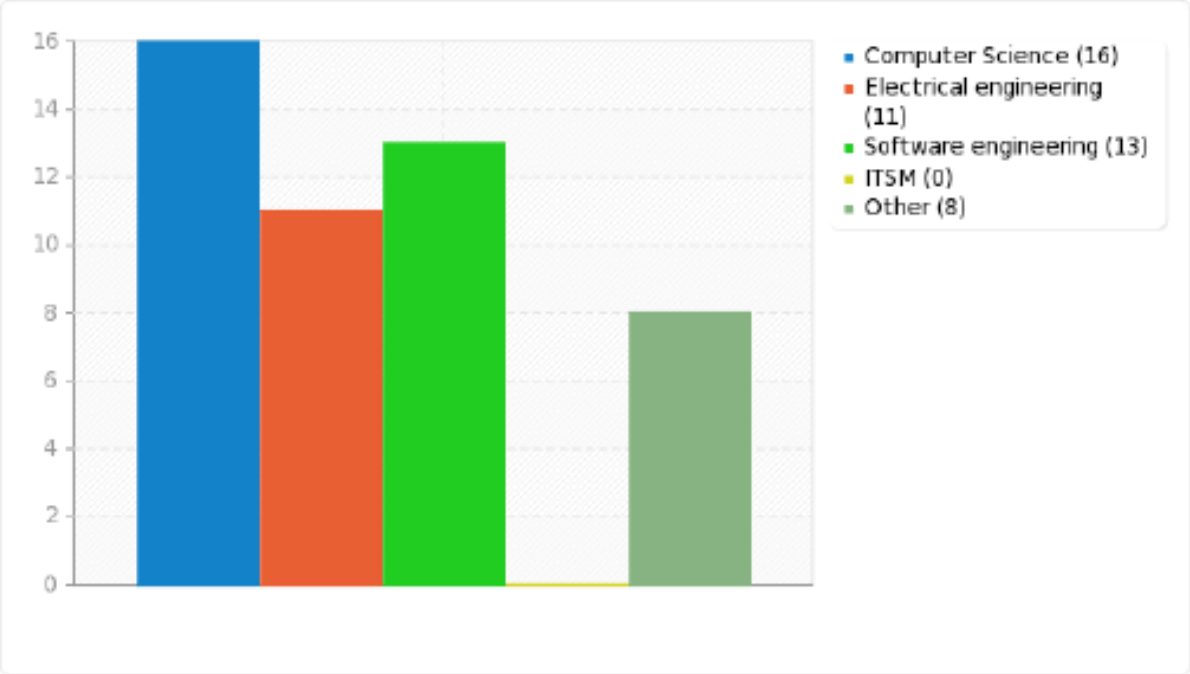


Appendix 6 Educational background of survey participants

Field summary for Q6

What is your educational background?

Answer	Count	Percentage
Computer Science (SQ001)	16	39.02%
Electrical engineering (SQ002)	11	26.83%
Software engineering (SQ003)	13	31.71%
ITSM (SQ004)	0	0.00%
Other	8	19.51%



Appendix 7 An expert profile: software engineer

ROY MANDELL

+31628374922

Roy.mandell@nl.thalesgroup.com

HENGELO, THE NETHERLANDS

PROFESSION Software engineer

DEPARTMENT: Software development

CURRENT TEAM: Team Visual

EXPERTISE

SOFTWARE SYSTEMS	PROFICIENCY LEVEL
Design: *UML*	*Limited working proficiency*
Programming: *Java*	*Full professional proficiency*
C++	*Intermediate*
Testing: *Cucumber*	*Intermediate*
NETWORK SYSTEMS AND PROTOCOLS *Internet; GSM/UMTS*	*Basic*
OPERATING SYSTEMS *Linux*	*Intermediate*
DATA AND INFORMATION *SQL/ PL-SQL*	*Limited working proficiency*
SOFTWARE MANAGEMENT TOOLS *Maven*	
PROJECT EXPERIENCE *Sensors* *Click or tap here to enter text.*	
OTHER EXPERTISE *Big data; Designing android applications*	PERSONALITY TRAITS *Detail – oriented ; Able to motivate people towards a common goal*

HOBBIES AND INTERESTS

Beginner course in ADA

Karting

THALES ENGINEERING COMMUNITY MEMBER OF

Apache, Celix, ASF, Kubernetes, ISTIO