OULUN AMMATTIKORKEAKOULU

Maarit Martikainen

CREATION OF AGILE TERMINOLOGY MANAGEMENT PROCESS AND TOOL FOR BILINGUAL AGILE SOFTWARE COMPANY

CREATION OF AGILE TERMINOLOGY MANAGEMENT PROCESS AND TOOL FOR BILINGUAL AGILE SOFTWARE COMPANY

Maarit Martikainen
Final thesis
Spring 2018
Master's Degree Programme in
Industrial Management
Oulu University of Applied Sciences

ABSTRACT

Oulu University of Applied Sciences
Master's Degree Programme in Industrial Management

Author: Maarit Martikainen
Title of thesis: Creation of Agile Terminology Management Process and Tool for Bilingual Agile Software Company
Supervisor: Jukka Jauhiainen
Term and year when the thesis was submitted: Spring 2018 Number of pages: 87 + 9

The purpose of this thesis work was to create a terminology solution for the needs of one department of an agile software company. The solution consisted of a process for developing and maintaining terminology in the case company as well as a basis of a terminology for the needs of the company. This thesis was a fundamental study in the case company because there was no terminology management process in use in the case department before. The case company department operates in the software field of healthcare and social welfare in Finland. The healthcare field, the social welfare field as well as the software development field are all very information-centric and use specialized languages both in Finnish and in English. The healthcare and social welfare fields have been under heavy national development in Finland for years and the common terminology is critical for the success of such development efforts.

This thesis project succeeded well in fulfilling the requirements set to it. The requirements were collected from end-users and terminology literature to guarantee a future-proof solution. An agile terminology management process was created and developed as well as a tool for the glossary. This project used agile development methods and one-week iterations. The complete solution was fully compliant with 24 of the set 27 requirements and two were partially compliant due to the brief usage period. The few requirements that were not fully compliant yet will be developed further by the Glossary Community of Practice (CoP) that was established in the case company at the end of the project to manage the solution in the future. The creation of the Glossary CoP sends a strong signal that the work is very much needed, and it will be continued in the case company. The Glossary CoP has responsibilities with the solution; the processes and the tool, but also with making the terminology work and the glossary known in the company. In the future, the case company could perform an end-user study to estimate Return on Investment (ROI) for the terminology work.

There is a great deal of changes currently happening at the national level of healthcare and social welfare and it also has implications to both technical and functional requirements of software operating in those fields. The existing national terminology is expanding and changing rapidly, so the existing terminological resources require updating. The traditional terminology work processes described in literature were considered slow for the quickly changing environment of software development. The agile process developed in this project provides a flexible solution for the case company to cope in the constantly changing situation.

Keywords: terminology, management, process, agile methods, software development

TIIVISTELMÄ

Oulun ammattikorkeakoulu
Master's Degree Programme in Industrial Management

Tekijä: Maarit Martikainen
Opinnäytetyön nimi: Ketterän terminologian hallinnan prosessin ja työkalun luominen ketterälle kaksikieliselle ohjelmistoyritykselle
Työn ohjaaja: Jukka Jauhiainen
Työn valmistumislukukausi ja -vuosi: kevät 2018 Sivumäärä: 87 + 9

Tämän opinnäytetyön tavoite oli luoda ratkaisu terminologian hallintaan erään ketterän ohjelmistoyrityksen osaston tarpeisiin. Ratkaisu sisälsi prosessin terminologian kehittämiseen ja ylläpitämiseen yrityksessä sekä sanaston pohjan yrityksen tarpeisiin. Tämä projekti oli perustavaa laatua yrityksessä, koska kyseisellä osastolla ei ollut aiemmin käytössä prosessia terminologian hallintaan. Kyseinen yrityksen osasto toimii sosiaali- ja terveydenhuollon ohjelmistoalalla Suomessa. Sosiaali- ja terveydenhuoltoalat sekä ohjelmistokehitysala ovat kaikki hyvin tietopainotteisia ja käyttävät erikoistunutta kieltä sekä suomeksi että englanniksi. Sosiaali- ja terveydenhuoltoalat ovat olleet voimakkaan kansallisen kehityksen kohteena vuosia ja yhteinen termistö on elintärkeä kyseisenlaisten kehitysponnistelujen onnistumiseksi.

Tämä opinnäytetyöprojekti onnistui täyttämään sille asetetut vaatimukset hyvin. Vaatimukset kerättiin loppukäyttäjiltä sekä terminologiaan liittyvästä kirjallisuudesta, jotta ratkaisu olisi aikaa kestävä. Projektissa luotiin ja kehitettiin ketterä terminologian hallinnan prosessi sekä työkalu sanastolle. Projekti noudatti ketteriä kehitysmenetelmiä ja käytti yhden viikon iteraatioita. Ratkaisu täytti lopulta täysin 24 vaatimusta asetetuista 27:stä ja kaksi täyttyi lisäksi osittain johtuen lyhyestä käyttöajasta. Noiden muutaman vajaaksi jääneen vaatimuksen kehittämistä jatkaa yritykseen projektin lopussa perustettu Sanastotyöryhmä. Sanastotyöryhmä hallinnoi terminologiaratkaisua jatkossa ja sen perustaminen on selkeä viesti siitä, että työ on todella tarpeellista ja sitä jatketaan yrityksessä. Sanastotyöryhmä vastaa ratkaisusta prosesseineen ja työkaluineen, mutta myöskin sanastotietoisuuden levittämisestä yrityksessä. Tulevaisuudessa yritys voisi myös tehdä loppukäyttäjätutkimuksen terminologiatyön tuottoarvioiden laskemiseksi.

Sosiaali- ja terveydenhuoltoalalla tapahtuu tällä hetkellä paljon muutoksia kansallisella tasolla ja niillä on vaikutuksia kyseisillä aloilla käytössä oleviin ohjelmistoihin sekä niiden teknisiin ja toiminnallisiin vaatimuksiin. Olemassa oleva kansallinen terminologia laajenee ja muuttuu nopeasti, joten olemassa olevat sanastotkin vaativat päivittämistä. Perinteiset kirjallisuudessa kuvatut terminologiatyön prosessit arvioitiin liian hitaiksi nopeasti muuttuvaan ohjelmistokehityksen ympäristöön. Tässä projektissa kehitetty ketterä prosessi tarjoaa yritykselle joustavan ratkaisun selvitä jatkuvasti muuttuvassa tilanteessa.

Asiasanat: terminologia, hallinta, prosessi, ketterät menetelmät, ohjelmistokehitys

PREFACE

I selected this topic because I would have needed a glossary to proceed with other final thesis topics in English in my area of expertise and I could not find a good one. This thesis is grounding work for others in this area. This topic is very well suited for me because I have done all my software related studies in English and worked for ten years in software development using English as the main language. For several recent years, I have worked in a Finnish healthcare software company in Finnish. In this thesis work I can combine my experience and build up my English proficiency in healthcare related software at the same time as helping others to cope in this challenging time of change in the healthcare industry.

I want to thank my family and colleagues who supported me through this process.

June 5. 2018, Oulu
Maarit Martikainen

CONTENTS

1 INTRODUCTION

The writers of the book “Handbook of Terminology Management” from year 1997 reason the use of word “Management” in the book title like this:

> *The emphasis on management underscored in these designations also reflects the efforts of manufacturers as well as terminology, documentation, and information specialists to come to terms with the contemporary information explosion, that is, to manage and control ever-increasing volumes of information. In quality management, for instance, much of what was once undefined chance or accident in manufacturing systems is now manifested as quantifiable data that can be manipulated to ensure control capability with respect to production results. These data constitute information, but the information does not just flow through database management systems like water in a pipe. It is carried by discrete, identifiable, retrievable vehicles called data elements. These data elements are rendered identifiable in that they have names, and the naming of data elements in information systems is an essentially terminological act. In fact, the harmonization of data element names and data structures is one of the most critical challenges facing information specialists today in their effort to create global systems that can "talk to each other" across hardware and software barriers. Consequently, the link between terminology management, information management, and quality management is one of intimate interdependence. (1 pp. 2-3.)*

The purpose of this thesis work was to create a terminology solution for the needs of one department of an agile software company. The solution consisted of a process for developing and maintaining terminology in the case company as well as a basis of a terminology for the needs of the company. The case company department operates in the software field of healthcare and social welfare in Finland. The healthcare field, the social welfare field as well as the software development field are all very information-centric and use specialized languages both in Finnish and in English. The healthcare and social welfare fields have been under heavy national development in Finland for years and a common terminology is critical for the success of such development efforts.

The terminology solution was needed in the case company because more and more national requirements apply to Finnish healthcare and social welfare software. The number of these national requirements increase due to the growing demand for system interoperability. Some of the terms in the new national requirements are similar or overlap with existing ones in the software, but the terms do not necessarily mean the same thing. Some of the used terms are defined in Finnish in some of the national specifications but glossary is not a standard section of national specifications. Care is needed to achieve the wanted operation of the software in these situations and to create the same understanding of the requirements within the company.

The requirements for the case company terminology solution were gathered from end-users in 2017 and 2018 through interviews and from terminology work related literature in the beginning of 2018. The process and tool development required studying theoretical and practical terminology work because the field of study was rather unfamiliar to the researcher. Information regarding the existing dictionaries used in the company as well as the previously created company and project specific collections of words were studied during the preliminary phase of the thesis work in 2017. The solution was developed iteratively between March and May 2018 using agile development principles to get it into use immediately and to receive feedback from the end-users as much as possible. The implementation phase of the project ended in May 2018. The national effort of terminology work related to national information management in the healthcare and social welfare sector was also participated in the beginning of 2018.

The next chapter describes the case company and its business environment. Also, the language situation in both the software industry and the healthcare sector are presented to give the reader a broader view to the current and future changes in the industry. The literature study section of this report focuses in terminology in theory and in practice to give the reader an idea of the field of study and the extent of it. Also, scaled agile methodologies are introduced briefly to provide an understanding of the operating environment. Next, the needs and requirements collection from the end-users is described and solution requirements from the literature study are collected. These requirements are used to compare existing practical solutions. In the implementation section, a suitable solution for the case company is developed and described. The solution is evaluated against the requirements, and finally conclusions are drawn, and further development suggestions are given.

2 BACKGROUND

This chapter draws the bigger picture where this thesis work is situated. First, the case company and the business it operates in are described. Then the current language situation in both the software industry and the healthcare and social welfare sector in Finland are presented. This information is needed to define the focus of the literature study.

2.1 Case Company

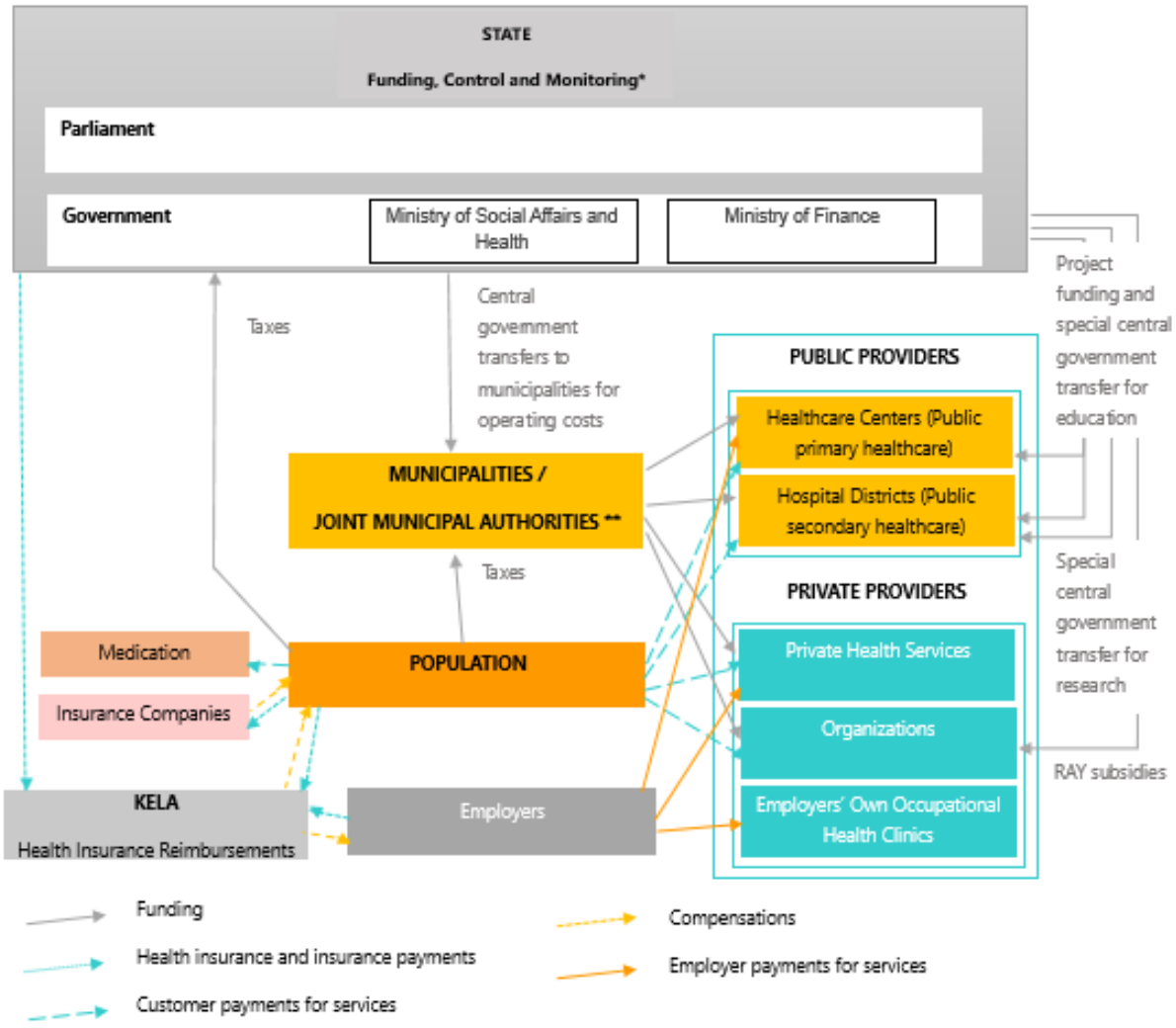
This study is done for a large international company and a department there that operates in domestic health and social welfare software industry. The company has a history of over 40 years in Finland including several acquisitions and mergers (2). The health and welfare department employ over 350 people in Finland currently (3). The company is significant in the domestic healthcare sector because their products keep electronic health records of over three million Finns (4). The company develops several specialized products for use in primary healthcare, secondary healthcare, private healthcare and in social welfare. Each of these health and welfare sectors have traditionally been differentiated by working methods and software needs, so that they have specialized software for patients' and customers' administrative as well as service and health data recording.

The case company uses scaled agile methodologies in their software development. Agile software development is becoming a standard in software industry, but it does not fit very well for large companies. Therefore, a scaled version of agile methodologies has been created where development efforts of several teams can be coordinated. This methodology is described in more detail in the literature study section of this report.

2.2 Business Environment

The health and social welfare sector in Finland has been a constant topic of discussions both for politicians as well as for the general population for several years now because of the much needed and prepared reform. The years of preparation time has brought uncertainty to the business environment when for example investments in the public health and social welfare sector have

been restricted with a special temporary law (5) that had to be amended (6) as well when the original schedules could not be kept. The reform is in a critical phase currently when several laws regarding it are to be passed this spring (7). One of the main objectives for the reform is simplifying the complex financial structure of health and social welfare (8). Figure 1 describes the current funding structure of the healthcare sector in Finland. The structure for social welfare is different to some extent but not simpler. A more detailed description of the current financing of health and social welfare sector can be found in a report (in Finnish) drawn during the preparation of the reform (9).



* Healthcare is monitored by Regional State Administrative Agencies, National Supervisory Authority for Welfare and Health (Valvira) and Finnish Medicines Agency Fimea. Information guidance is the responsibility of specialist institutions National Institute for Health and Welfare (THL), Finnish Institute of Occupational Health (FIOH) and Radiation and Nuclear Safety Authority, Finland (STUK).

** Municipalities are responsible for organizing health services to their population. Primary healthcare must be organized in municipalities of at least about 20 000 people or in a joint municipal authority. To fulfill the organization responsibility of secondary healthcare, a municipality must belong to a hospital district.

FIGURE 1. Organization, funding, provision and monitoring of health services (10)

The customers of health and social welfare software are the public and private provider organizations shown in Figure 1. This group and their financing is changing dramatically in the reform and it inevitably influences the companies operating in this specific software sector. The National Institute for Health and Welfare describes the reform briefly (11):

> *An overhaul of the structures of the social welfare and health care services system has been going on in Finland for several years. The need for this reform emerged from problems in ensuring equal and adequate social welfare and health care services for the population under the existing municipality-based service structure as the dependency ratio changes. Small and financially weak municipalities have encountered significant difficulties in organising and producing services. In the present reform, responsibility for providing social welfare and health care services is being transferred to larger and hence stronger administrative entities.*
>
> *The Government has outlined the creation of autonomous areas for the purpose of organising social welfare and health care services. The objective in this operation is not only to create financially more viable bodies as service organisers, but also to achieve complete horizontal and vertical integration of social welfare and health care services.*

The current schedule for the reform is described in Figure 2. There has been years of preparation and delays in the reform by now and the schedules have been changed. The figure shows that there are several critical steps still to take later this year to keep the current target of having the main part of the reform done in 2020.

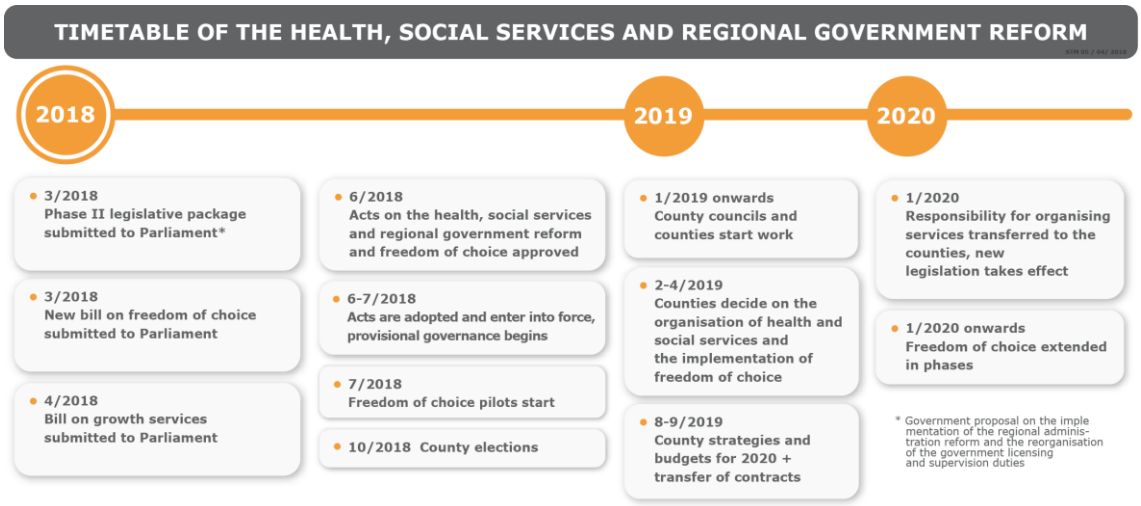


FIGURE 2. Current schedule for the health and social welfare reform (7)

There is a considerable number of regulations in European Union level as well as in the national level that apply to healthcare and social welfare as well as the software that is used in providing the services. For example, the European Union Directive for Medical Devices applies also to electronic health records software (12). An example of a national level regulation is the Act on

Electronic Handling of Customer Data in Health and Welfare (13). The national regulations vary between countries even with the increasing number of European Union regulations that apply to healthcare and social welfare.

Software industry in general is rather global, but the country-specific regulations have for the most part prevented globalization in health and social welfare software industry and even the Finnish acts and decrees have not been translated to English for the most part (14). For example, the Act on Electronic Handling of Customer Data in Health and Welfare does not have an English translation, but it specifies in Section 19c that the information and instructions meant for the health and welfare staff using an information system must be in Finnish and Swedish (13). In addition, all the national requirements imposed on health and welfare software are only available in Finnish (15). These kind of language regulations as well as the current distinctive organization and financing structure (16) have kept the software used in health and social welfare domestic for the main part. The previously described reform will have an influence in this as well when the organization and financing structure of the services will become more like those in other European Union countries.

There are also other efforts to standardize healthcare and social welfare internationally besides the European Union regulations. These efforts have been going on for centuries internationally and have managed to introduce more structure and standard ways of documenting for example diagnoses by doctors and nurses. This started already in 1893 when the International Statistical Institute adopted the International List of Causes of Death (17). The World Health Organization (WHO) became responsible for maintaining it when the organization was created and the first version to include also morbidity was published in 1948 (17). This was called the international classification of diseases (ICD) version 6 (17). Since 1990 the version is ICD-10 and it is used in over 100 countries globally (17). In Finland ICD-10 was taken officially into use in 1996 when it was translated, and the first edition got published in 1995 (18 p. 7). WHO has been working on ICD-11 for years and the Beta phase has been running since 2012 (19). This version was expected to be released already in 2012 (18 p. 81) but currently the estimate is June 2018 (19). A lot has happened in information technology since 1990 and major changes are expected also for the ICD and its usage but apparently it is difficult to make major changes to a classification that is used globally.

Another example of international efforts of standardizing healthcare is the non-profit organization Health Level Seven International (HL7) that was founded in 1987 to standardize the exchange,

integration, sharing, and retrieval of electronic health information (20). HL7 Finland was founded in 1995 as the 5th International Affiliate of HL7 and it is an open association that promotes interoperability standards, produces national interoperability specifications and provides guides and trainings for localizing and implementing solutions following the standards (21). HL7 standards are used in Finland for example in Kanta Services like Patient Data Repository and Electronic Prescription Centre (21, 22).

The previous examples of international standardization efforts have been adopted also in Finland and are in active use and development nationally. An example of an international effort that has not yet landed in Finland is the SNOMED CT. SNOMED CT is a clinical terminology created, developed and maintained by a range of healthcare specialists in the International Health Terminology Standards Development Organisation (IHTSDO) (23). The terminology supports clinical decision-making and analytics in software programs (23). This terminology is in use already in over 30 countries and in all the other Nordic countries except Finland (24).

The National Institute for Health and Welfare (THL) is preparing currently for the possible adoption of the SNOMED CT terminology in Finland and the Ministry of Social Affairs and Health is expected to decide about it in summer 2018 (24). This is not a new development in Finland because already in 2004 Matti Ojala, the director of Classification Centre of Stakes (predecessor of THL) considered SNOMED CT in an article of Terminfo (25). He anticipated at the time that if SNOMED CT was to be taken into use in Finland it would take considerable resources and still take years (25). Later, in year 2010, Åsa Holmér reported in the same journal that it was a massive project of three years to translate SNOMED CT into Swedish with the help of experts of several special fields (26). At that time there were 315 000 terms and expressions in the terminology (26). The number of concepts in SNOMED CT continues to grow, and the January 2018 release contained 341,105 active concepts (27). Concepts, terms and expressions are covered in more detail in the literature study of this report.

2.3 Languages in Software Industry

Software industry has been global for decades, if not always. English is the common language of computing and programming (28, 29). This, in part, could be the result of having English words, or at least words derived from English, as instructions or commands in the most popular programming

languages. Some software companies and open-source organizations have programming guidelines commonly known as coding conventions stating that all comments in the source code must be in English (30, 31). In some others, the second language i.e. the natural language in source code in addition to the programming language i.e. formal language is completely omitted either because it is taken for granted or it does not make any difference in that environment. Some guidelines take the specification of the natural language to an extreme, like the Drupal coding standard: “One overall note: comments and names should use US English spelling (e.g., "color" not "colour"). (32)” The specification of the natural language to use in source code is commonly justified by better maintainability and transferability especially in international contexts (31) but is very much valid also for companies and organizations wishing to become international in the future. Software development and maintenance is easy to move from country to country when comparing to other industries.

There has been some research recently done about the natural language in source code. Pawelka and Juergens conducted a study in 2015 where they studied the comments and identifiers of both open-source and industrial software (29). The results of the study showed that a significant amount of the studied industry software contained comments and identifiers in more than one natural language and none of the open-source software in the study had this problem (29). The study was conducted in Germany and German language is historically one of the main languages in Europe (33) which could well explain the usage of German in source code. The study anyhow does not tell whether the industrial systems that were studied had any coding standards or even if the companies studied were international or local. It is likely that if such a study were conducted in Finland, the results would be different. In Finland many of the international software companies have a rather high percentage of foreign employees. One example is a game industry software company Unity that recently stated having 38% foreign employees in the Helsinki office and them representing 24 different nationalities (34). Even the originally Finnish software companies tend to use English in documentation and source code if there are any wishes to grow outside of Finland for example through company acquisition in either direction.

The language used when developing software is specialized and constantly evolving and changing due to the very nature of the industry creating something new all the time. In addition, the methods and tools that are used in developing software are also developing and that is also changing the language. If a Finnish software industry company tries to use Finnish as a working language, it requires inventing new Finnish words constantly. And Finnish software development people do

invent words to discuss the items they are working on with their coworkers. The words however do not often get standardized even within the company and each person gets to invent their own words for example to architectural elements or several types and sizes of objects created in object-oriented programming languages. There have been discussions about classes (“luokka”) and methods (“metodi”) and applications (“sovellus“) but also about “palikka” (“block”), “härpätin” (“contraption”) and “kikkare” (“dumpling”). Especially when using agile development methods and hence minimizing documentation, these words usually do not end up documented at all but are normal spoken language in a Finnish software company. Special language issues like these are covered in the literature study section of this report.

2.4 Languages in Healthcare Sector

This chapter introduces the general languages present in healthcare sector in Finland. The professionals of healthcare sector use a highly specialized language and that is studied and introduced in the literature study section of this report along with the previously mentioned specialized language of software industry.

The customers in healthcare sector in Finland comprise mostly of the population of Finland and for that reason the languages present in the population are presented here. For the purposes of this thesis, it is more relevant to focus in native languages of people rather than immigration, emigration or nationalities, even though nationalities in some cases do have an impact on healthcare service fees and work permits for example. Figure 3 shows how the number of foreign-language speakers living in Finland has increased over the last decades from an insignificant number to such that it is noticeable. In the end of year 2017, there were 373 500 foreign-language speakers living in Finland. That is 6,8 per cent of the population. Foreign-language speakers here refer to people whose native language is not Finnish, Swedish or Sami. The amount increased from previous year by over 19 000 people, that is a 5,5 per cent increase. (35.)

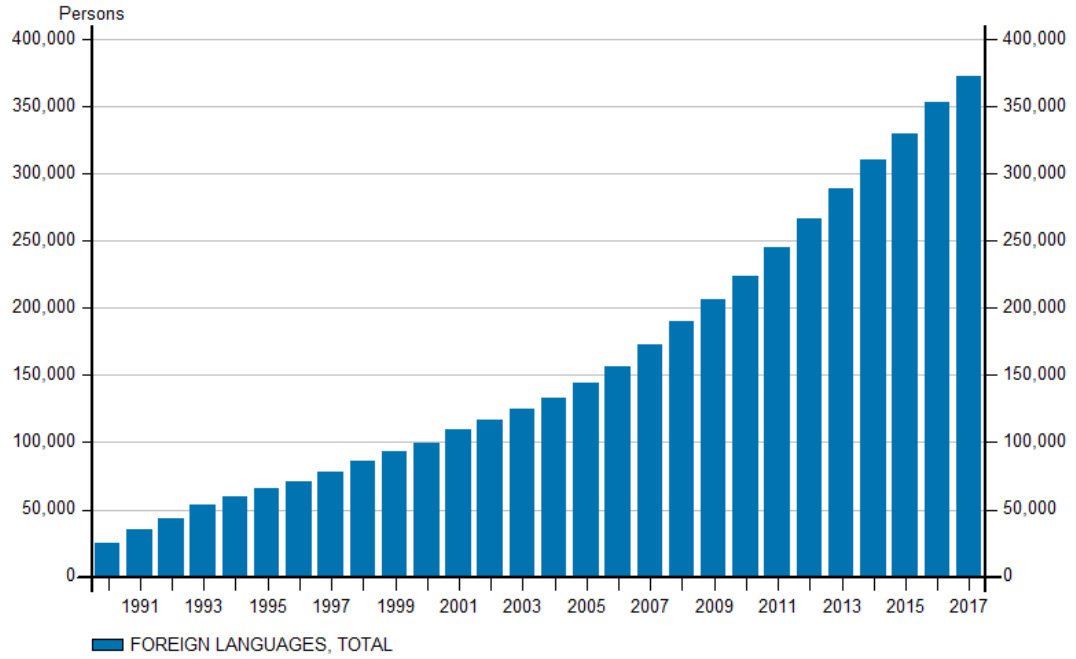


FIGURE 3. Population in Finland by language 1990-2017, total foreign-language speakers (36)

The increased number of foreign-language speakers in Finland has been noticed also in day-to-day work in healthcare sector. This can be seen for example in the number of thesis topics in this decade focusing on ways to cope with the unfamiliar cultures and languages among clients. Some examples of those theses are: developing a website describing the Muslim culture and how to handle it in healthcare settings (37), designing a multicultural learning package for public health nurse students (38) and creating a pocketsize Finnish healthcare vocabulary for seven languages (39).

Increased tourism in Finland and the so-called healthcare tourism also increase the number of foreign languages among healthcare clients. Finland received 8.3 million foreign visitors in 2017 with an increase of 13% to previous year (40). Not many of the tourists need healthcare services in Finland but accidents do happen also on holidays and the increased tourism is bound to show also in healthcare services. This group of people does not anyhow seem to have inspired many studies or noticeable research yet.

So far, most of healthcare tourists come to Finland from Sweden, so they do not introduce new languages but just increase the number of Swedish speaking patients mostly in private healthcare (41). This rather new and increasing trend is partially taking place due to the free choice of

healthcare providers for patients that has been in force in Sweden for some years already. The county councils restrict the selection of service providers by agreements, but patients can freely reserve appointments with those specialists without referrals. (42 p. 17) Patients in Sweden are therefore accustomed to choosing their healthcare professionals. In the north part of Sweden, Finland may be easier to access for geographic reasons than a specialist further away in south Sweden (41). Another reason for increased healthcare tourism is the Directive 2011/24/EU of the European Parliament and of the Council on the application of patients' rights in cross-border healthcare (43). The transposition time for this directive was until 25 October 2013 (43). Finland complied with the directive with Act on Cross-Border Health Care on 30 December 2013 (44). The directive was adopted at varying times by the European Union countries and patients were unaware of this possibility of cross-border healthcare in 2014 (45 pp. 7-9). There will be a new report on the progress of patient mobility within the European Union later this year (45 p. 3). The directive is likely to increase the internationalization of healthcare in Finland in the long run.

The professional side of healthcare in Finland has stayed domestic for a long time due to the difficult national language combined with strict language proficiency requirements. The Finnish Health Care Act, Section 6 deals with the provision of health care services in different languages (46). It is stated that the language of service is determined mainly according to the language of the local municipality, but bilingual municipalities must provide service in both Finnish and Swedish so that the clients and patients can choose their preferred language (46). The Finnish Government determines the languages for municipalities according to minority percentages for ten-year periods at a time by a Decree (47). Currently there are 311 municipalities in Finland of which 16 are Swedish-speaking, 33 bilingual and the rest are Finnish-speaking (48). The employer of a healthcare professional is responsible of ensuring adequate language proficiency of the employee for their tasks (49 Section 18a) and with a recent amendment the national authorities reserve the right to check the adequate proficiency level of Finnish or Swedish as a prerequisite before granting professional rights to citizens of other member states of the European Union (49 Section 8b). This means that most of the healthcare professionals must have professional level Finnish skills but combined with increasing number of foreign-language speakers in Finland, also the need and space for foreign-language professionals increases.

In addition to the language requirements of professionals there is also the issue of recognizing professional qualifications in the highly regulated professions of healthcare sector. The Health Care Professionals Act was amended in 2007 to provide for the Directive of the European Parliament

and of the Council on the Recognition of Professional Qualifications 2005/36/EC. (50). It describes how professional qualifications obtained in the European Union and elsewhere are recognized in Finland. The citizens of European Union are granted rather free access to the labor market in Finland but each healthcare professional in Finland must be registered nationally. Professionals coming from other countries have a more complicated application process.

Despite of the previously mentioned requirements for healthcare professionals entering the Finnish labor market, the number of foreign-language professionals has been increasing in the recent years. In year 2016 there were 1818 physicians of working age licensed in Finland with a mother tongue other than Finnish or Swedish (51). A year before there were 1436 of them (52). The total increase of physicians in Finland in that time was 567 and 382 of them were non-native by mother tongue (51, 52). It is to be noted here though that not all the licensed physicians are working in Finland.

3 LITERATURE STUDY

Both the medical field as well as the software development field are very information-centric and because of that specialized languages have developed in those fields to communicate that knowledge. This chapter first introduces varying terminology related words and how they relate to terminology work. Then, a slightly deeper dive is done to principles and methods of terminology work through the international standard and a more practical approach to it with literature of experienced terminologists. Theoretical science of terminology is explained as little as possible in this report because it is out of scope of this project.

Scaled Agile Framework (SAFe®) is introduced in this chapter because it is used in the case company and the working methods of a company affect the choice of tools and methods to use. This introduction is done based on material from Scaled Agile, Inc. that owns the trademark and methodology (53). This source serves the purposes of this thesis project adequately.

This literature study used a qualitative and highly selective method for the case study purposes of this thesis project. Searches especially for the practical terminology work were made in Google Scholar, library databases and other online publication databases like IEEE Xplore, ScienceDirect and Ebsco. Also, the Finnish online magazine Terminfo published by the Finnish Terminology Centre TSK is used for the practical approaches of terminology work to get some Finnish perspective to the topic.

3.1 Terminology Related Terminology

3.1.1 Special Languages

General language is produced by selecting words from everyday language and combining them logically. Lexicographic dictionaries that are used in general language document the global range of words that are used throughout the community and language. Terminology management and the produced terminological resources focus on terms that are used in special languages. Terms in these special languages represent mostly nouns and verbs. In some subject fields, also adjectives, adverbs and specific phrases can be terms as well. The terms are combined with

connective words of general language to create complete sentences in special language. Identifying and selecting terms is the first step and a continuous effort in developing terminological resources. (1 pp. 13-14.)

Lexical subsystems that are intended for unambiguous and clear communication in a subject field are called special languages (1 p. 330). Varantola cites several authors in her article on special language and general language for the definition of a special language (54). Hoffmann is quoted to scope it as "…a complete set of linguistic phenomena occurring within a definite sphere of communication and limited by specific subjects, intentions and conditions". Sager et al. is quoted in the article as well with their definition as "Special languages are semi-autonomous, complex semiotic systems based on and derived from general language; their use presupposes special education and is restricted to communication among specialists in the same or closely related fields."

3.1.2 Collections of Words

According to the International Organization for Standardization (ISO), the set of designations belonging to one special language constitutes the terminology of a specific subject field (55 p. v). The Handbook of Terminology Management considers "a terminology to be a structured set of concepts and their representation in a specific subject field" (1 p. 325). The handbook explains some of the near synonyms of a terminology as well to make the definition clearer. Here is a very brief description of those (1 pp. 325-326):

- A vocabulary is a list of words used for example in a language, in a book or in branch of science. A terminology usually refers to engineering, law, art or other serious disciplines and vocabulary can include both special language and general language words.
- A glossary is an alphabetical list of words related to a specific topic. Explanations of words may be included. A glossary resides in backmatter attached to books and other publications. Glossaries are sometimes considered less scientific than terminologies and vocabularies.
- A dictionary is usually a book that lists and defines words in a language. If the word collection is not large enough to be a book, it cannot be called a dictionary either.

- A lexicon in English is a high-register word for a dictionary or it can be the vocabulary of a person, language, or domain. Internationally lexicon is not commonly used for these purposes because in other languages it contains broader encyclopedic knowledge.
- A nomenclature is an authoritative system of terms following strict, systematic naming rules and practices in a specified field.

3.1.3 Ontology vs. Terminology

A term that can cause confusion in the context of this study and therefore needs to be explained is ontology. Traditionally, the word ontology means the philosophical study of being and what entities exist in the universe and how they can be categorized. It is a branch of metaphysics and studies the essence of things. In recent decades, in information technology, the word ontology has been adopted to mean the working model of entities and interactions in a domain of knowledge like for example electronic commerce. In artificial intelligence (AI), ontology specifies the conceptualizations that enable software and people to share knowledge. More specifically, an ontology is a set of specified concepts: things, events and relations. This ontology is used as a vocabulary in information exchange. (56.)

Ontologies are common in the internet ranging from large taxonomies categorizing websites with their contents to categorizations of products and their features in online stores. Many organizations are working to get standardized ontologies for specific domains that can be used to share and annotate information. The SNOMED CT mentioned in the Background chapter of this thesis is one example of such a standardized, structured vocabulary. An ontology includes machine-interpretable definitions of concepts in the domain and relations among them. (57.)

The modern understanding of the term ontology is closely related to terminology and the work done in either should be reflected in the other. The most notable difference is that a terminology and a terminological resource is not commonly machine readable i.e. formal language, and that is the distinct feature of an ontology. According to Sauberer (58), the 2010 conference of Terminology and Knowledge Engineering (TKE) had a pre-conference workshop “Establishing and using ontologies as a basis for terminological and knowledge engineering resources”. The description of the workshop gives an understanding of the relation of the two disciplines:

> *“For knowledge engineers with a background in artificial intelligence, knowledge domain concepts and the relations established between them can only be represented by using a*

formal language. For terminologists with a background in linguistics, concepts and their relationships are represented in discourse by means of natural language terms, these occupying a place in concept systems relevant to the ontology-building stage. With a view to knowledge communication, representation and sharing purposes, this workshop will take into account the linguistic and conceptual dimensions which take place in the different stages of ontology development with special emphasis on the contribution of terminology."

3.1.4 Purposes of Terminology

Terminology management and terminology work can generally be categorized either descriptive or prescriptive according to their purpose. Descriptive terminology management and work helps writers and translators make wording choices, but this type of a terminological resource does not dictate their choices. This approach is usually practiced by translators, technical writers, and social scientists who document terminology for translation and writing, but generally do not themselves decide on the selection of words in the subject field. Standardizers perform descriptive terminology work if they prepare for standardizing terminology in a subject field by collecting existing terms and their multiple definitions. (1 p. 329.)

Standardizers, government regulators, nomenclature specialists and language planners generally work with prescriptive terminology management. Language planners create terminology and words for subject fields to enable for example technology transfer between linguistic communities. A standardized terminology in a subject field ensures that standardizers and stakeholders understand each other when they negotiate and agree on technical specifications and that in turn enables implementing for example standards or legal regulations appropriately. Translators and other non-experts in terminology regulation can perform prescriptive terminology work when they specify in-house usage of terms. (1 p. 329.)

3.2 Introduction to Theory of Terminology

Terminology work has been standardized by the International Organization for Standardization (ISO). According to the standard, terminology is multidisciplinary and is related to several disciplines like logic, linguistics and information science. Terminology is a study of concepts and their representations in special languages and general languages and it uses many theoretical approaches to describe, order and transfer knowledge. The main activities of terminology work and terminology management are:

- Identifying concepts and their relations
- Analyzing and modelling concept systems
- Establishing representations of concept systems with concept diagrams
- Defining concepts
- Attributing designations to concepts in one or more languages
- Recording and presenting terminological data

Objects, concepts, designations and definitions are fundamental to terminology. Objects are perceived or conceived and abstracted into concepts which, in special languages, are represented by designations and described in definitions. A set of designations belonging to one special language constitutes the terminology of a specific subject field. (55 p. v.)

3.2.1 Conceptualization

Objects in this context can be anything perceived or conceived. They can be material or immaterial, concrete or abstract or even imagined. Examples of objects are a unicorn, a diamond, gravity and conversion ratio. Each individual object does not get named anyhow, every diamond does not have a differentiated name. After observation, an abstraction process called conceptualization occurs and the observed object gets categorized into a mental construct or a unit of knowledge called concept. A concept then can be represented in various forms of communication through a designation. In natural language the most common forms of designations are terms, appellations and definitions. In other language types designations can be icons, diagrams, formulae, sign language or body language. (55 pp. 2-3.)

The conceptualization process involves observing the properties of an individual object in a context of the subject field and then abstracting them as common characteristics that apply to a group of objects in that subject field. Several similar objects should be studied to obtain a comprehensive list of properties before the abstraction to characteristics. Characteristics can be grouped into types like composition, shape, function, use, movement, color and location. Characteristics should be also closely related to the specialized subject field knowledge and identification of characteristics often requires research if the terminologist is not familiar with the special field. A specific combination of characteristics forms a concept and characteristics should be used in the definition creation of a concept. Some characteristics are delimiting i.e. essential in distinguishing the concept

from another concept. Characteristics are critical for building concept systems and relations of concepts depend on their similar and distinct characteristics. (55 pp. 4-7.)

3.2.2 Concept Systems and Diagrams

Concepts in a concept system can have several types of relations and concept systems are often visualized by drawing concept diagrams. Two high-level types of concept relations are hierarchical and associative relations. Hierarchical relations can be either generic or partitive. Associative relations are thematic, for example cause and effect or action and actor. In a concept system of generic hierarchical relations each level generalizes their subordinate level and specifies their superordinate level. A specific concept has all the characteristics of the generic concept and at least one additional delimiting characteristic. In partitive hierarchical relations, the superordinate concept is a whole made of the parts represented by subordinate concepts. These are called the comprehensive concept and the partitive concept. In both the generic and partitive hierarchical relations, items on the same level are coordinate concepts. Coordinate concepts have distinct delimiting characteristics and with partitive coordinate concepts they can be also distinctly different. (55 pp. 8-9, 13, 17.)

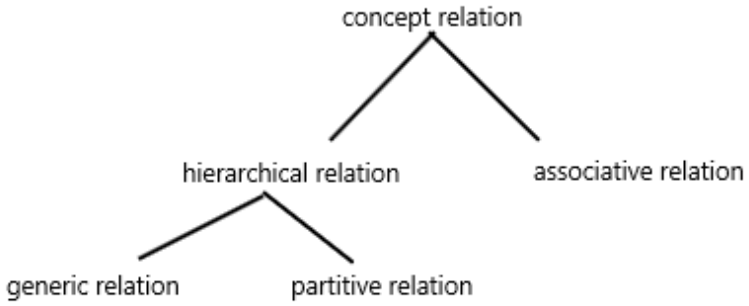


FIGURE 4. Relations of Concept Relations in a Tree Diagram

Figure 4 illustrates the generic relations of concept relations in a tree diagram like the generic relations often are depicted. Another common way to illustrate generic relations is an indented list with numbered nodes and levels like shown in Figure 5.

1 concept relation

1.1 hierarchical relation

1.1.1 generic relation

1.1.2 partitive relation

1.2 associative relation

FIGURE 5. Relations of Concept Relations in an Indented List

Partitive relations are indicated with a rake diagram or an indented list with dashed node numbers. Examples of these are shown in Figure 6. Associative relations are depicted in diagrams and lists with bidirectional arrows ←→ (55 pp. 18-19).

1- unicorn

1-1- tail

1-2- head

1-2-1- mane

1-2-2- muzzle

1-2-3- horn

1-2-4- neck

1-3- hoof

1-4- body

FIGURE 6. Relations of a unicorn in a rake diagram and an indented list

Concepts can be subdivided using different criteria even in the same concept system (55 p. 8). The different criteria are based on types of characteristics, for example composition, function or usage (55 p. 10). Division by one criteria is called a dimension and if the same concepts are divided by several criteria, the system is considered multidimensional (55 p. 8). The criteria to use depends

on the subject field and purpose of the terminology work (55 p. 8). Then, the system can use either one of the mentioned relation types in which case the system is typed according to it; or several of them in which case the system is called a mixed concept system (55 p. 19).

3.2.3 Designations

There are three types of designations categorized in the terminology standard: terms, appellations and symbols (55 p. vi). A designation is a representation of a concept or its definition (55 p. 22). A definition of a concept is valid if it can replace a designation in a text without loss of or change in meaning (55 p. 25). A term is a designation consisting of one or more words representing a general concept whereas an appellation designates an individual concept (55 p. 3). A general concept represents several similar objects and an individual concept represents one single object (55 p. 3). A symbol can designate both individual or general concepts and they are understood independently of any given language, so they work well in international communication (55 p. 41). Table 1 presents some examples of the mentioned designations.

TABLE 1. Example Designations

Designation Example	Designation Type	Concept Type
recyclable	Term	General Concept
-	Appellation	General Concept
	Symbol	General Concept
-	Term	Individual Concept
Oulu University of Applied Sciences	Appellation	Individual Concept
	Symbol	Individual Concept

3.3 Introduction to Practice of Terminology

3.3.1 Starting Terminology Work

Seija Suonuuti, who built and lead Nokia's global terminology work for over 30 years (59) and teaches still at the university of Tampere (60), wrote an article in Terminfo in 2013 about starting terminology work in a company (61). In the article Suonuuti presents aspects that should to be considered before making the investment decision about starting terminology work in a company.

According to Suonuuti the benefits of terminology work depend on the organization and environment, so first it is important to study what kind of terminology work is needed in the organization. A thorough preliminary study is critical before starting terminology work to avoid big mistakes like investing in terminology software without knowing what the content will be or who will be doing the terminology work. The study should also consider various solutions. (61.)

The purpose of terminology work is always to guide the usage of terms and if the terminology is built just by listing and describing used terms without larger scale objectives, limiting or harmonizing, the benefits of the work will stay minimal. The terminology may look fancy externally, but it does not solve the problems, give answers to questions or guide the usage of terms. This kind of a terminology may create false feelings of satisfaction and security to the occasional user when all the terms they search for are in the terminology. A more advanced user might notice that the terminology gives several conflicting answers to questions. This kind of a terminology does not change anything or benefit the texts it is used for and the same result could be achieved with using search machines. Terminology work can influence for example quality, coherence of language, localization and costs. (61.)

> *When the benefits of terminology work are considered, quality is often the first thing that comes into mind, and undoubtedly terminology work affects the quality of operation and language. But if there have been no distinctive problems in quality, the improvement of quality is not an adequate argument for decision-making. Quality is a working alternative when it has already been noticed that language should be developed. Wrong or bad terms can hinder the understanding of texts, give unprofessional image of the product or service, or hinder information retrieval. Quality aspects are often linked with the uniformity of language and how well it can be localized. (62.)*

Consistency of the used terms affect text often considerably more than the quality of language as such. It is easier for a reader to notice inconsistent use of terms than poor terminology. Inconsistent terminology makes understanding and reading texts difficult and that deteriorates the product image. Poor terminology can consist of using ambiguous or wrong terms or using loan words instead of terms in the text's language. If this is done consistently, an occasional reader might not notice it, but an expert sees a poorly written text. Consistent terminology can be of critical importance in texts where quick and precise understanding is vital for example for safety reasons. (61.)

> *Coherent terminology in which one term, or at least a limited amount of synonyms, is used for one concept is very important if some text is changed or corrected. A coherent text is much easier to correct when some term is changed. A coherent text is also easier to localize when the translator is not forced to guess whether it is a question of a synonym or a totally different concept. For the user, coherent terminology creates quality and gives an impression of a good-quality product. (62.)*

Consistent terminology facilitates for example information retrieval both in search engines as well as in the product texts. This enables also the writer to edit the text for different target groups. It is common that the same concept is known by different terms in different areas, and it is possible that separate versions of the product are required for the areas. Consistent term usage in texts and the knowledge of area specific terms facilitate this type of versioning. (61.)

> *Different companies have reported savings after they have started terminology work. The savings have been approximately 20%. The information on savings is usually based on interviews and exact calculations have not been given or they are not publicly available. The savings depend on the starting point and how big problems the lack of terminology work has caused. The availability of public terminologies used in the field can affect savings some terms have already been clarified and probably in use.*
>
> *It depends on the nature of terminology work what kind of savings can be achieved with it. The savings of monolingual terminology work remain often smaller than the savings of multilingual work. The savings do not depend only on the amount of languages. Other influencing factors are e.g. the number of users of a terminology, number of terms to be handled and how established the terms are. (62.)*

Estimations of savings provided by terminology work can be used as an argument for starting terminology work. Measurable results can only be calculated after around 1 to 3 years after starting but calculations can be used to make reliable and justified estimations beforehand. For these calculations, the time used for finding terms, discussions and checks for the right term and correcting wrong terms can be estimated. According to Suonuuti, there are commonly disputes

about terms but if the results are undocumented, the same problems are solved several times. Suonuuti suggests that a questionnaire about time usage could be done to the writing staff to support this estimation. If, for example the questionnaire results show that 5% time saving per person can be achieved with terminology work and the writing staff is 50 people with monthly costs of 3 000€ per person, the savings with terminology work become 7 500€ per month. Further, if it is assumed that the personnel cost of terminology work is also 3 000€ per man-month, a person could be employed for terminology work and there would be an additional 4 500€ per month for amortization of start-up costs and investments of terminology work without additional costs. The example would have 54 000€ for other costs with one-year amortization period. (61.)

Return on Investment (ROI) for the terminology work can also be calculated based on the previous example calculations. Assume for the calculation that the investment cost of terminology software is 20 000€ and the critical level of content in the terminology is reached after one year of work. The total costs would be 20 000€ + 12 * 3 000€ = 56 000€. ROI is calculated for the traditional 12-month period with the following equation:

$$ROI = \frac{(investment\ profit)}{investment\ cost} \times 100$$

Investment profit = total savings – total costs = (12 * 7 500€) – 56 000€ = 34 000€
ROI = (34 000€ / 56 000€) * 100 = 60,7%
ROI is positive already with the 12-month period and above 20% that is considered the limit of good ROI. The calculation can be used also for estimating the investment amount available for the terminology software. (61.)

The previous example calculation does not consider the possible effects of multiple languages. In practice, the costs per language decrease when the number of languages increases. This is the case when terminology work is done in one source language and only equivalent terms are searched for in other languages. According to Suonuuti's experience, the time needed for one term and definition in the source language varies between 20 minutes and 180 minutes depending on the experience of the terminologist, the complexity of the terminology, how established the terms are and the available sources. If the source language section of the terminology work is done properly, the time needed for searching equivalents in other languages is considerably shorter. The factors affecting the time needed for equivalent searching are the language itself, how established

the terminology is in the topic area in the language, as well as the experience and proficiency of the translator. Suonuuti herself reports using usually two to 20 minutes for an equivalent. (61.)

> *Besides time management, terminology work can also affect other investments and costs. When there is one source for terminology, the need for other sources will decrease. The time used for negotiations and proofreading may become shorter and the need to rewrite or re-translate decreases. (62.)*

Suonuuti states that after the benefits of terminology work are clear in the organization, it is time to study the requirements for terminology work. There are several types of terminology software available for different purposes and it is best to select the software and make the investment only after compliancy to the organization's requirements is evaluated. One of the most critical issues when planning terminology work is the purpose of the resulting terminology. The varying contexts where the terminology can be used define quite different requirements for the terminology and its usability. If the main part of usage is mono- or bilingual and there are no localization requirements, the software terminology requirements are quite different compared to terminology work with localization and use of many languages. According to Suonuuti it is advisable to use the terminology as widely as possible to reach good cost efficiency. Considering also the quality aspects, it is better if the terminology usage is not limited to only one part of the organization. The wider the usage, the better effects to the quality and maximum savings. (61.)

> *The terminologies and terminology work must also meet the users' needs or the terminologies will be curiosities and no benefits will arise. The terminologies must be linked to everyday work and function in the same environment as the tools used for writing. The users must be able to find the information they need and to link it as part of their work. (62.)*

Suonuuti summarizes four issues in the article that define the basis for the start and planning of terminology work in her opinion (61):

1. What problems and challenges the terminology work should answer?
2. What is the ratio between costs and benefits?
3. What are the user requirements?
4. What can be done with the available resources?

3.3.2 Business Process View to Terminology Work

Seija Suonuuti examines terminology work as a business process of a multilingual company in an article of Terminfo in 1998 (63). According to Suonuuti's article the processes of a company are generally categorized as core or support processes. Core processes are the ones that realize the business idea of the company and produce the profits. A disturbance in a core process has immediate impacts in the company performance but the absence or stalling of support processes may influence only after a brief period. Terminology work is clearly a support process when compared to other processes of a company, unless the company provides terminology work services to others. (63.)

Terminology work is a support process and its significance to the company operation is difficult to measure. The influence of terminology work is often only indirectly connected to the economic operation of the company and the significance of terminology work can be seen years afterwards. Certain critical level of terminology work and usage of the resulting terminology must be reached before it is feasible to measure the influence. The results of the terminology work must cover a major portion of the used terminology in the field and the results must be used in most of the produced documentation before the influence can be measured. Terminology work must be comprehensive enough and the resulting terminology be used widely enough in the supported process for it to make a difference. It is not enough if terminology work is done and utilized only in a small group of enthusiasts. This is anyhow the common way how terminology work is started in companies, states Suonuuti. (63.)

Defining the relation of terminology work to other processes can be challenging according to Suonuuti. The processes that are supported by terminology work must be known and the way they are supported by terminology work must be known. Also, it must be known what the results of terminology work are used for in the supported process. Obviously, multilingual terminology work supports translation process, but there is need for monolingual terminology work as well. Suonuuti states that efficient and high-quality translation process requires high-quality source text with good terminology. This eventually means that terminology work must support all processes that produce documentation and there are very few processes that do not at least describe and verify the results of the process through documentation. (63.)

The support provided by terminology work to other processes can be of diverse types and levels and be situated in varying phases of the supported process. For some processes, terminology work results are needed mainly in the beginning phases and wrong term selection in this phase can result in bigger economic effects than the same mistake in later phases. Correcting and harmonizing terminology afterwards can be a lengthy and challenging task, so it is critical to schedule terminology work correctly in relation to the supported processes. (63.)

Suonuuti presents an example of a design process and the writing and translation process of related customer guides. In the example it is assumed that the processes use the same consistent terminology because they describe the same product. Some part of the terminology is bound to be the same in these processes but according to Suonuuti's experience, a surprisingly large part of the terminology may be developed only during the guide writing when the technical solutions are illustrated and popularized through examples, customer specific applications and user interface descriptions. In the example, terminology work in the design phase is focused in monolingual definition related terminology and in the guide writing phase it is focused in multilingual customer friendly terminology. Combined, this terminology can form an irreplaceable knowledge base and a valuable entity describing the product and its use. (63.)

When planning the process for terminology work and when running the process, the varying needs of the supported processes must be considered. The design process of the product is easier to perform if the needed terminology is well defined and the relations between concepts are clear. In the example, the terminology work should happen simultaneously with the design process and be ready before the guide writing and translation process. The terminology process occurring during the design process is significant in the language used for designing and possibly the languages used for example in the applicable standards. Often terminology work can be monolingual during the design process and emphasis can be put on definitions of terms. For the guide writing and translation process it is ideal if most of the terminology is already clear before starting. This facilitates a consistent terminology both in the source text as well as the translations and shortens the time needed for the translations significantly. In these example processes the multilinguistic aspect as well as the availability of correct and usable terms in the terminology are highlighted. (63.)

An ideal terminology process supports other processes at the right time, for example providing accurate term definitions for the design process and all the needed language variants before the

guide writing and translation process. The terminology process should also be profitable, the benefits should exceed the investments. Suonuuti suggests that the terminology process should be phased, and the target user group and content be limited or at least extended only gradually. One option according to her is to divide the user group into main customers and other customers. Another option is to have several terminology processes, for example one for monolingual product related process and a multilingual translation related process that focuses in equivalent searching and harmonizing the multilingual concept system. Care is needed in planning either of these options to guarantee wide acceptance of the terminology process results and active usage of the terminology. (63.)

A well-planned terminology process can achieve quality and economic objectives set to it, for example saving time in documentation and translation, harmonizing the used language and decreasing errors in term selection and translation. A good terminology process offers solutions to all the customers, the results of the work are appreciated, and the terminology is considered common property in the sense that it can be evaluated and criticized for further development. Suonuuti suggests that time is used for developing the terminology process and evaluating its status. The risk is that even a professionally created terminology becomes outdated and is not eventually used at all if it does not reach its customers appropriately. Terminology work can be experienced also as a hindrance to the other processes and causing for example rewriting needs. A good terminology process requires continuous development, evaluation of working methods, receiving critique and communicating with customers. According to Suonuuti, the biggest challenge with terminology work both inside companies and outside, is reaching the customers and creating a working and productive interaction with them. If the process fails to do this, it can be thought of as a burden and a necessary evil. (63.)

3.3.3 Terminology Workflow

According to a qualitative interview study of professional terminologists in the legal and administrative organizations by Chiocchetti et al. (64) the theoretical workflows and steps of terminology work described in the literature give a simplified picture when compared to the reality of terminology work. The study revealed that in practice, the roles and process steps are not clearly defined or separate and one person often has several roles and the workflow tasks sometimes

overlap. The study found that the practical process for terminology elaboration has the following steps:

- needs analysis
- documentation
- term extraction
- term selection
- elaboration of terminological entries
- revision and quality assurance
- dissemination
- maintenance

In the needs analysis step the most critical aspects are the time frame and the terminological problem to solve in that time frame. Ad-hoc terminology work can be done if the time is very limited. In this working method, the terminology is compiled simultaneously with the other activity that requires the terminology work, for example translation or documentation. If there is more time for the terminology work, either proactive or systematic approach can be used. In these approaches, the terminology work is done first, and the results are utilized afterwards. Proactive terminology work is done if results are needed medium term and the subject or term group is limited enough. For long term needs and bigger entities, for example a new language or domain in the terminological resource, the systematic approach is suitable. According to the study by Chiocchetti et al. most of practical terminology work is done ad-hoc even when trying to be proactive and very few professionals have time for systematic terminology work. (64.)

The documentation step is the data gathering of the relevant terminological information generally from documentation of varying kinds. It can include also cooperation with domain experts. The information sources should be evaluated for relevance against the purpose, content and target users of the terminological resource. Next is the term extraction step, where the candidate terms are extracted from the documentation collected in the previous step. This step can be done manually, semi-automatically or automatically by utilizing term extraction tools but according to the study, the professionals preferred manual term extraction due to poor usability of the tools at the time of the study. In the following term selection step, the previously collected candidate terms are validated by terminologists or domain experts before they should be further elaborated and included in the terminological resource. (64.)

The elaboration of terminological entries -step includes adding the selected terms into a terminological resource with further details like domain attributions, definitions, contexts of use, equivalents in other languages, synonyms and variants in the same language, sources of definitions and contexts and any other additional information. The following revision and quality assurance -step is essential for all high quality terminological resources. This step includes three different validations: spellcheck, completeness check and correctness check and can be performed by people in various roles. Dissemination is the last and most important fixed position step in the terminology workflow according to Chiocchetti et al. In this step the results of the terminology work are delivered to the users of the terminological resource in one or several ways. (64.)

Maintenance, listed previously as the last step of the workflow, can be done any time and be triggered by several types of things like a spelling reform or a quality check. The frequency varies between organizations and the triggers used in the organization. The step includes also for example consolidation efforts and removal or merge of duplicate entries and handling of legacy entries. The study by Chiocchetti et al. revealed that maintenance of a terminological resource is very important, but it takes a lot of time and resources. (64.)

3.3.4 Roles in Terminology Workflow

Chiocchetti et al. studied also the roles related to terminology work in their interview study (64). Terminology work is usually done in teams of people with diverse linguistic, professional and technical competences. Only large terminology centers can separate all the roles distinctly. The terminology team should have people with terminology-related expertise, for example terminologists, translator-terminologists and quality controllers. These people are the main actors participating actively in all the steps of the terminology workflow because they are responsible of the terminological entries and concept systems. The study revealed that there is need for more competent terminologists in the organizations, but the report refers to a 2012 study by Chiocchetti and Ralli to explain that “the limited availability of qualified staff is due to the lack of university courses and specific trainings focused on terminology”. Members of a terminology team with management-related expertise coordinate the work and projects and cooperate with decisionmakers in the organization, as well as with customers and end users. Another significant role in a terminology team is that of an information technology (IT) expert. They can provide

technical support in each of the workflow steps, for example as database administrators, tool developers or IT specialists. (64.)

Domain experts bring subject knowledge to a terminology team in most steps of the terminology workflow and act as consultants and validators for terminologists. Their domain knowledge enables high-quality terminological entries. For example, domain experts bring up terminological needs, suggest source documentation and terms for elaboration, check concept systems, consult on content, create definitions and suggest or approve translation proposals. Domain experts mostly contribute to terminology work in their native language and they rarely have terminological background. They usually need guidance in the terminological aspects of the work, for example explanations on what a designation is or how a terminological definition is created. The study by Chiocchetti et al. showed that in most practical cases domain experts were involved in terminology work informally through terminologists' personal contacts, when ideally domain experts should be involved in the terminology workflow formally and regularly. (64.)

3.3.5 Translation-Oriented Terminology Work

The Handbook of Terminology Management describes in the Terminology Management for Technical Translation section the distinctive methods and features related to translation-oriented terminology management. According to the handbook, especially out-sourced translators compared to other terminology workers, must do ad-hoc terminology work because they often are not provided enough contextual material for the translation and do not work with the same topic and field for long. Usually translators do not have time for the preferred systematic approach for terminology management. The systematic approach is subject-field-driven as opposed to the ad-hoc terminology work that is text-driven. This results in a different sequence of activities in the workflow. A comparative view of the order of activities in systematic terminology work and ad-hoc terminology work is presented in Table 2. (1 pp. 147-150.)

TABLE 2. Workflow in Systematic Terminology work Compared to Ad-hoc Terminology Work (1 p. 150)

Systematic Terminology Management	Ad-hoc Terminology Management
1. Collect terms and concepts from global field	1. Identify terms occurring in isolated texts
2. Construct a concept system or systems	2. Create starter term entries
3. Craft well-structured definitions	3. Document available contexts
4. Create term entries	4. Research greater context, within time restrictions
5. Link entries to conceptual structure, reflecting the concept system(s)	5. If time and opportunity allow it, reconstruct the concept system based on available fragments

Translators should aim for a reliable and accurate target translation even with their limited knowledge of the domain or the discipline of the source text. To do that they must create some level of terminology documentation to ensure consistent and correct term usage in the translated text. (1 p. 149.)

Translators themselves can define what they record in minimum for this purpose for each entry in their own translation-oriented terminology documentation. Term-equation entries that are often used in glossaries i.e. Term A in L1 = Term B in L2 are adequate only if the translator knows and remembers the original context where the translation was valid. Usually this is not the case and classification according to domain and possibly subdomain categories is needed, so that the terminology documentation is reusable. Brief contextual references i.e. example phrases can be included for context definition and usage information i.e. discourse-related collocational information. (1 p. 151.)

For an out-sourced translator it can be impossible to formulate adequate definitions for concepts, but context recording should be feasible. According to the writers' (Wright and Wright) experience, also source, date and responsibility should be included in the entry record for future needs even if the future need is not known when creating the entry. Classification of the entries and selecting appropriate categories for them is essential for retrieval and manipulation in future. Wright and

Wright state that ideally all the previously mentioned data is recorded for each entry but more commonly some stay empty, often definitions and finer classification data at least in one side of the language pair. (1 p. 151.)

3.4 Introduction to Scaled Agile Framework

Agile software development is an umbrella term and somewhat standard currently in software industry. This introduction to a scaled agile approach assumes that the reader is familiar with the general ideologies of agile and lean development. If not, some educational reading could be agilemanifesto.org (65) and a book “Lean Software Development: An Agile Toolkit” by the Poppendiecks (66).

The case company uses Scaled Agile Framework (SAFe®), a trademarked software development methodology owned by Scaled Agile, Inc (67). SAFe has evolved in the software development industry as a proven methodology for developing complex systems and software in a Lean-Agile manner and it has three primary sources of knowledge: Agile development, systems thinking, and Lean product development. The main purpose of the methodology is synchronizing alignment, collaboration, and delivery of multiple Agile teams. Currently there are four types of SAFe configurations to fit different sizes of organizations and development environments (67):

- Essential SAFe
- Large Solution SAFe
- Portfolio SAFe
- Full SAFe

Figure 7 presents the structure, roles, methods and events of the Full SAFe configuration. It is the most complex of the configurations and includes all the other configurations, so the whole system can be introduced with it. The Large Solution SAFe and Portfolio SAFe are parallel in the sense that both include the Essential SAFe configuration but not each other.

The Essential SAFe configuration is the simplest and presents the basic implementation that exists in all the configurations. It describes the most critical elements of SAFe, the Team and Program Levels that form an organizational structure called the Agile Release Train (ART), where Agile teams, their key stakeholders, and other resources are dedicated to an important, ongoing solution

mission. The Large Solution SAFe configuration is for developing large, complex solutions for example in automotive or government sectors. These usually have multiple ARTs and Suppliers and building these solutions require additional roles, artifacts, activities, and coordination compared to the Essential SAFe configuration. If the solution is independent of others or requires only a few hundred people working on it, the Large Solution SAFe configuration is not what is needed. The Portfolio SAFe configuration organizes Agile development around the flow of value, through one or more value streams. It provides tools for breaking enterprise strategy down to portfolio level execution. (67.)

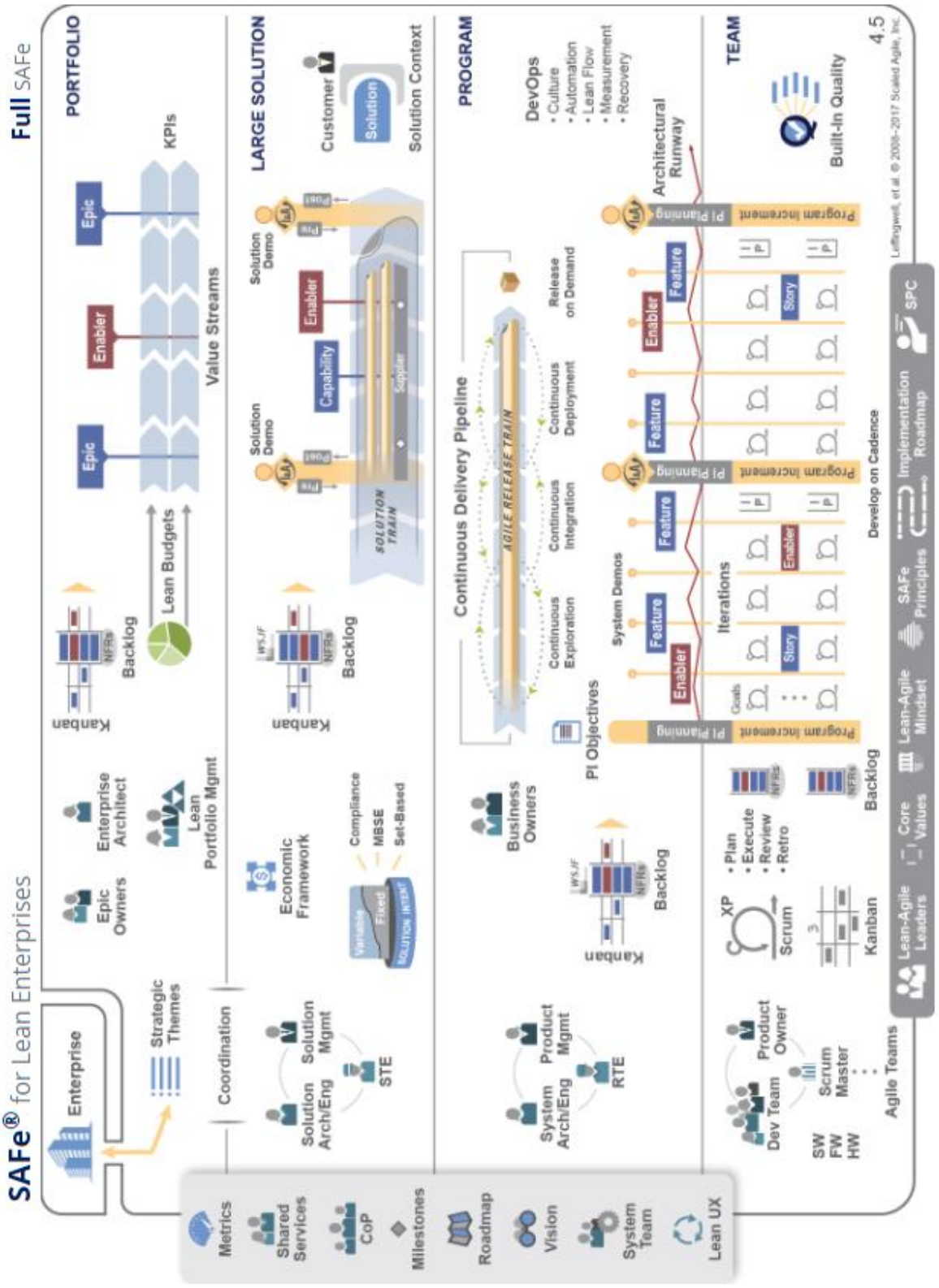


FIGURE 7. Full SAFe Configuration Big Picture (67)

The Spanning Palette and the Foundation are the grey areas around the Big Picture of SAFe in Figure 7 and they exist in all the configurations. The Foundation includes the principles and culture with what SAFe is run and the Spanning Palette includes the shared support functions and artifacts that are necessary but do not belong to any one ART (68 pp. 4-5). The Spanning palette includes for example metrics, milestones and roadmap but also Lean UX (User-Experience), System Team and CoP (Community of Practice) (68 p. 4). The Foundation holds for example the four Core Values of SAFe: alignment, built-in quality, transparency, and program execution (68 p. 6). It also holds the nine principles of Lean-Agile SAFe (68 pp. 11-12):

- Take an economic view
- Apply systems thinking
- Assume variability; preserve options
- Build incrementally with fast, integrated learning cycles
- Base milestones on objective evaluation of working systems
- Visualize and limit the amount of work in progress, reduce batch sizes, and manage queue lengths
- Apply cadence, synchronize with cross-domain planning
- Unlock the intrinsic motivation of knowledge workers
- Decentralize decision-making

The most notable events happening in all the SAFe configurations in Agile Release Train level are (68 pp. 15-16):

- System Demo - Every two weeks, the integrated working solution of all the Agile teams on the ART for that iteration is demoed to stakeholders. Stakeholders provide the necessary feedback to the ART to stay on course and make changes. This is the primary measure of the ART's progress.
- PI Planning - Each Program Increment (PI) is started with a two days planning event that ensures alignment for the Agile teams in the ART. PI Planning sets the common mission, vision, and purpose for the ART for the next ten weeks Program Increment. This is one of the major cornerstones of SAFe.
- Inspect and Adapt - The I&A event takes place once in each PI, usually around the planning. It assembles teams and stakeholders to assess the solution and plan on the improvements needed to increase the velocity, quality, and reliability of the next PI.

- Innovation and Planning Iteration - The IP iteration is the last two-week iteration at the end of each PI and it acts as an estimating buffer for meeting PI objectives. It also provides time for innovation, education and PI planning and I&A events.

Large Solution SAFe configuration has four additional notable events in the Solution Train level to coordinate several ARTs and possible suppliers (68 p. 20):

- Pre- and Post-PI Planning - These events are needed for preparing for, and following-up after, PI Planning of each ART and Supplier of a Solution Train.
- Solution Demo - This demo takes place once every PI and presents the integrated development efforts of all the ARTs and possible suppliers to customers and other stakeholders.
- Inspect & Adapt - This higher level I&A also takes place once every PI and assembles representatives from all the ARTs and suppliers to reflect and identify improvement backlog items.

Some of the most important roles and organization elements in SAFe are (68 pp. 14-15, 18, 20):

- Real Agile Teams and Trains have all the needed resources for each increment of the solution. They must be self-organizing and self-managing and have all the needed functions and roles to minimize overhead and maximize flow of value. Product Owners and Scrum Masters help the Development Teams meet their objectives.
- System Architect/Engineer is a person or small team that defines the overall architecture of the system including Nonfunctional Requirements (NFRs) and the major elements and subsystems. They also define interfaces and their usage.
- Product Management is the voice of the customer in the ART and they cooperate with Product Owners and Customers to mediate needs, define system features, and participate in validation. Product Management is responsible for the program backlog and they prioritize features and enablers for the next PI.
- Release Train Engineer (RTE) is the chief Scrum Master of an ART. They run for example the Program Kanban, Inspect & Adapt (I&A) workshop and PI Planning.
- Business Owners are a small group of stakeholders who are responsible for fitness for use, governance, and return on investment for the solution developed by an ART. They are key stakeholders and participate in certain ART events.

- Customer is the ultimate decider of value. Customers are essential for the development process and value stream and they have defined responsibilities in SAFe.
- Solution Architect/Engineer is a person or small team that defines a common technical and architectural vision for the Large Solution SAFe organization.
- Solution Management is responsible for the content in the Large Solution level. They cooperate with Customers to create the solution vision and roadmap and define requirements. They use a Solution Kanban to guide work.
- Solution Train Engineer (STE) is a servant leader and a coach in the Large Solution level who facilitates and guides the work of all ARTs and suppliers of the solution.
- Lean Portfolio Management (LPM) is responsible for the SAFe portfolio and are also financially accountable for it. The LPM decides on Strategy and investment funding, Agile program guidance and Lean governance.
- Epic Owners are responsible for coordinating portfolio epics in the Portfolio Kanban.
- Enterprise Architect is a person or group of people that work across value streams and programs to provide strategic technical direction to optimize portfolio results. They often act as epic owners for enabler epics.

3.5 Conclusions of Literature Study

There is much research done and studies to be found internationally regarding terminology work and the related translation work. Terminology is a multidisciplinary science or a field of study of its own, but it still struggles to be known generally. The field of study was briefly presented in this report through a highly selected set of sources from different viewpoints with the emphasis on practical work and starting terminology work in an organization. Some basic requirements for terminology work are extracted from these chapters in the Requirements section of this report to take them into account in the implementation and development of the solution.

Scaled Agile Framework (SAFe®) was introduced in this section to provide the reader with a context where the solution process and tools must fit. This working methods knowledge is needed and used in the requirements gathering and the implementation and development parts of this project. The operational framework affects both the way this project is run as well as how the solution should work. This methodology affects the further development and future aspects of the resulting solution of this thesis as well.

4 REQUIREMENTS

This chapter first describes the end-user requirements collection method and process. The main requirements are then listed and prioritized using two different requirement prioritization methods to get different views to the requirements.

Requirements from the Literature Study section of this report are extracted and listed at the end of this section. The requirements are for the solution itself as well as the process and management related to it. These requirements are collected separately from end-user requirements, because the end-users of the solution are not expected to be experts in terminology management and their requirements can conflict or ignore some very elementary terminology management related requirements. At least the most critical terminology management related requirements should be considered in the solution for it to be a working solution from the beginning and for some time.

4.1 End-user Requirements Collection Method

The main requirements in this study are collected from different type of end-users using non-formalized interviews as a method of qualitative research interview. This method is recommended in research methods studies for requirements gathering and criteria definition (69 pp. 140-141). Unlike Järvinen described the non-formalized interview (69 p. 141), in this study the interviews were carried out without audio or video recording to avoid social dissonance and enable openness and ensure confidentiality. These objectives followed the guidelines for qualitative interviews by Myers and Newman according to Järvinen (69 p. 141). Instead of audio or video recordings of interviews, the interviewer took notes of only the key results that were sent to the interviewees for acceptance after the interview session.

Most of the interviews in this study were carried out using a theme-based interview-guide like Järvinen described the non-formalized interview (69 p. 141). The key topics to discuss in the interview were listed by the interviewer before the interview sessions. This helped in bringing some structure to the mostly free-form interaction and ensured that the objective of gathering specified information in the interview was met. The first interviews were less formalized when the initial list of main requirements was gathered. This followed the description of building case in the

constructive approach according to Järvinen (69 p. 140). Later in the process, the interviews grew more into the direction of formalized interviews and evaluation case of the constructive approach (69 p. 141). In this case, it meant that the interview-guide consisted of more detailed themes but still each interviewee could use their own possibly ambiguous criteria for the stated themes.

These approaches to interviews were taken because they were for the most part familiar to people in the case company. Similar sessions are held when developing software with agile methodologies. Using a familiar method for the interviewees enabled minimizing the social dissonance and improving the quality of disclosure (69 p. 141). Other ways to minimize social dissonance in the interviews and get the true needs discussed in this study were a positive and relaxed atmosphere, familiar interviewer and interviewer's genuine interest in the interviewees' problem.

4.2 End-user Requirements Collection Process

The initial list of end-user requirements for this thesis work was collected in 2017 using the previously described theme-based non-formal interview method. It was known before the interviews that there are problems in the case company communicating with an offshore team and in multisite teams within the same country. The people reporting these problems worked in varied roles in software product development, for example requirement management, customer service, software development and product management. In the beginning of this thesis project five people in different roles were interviewed in varied group settings and some also one at a time. The themes that were covered in the first interviews were:

- the problem
- solution ideas
- requirements for the solutions

The interviews generally ended in discussing and brainstorming about the possible solutions like vocabulary including simple word for word translations between Finnish and English and how it should be. Others ended with wondering what some specific professional Finnish words mean exactly and some sessions ended in laughing at the imagination and humor used in naming objects in English in source code. One of the main end results of these interviews was that people generally did not know where to look for the translations or explanations of words and the services or places

they had checked did not give appropriate answers. It was also clear that people were busy doing their core tasks and could not use much time with each unfamiliar word. People were for example creating their own word lists for their own use and storing them in their laptops because that is what they could do quickly to help their own work.

The first interviews suggested that some kind of a shared specialized vocabulary, dictionary or glossary would help the situation of many of the user-groups. It was also evident that people needed help for a couple different but related problems in communication. The next interviews then had more focused themes:

- purpose of the glossary
- usability requirements
- technical requirements

The later interviews produced short lists of necessary features and requirements that could be used later in the thesis work. There were not many detailed requirements given but it was very clear that time is of essence and the quicker the solution is available the better. The collected requirements are described in the next chapter. Two main purposes for the glossary became clear in the interviews:

- Help translating Finnish national regulatory requirements as well as customer requirements into English source code
- Help understand Finnish specialized words used both in software development and healthcare and social welfare

Some requirements are extracted from these purposes as well in the next chapter.

4.3 Collection of End-user Requirements

The following list of high-level requirements was collected from the interviews:

- Company internal
- Must include global, national and company terms with descriptions
- Must include terms used in source code and product user interfaces with descriptions
- Must include terms in Finnish and English
- Easy to use
- In a place where everyone can access it

- Easy to maintain and expand by several people
- Get the solution into use as soon as possible

Most of the requirements were refined further during the interviews to understand what exactly is meant by them. For example, “easy to use” means different things to people.

Company internal meant that the solution should not be delivered to customers for example. The glossary should be internal to health and social welfare department, so that source code terms and such items that are part of the intellectual property of the company could also be included without worrying about disclosure. On the other hand, it was requested that the solution should include also global and national terms and not only company internal ones. This was explained by the fact that people did not like the idea of having to search a word in various places or services. That would take too much time and not be user-friendly in this setting. People need the global and national terms in their daily work as well as the company specific ones. The “Company internal” requirement is refined to “Must be accessible only within health and social welfare department” but the terminology to be included did not get further refined.

Finnish and English languages were needed but Swedish was excluded as unnecessary at this point. It became clear also that not all words need a translation between Finnish and English, but some need it between Finnish and Finnish. An example of this kind of a word could be a Finnish jargon word in software development that is used only with one product and not understood in other teams. To ease communication between product teams these kinds of words should be defined and linked to a more generally known word for the same concept if possible. The interviews revealed also that the two purposes of the glossary solution mentioned in previous chapter - translations between Finnish and English and definitions of words, were equally important. It was mentioned that a translation is often difficult to make without a definition in the source language. Translation without definition is possible for general type of words, but the glossary is more critical for specialized words. The languages and word definition requirements specify some of the information that the users need for each word in the glossary solution. The solution must be such that it is not necessary to have both languages for all the words. With this requirement it is possible to harmonize the used terminology between multisite teams in the longer run. The requirement is thus refined to form “Must support two languages but enable also single language usage with definitions”.

Usability came up often in the interviews and accessibility as well as maintainability were mentioned as aspects of usability. Those are so significant issues that they became separate requirements already in the interviews. It was mentioned that simplicity should be one objective because people with varying levels of technical background must use the solution. It was mentioned that preferably the solution would not introduce any new software so that there would not be change resistance or threshold for taking it into use. Maintenance of the solution and expanding the glossary should not depend on one person but rather a group of people, or even allow maintenance for everyone. It must be possible for a non-technical person to maintain the glossary. After this requirement, the previously mentioned simplicity becomes even more important. Requirement “easy to use” refined in this consideration into “Must be simple to use” and requirement “Easy to maintain and expand by several people” refined into “Maintenance must be simple” and “Maintenance must not depend on one person”. A new requirement arose as well from refining good usability, “The solution should not introduce new software”.

Accessibility is an issue in a large company where some teams work behind locked doors with classified patient or customer data and there are various tools of different accessibility levels in use. It was stated in the interviews that the solution should be automatically accessible to everyone within the health and social welfare department and not elsewhere, and it must not have additional user identification attached to it. This means that the requirement “In a place where everyone can access it” refined into “Must use existing user access controls”.

The interviewees reported that the need for a glossary solution was growing in 2017 and 2018 with the increased communication and cooperation between teams. The solution must be taken into use as soon as possible, so that people can work more efficiently and avoid creating their own solutions any further. A common solution was basically needed immediately. The requirement “Get the solution into use as soon as possible” turned into “Must be usable immediately” when discussed further.

A refined list of end-user requirements that will be prioritized in the next chapter is:

- Must be accessible only within health and social welfare department
- Must include global, national and company terms with descriptions
- Must include terms used in source code and product user interfaces with descriptions
- Must support two languages but enable also single language usage with definitions
- Must be simple to use

- The solution should not introduce new software
- Must use existing user access controls
- Maintenance must be simple
- Maintenance must not depend on one person
- Must be usable immediately

4.4 Prioritization of End-user Requirements

Like in any product development or creation project, the end-user requirements must be prioritized in some way. In this small project, the end-users did not provide a huge list of requirements but the ones they did give were of very high priority. This is a very normal problem in product development and that is why several methods for prioritization of requirements have been developed. For example, some of them have been analyzed by Aasem, Ramzan and Jaffar (70) and presented at the 2010 International Conference on Information and Emerging Technologies. For the limited number of requirements in this project, a simple prioritization method is sufficient. After the interviews with end-users, the requirements follow the MoSCoW method already in their wording, so another method is used in addition to evaluate the requirements further (71). John McIntyre combines using the four levels of MoSCoW with the three types of Kano in his blog post and that is what will be used here as well (71).

The two methods are described here very briefly before using them. In the MoSCoW method each requirement is labelled with either Must have, Should have, Could have or Won't have according to the evaluated importance of that requirement to the success of the delivery. Missing a Must have requirement will result in a failure of the project but often Must have requirements can be negotiated into Should haves. In the Kano method requirements are similarly categorized but into three main groups according to their effect on customer satisfaction: Baseline Expectations, Linear Satisfiers, and Delighters. In this method, the baseline expectations are such that the customers take for granted in the product category, without these the product does not exist in that category or cannot get to markets. The linear satisfiers are those requirements that increase customer satisfaction if implemented well and decrease it if implemented poorly or not at all. Delighters are the wow-effect that the customers do not know to miss but will increase customer satisfaction if implemented. A mix of requirements from each of these categories should be implemented in each delivery when

following the Kano method but emphasis should be more in baseline requirements in the first deliveries and transfer more into delighters in later ones. (71.)

In Kano method, the requirements are mapped on a graph with plots of the three categories. To do that, each of the requirements is given a letter to name them for the graph:

A. Must be accessible only within health and social welfare department
B. Must include global, national and company terms with descriptions
C. Must include terms used in source code and product user interfaces with descriptions
D. Must support two languages but enable also single language usage with definitions
E. Must be simple to use
F. The solution should not introduce new software
G. Must use existing user access controls
H. Maintenance must be simple
I. Maintenance must not depend on one person
J. Must be usable immediately

Figure 8 is a graph showing the relationships of the three categories of Kano method and the requirements of this project mapped on them. For the purposes of this project, the mapping done by the researcher is adequate when considering that the interviewed end-users prioritized the requirements already with one method. Logical deduction was used to categorize each of the requirements for the mapping in the Kano method keeping in mind the descriptions of the categories and the product group in question. It is noteworthy here, that Figure 8 does not show the order of implementation of the requirements and in this case project, when the requirements have not been implemented, the graph does not show the current customer satisfaction level produced by the requirements either. The graph shows the possible influence of the requirements on customer satisfaction in a color-coded fashion.

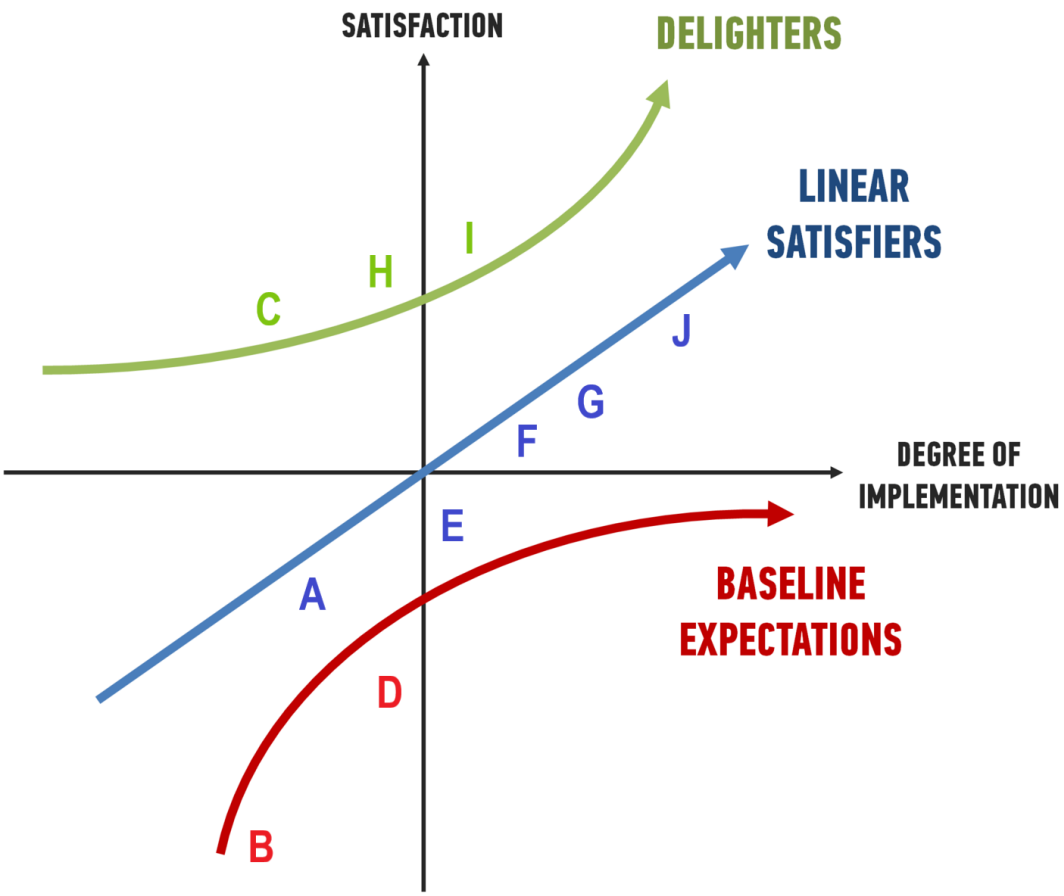


FIGURE 8. Kano categories plotted (71) and project requirements mapped on them

As can be seen in the Figure 8, only two of the end-user requirements were mapped as baseline expectations. This reveals that there is a risk that the end-users did not mention some of the most basic requirements in the interviews. Also, only three requirements got labelled as delighters. This could be because of the urgency of having a solution of some kind and not being able or willing to request special wow-features into it. Half of the requirements got labelled as linear satisfiers. This is important to take note of because if the users evaluate these requirements poorly implemented, they might not continue using the solution after a while.

The solution according to this evaluation must fulfill requirements B and D for it to be a solution at all. Requirements A, E, F, G, J should be implemented carefully to ensure user satisfaction and requirements C, H and I add value if they are available but missing them will not prevent or stop using the solution. In the MoSCoW method, all the requirements were labelled Must have except for requirement J, that was a Should have. The Kano method gave a good additional view to the requirements in this case when compared to just using the MoSCoW method.

4.5 Requirements from Literature Study

Requirements from the Literature Study section of this report are extracted in this chapter. The requirements are reflected against and restructured with information from the previously presented end-user requirements. The end-user requirements are not covered here for the most part but only their effect in prioritizing the basic terminological requirements. The requirements are prioritized using the MoSCoW method that was used for the end-user requirements as well. The prioritization is done by the researcher based on the end-user requirements and information from the literature study. Requirements for the tool, the process and terminology management are listed separately.

The tool related requirements in priority order by classification are:

1. The solution must enable entries of several special languages, like medical terms, social welfare terms, client administration terms, software development terms and programming terms.
2. The solution must enable entries of general language.
3. The solution must enable both descriptive and prescriptive terminology entries. Both purposes of terminology work must be supported because of the end-user requirement of including both national and in-house terms. Nationally defined terms should be considered prescriptive and some in-house terms can become prescriptive as well, but for the most part the terminology should be descriptive and guiding term usage.
4. The solution must support terms and appellations. This is an end-user requirement for terms, but appellations are included here as well because they do not introduce new technical requirements.
5. The solution must support minimum entry information: Term, definition or example textual context, domain and subdomain, source, date, responsibility.
6. The solution could support concept systems and diagrams or in minimum not prevent extending to this direction later. They were not required by end-users, but they are rather basic elements in terminology work.
7. The solution could support symbols. The end-users talked only about words; terms and appellations are covered with that, but symbols pose a technical requirement on the solution that could be considered even though they were not mentioned by the end-users.
8. The solution should not be machine readable i.e. an ontology. The solution is meant for human usage only.

The process related requirements in priority order by classification are:

1. The process must define the terminology workflow, for example: needs analysis, source documentation, term extraction and selection, term entry elaboration, quality check, dissemination, maintenance.
2. The process should ensure limited number of synonyms for / in entries to facilitate consistent term usage in documentation.

The terminology management related requirements in priority order by classification are:

1. Terminology work must have defined objectives.
2. Terminology work benefits should be maximized by having only one source for terminology.
3. Terminology work benefits should be maximized by as wide as possible usage.
4. Terminology work benefits should be maximized with a thorough preliminary study before big investments.
5. Terminology work and process should be developed and evaluated continuously.
6. The core processes supported by the terminology work should be identified.
7. Terminology work should be fitted and scheduled to the supported core processes in the company.
8. Terminology work related roles should be in minimum roughly defined.

Much more requirements were extracted from the literature study when compared to the number of end-user requirements. Also, the requirements were much more detailed and for the most part concerning different things than the end-user requirements. Together, the requirements form a rather good set of requirements for later phases of the project.

5 ANALYSIS OF EXISTING SOLUTIONS

Existing practical word collections are analyzed in this section for compliancy against the previously defined requirements. To do that, a modified list of requirements applicable for the compliancy analysis is created first. Then, a selection of dictionaries, vocabularies, thesauri and ontologies as well as previously created company internal vocabularies are briefly described. Finally, the compliancy analysis is done, and conclusions are drawn from it.

5.1 Requirements for Compliancy Analysis

Not all the collected requirements are valid for the existing word collections and some must be divided into several requirements to be able to define compliancy against them unambiguously, so a modified list of requirements for compliancy evaluation is created first. For example, the inclusion of different languages, special languages and different contexts of terminology are divided into distinct requirements to evaluate what the existing solutions include.

The requirements related to terminology management can be ignored here, because they do not apply to a collection of words but apply on a higher level. The process related requirements can be combined to form one requirement of enabling process development, because the solution is not the process, but they do influence each other. User access related requirements are combined to one requirement. It tells whether the solution is freely available or limited somehow. Enabling user access control is required only if the solution includes in-house terminology. Maintenance related, and further development related requirements are combined into one, namely enabling extension and maintenance for the company. For existing solutions this is more critical than the details of how it is done. Existing solutions are ready for use, but the commercial ones require procurement and related activities, so the requirement for immediate usability separates the commercial ones from the ones freely available online.

This list of 30 requirements is used for compliancy analysis and comparison of a selection of existing word collections later in this chapter:

1. The solution must enable controlling user access.
2. The solution must include global terms.

3. The solution must include national terms.
4. The solution must enable including in-house terms.
5. The solution must include source code terms.
6. The solution must include product user interface terms.
7. The solution must include special language of medicine.
8. The solution must include special language of social welfare.
9. The solution must include special language of client administration.
10. The solution must include special language of software product development process.
11. The solution must include special language of programming.
12. The solution must include general language.
13. The solution must include Finnish language.
14. The solution must include English language.
15. The solution must include descriptions for terms.
16. The solution must include descriptive terminology entries.
17. The solution must include prescriptive terminology entries.
18. The solution must include appellations.
19. The solution must include entry domain and subdomain information.
20. The solution must include entry source information.
21. The solution must include entry date information.
22. The solution must include entry responsibility information.
23. The solution must enable process development.
24. The solution must be simple to use.
25. The solution must enable maintenance and extending.
26. The solution must be usable immediately.
27. The solution could include symbols.
28. The solution could include concept systems and diagrams.
29. The solution should not introduce new software.
30. The solution should not be machine readable i.e. an ontology.

5.2 Descriptions of Existing Word Collections

This chapter briefly describes a selection of dictionaries, vocabularies, thesauri and ontologies as well as previously created company internal vocabularies. The company internal solutions are left

out of the compliancy analysis but are presented here briefly to create a baseline for the implementation phase. Rest of the selection is based on information from the interviewees as well as searches done online by the researcher. Only digital solutions are selected for usability reasons, and the solutions must include more than a trivial number of terms with definitions in at least one of the required languages. Also, if the solution does not seem to have been updated in this decade, it is ignored as outdated.

5.2.1 Company Internal Vocabularies

In-house vocabularies of different domains in varying sizes and technical solutions were discovered during the project. One technical solution was found in the product development management system and there were at least ten separate implementations of it with varying and partially overlapping content. The implementations contained words with definitions and some also with translations. The number of words varied in the implementations between 0 and 900. Technically the implementations consisted of term specific objects and pages. Two different solutions were found in the company wiki for the department in question and one for a higher organization level. These solutions included word amounts between some tens and slightly over 2 000. The technical solutions were a simple table on a wiki-page and a spreadsheet attached to a wiki-page. All these vocabularies are left out of the following compliancy analysis but are presented here as a baseline for the implementation phase.

5.2.2 Code Service

Code Service provides the nationally standardized data structures required by the electronic client data systems in social welfare and health care. The Code Service forms a part of Kanta, the national patient data system. The service is provided by the National Institute of Health and Welfare (THL) and the technical maintenance of the code server is provided by the Social Insurance Institution of Finland. The data structures include code sets, classifications, form structures, texts, register data as well as vocabularies and terminologies related to them. The Code Service usage is free of charge and the service is in Finnish. (72.)

5.2.3 Duodecim Dictionaries

Duodecim Dictionaries is an online collection of medical dictionaries provided by Duodecim Medical Publications Ltd. The included dictionaries are Medical Terms, Medical Finnish-English-Finnish -dictionary and MeSH. The Medical Terms includes Finnish terms and definitions and provides equivalents in Latin, English and Swedish. The Medical Finnish-English-Finnish -dictionary contains roughly 84 000 terms in Finnish and 74 000 in English. The MeSH content is included in the dictionaries' common search function and contains translations to seven languages including English definitions and guidance. Duodecim Dictionaries have user-access control and the service is available to their customers online. (73.)

5.2.4 Finto

Finto is a Finnish online service for publishing and using vocabularies, ontologies and classifications. The service is currently developed by the National Library of Finland, the Ministry of Finance, and the Ministry of Education and Culture. Finto started as a research project called FinnONTO in the Aalto University and the University of Helsinki in 2003. The service has an online user interface for browsing vocabularies and open interfaces for using them elsewhere. (74.)

5.2.5 Hoidokki

Hoidokki is a thesaurus published by the Finnish Foundation of Nursing Education (SHKS) in 2005. Hoidokki can be used to search and to index publications of nursing science. Hoidokki consists of ten themes with hierarchically structured terms of up to five levels and definitions in Finnish, Swedish and English. Technical solution of Hoidokki is created and maintained by MI Tietorakenteet Oy and the service is available online. (75.)

5.2.6 Kielitoimiston Sanakirja

Kielitoimiston sanakirja is an online dictionary of standard Finnish and it is maintained by the Institute for the Languages of Finland. The dictionary contains words in contemporary Finnish with information on the meaning, usage, register and style. Kielitoimiston sanakirja follows the decisions

and recommendations of the Finnish Language Board. The dictionary contains currently over 100 000 entries. (76.)

5.2.7 MOT Language Services

Kielikone Ltd. provides digital and online dictionary solutions and services commercially in Finland. Companies can add their own internal vocabularies into their MOT language service. All the MOT services are created in co-operation with world-class dictionary publishers and language experts and the contents are kept up to date. MOT language services are accessible on computers as well as mobile devices. (77.)

5.2.8 Oxford Dictionaries

Oxford Dictionaries belongs to Oxford University Press (OUP), a department of the University of Oxford and they have been making dictionaries for over 150 years. Oxford Dictionaries is a global organization and provides dictionaries in several languages online, mobile and in print. The dictionaries provide definitions to words as well as history, usage and relationship to other languages. (78.)

5.2.9 TEPA Term Bank

TEPA is a free-of-charge term bank containing about 355 000 terminological entries with terms and definitions in several special languages. TEPA is maintained by the Finnish Terminology Centre TSK (Sanastokeskus TSK) who receives financial assistance from the Finnish Ministry of Education and Culture for maintaining TEPA and for keeping it freely available. TEPA contains vocabularies, dictionaries and terminological databases compiled by both TSK and other experts. The most common languages in TEPA are Finnish, Swedish, English and German. (79.)

5.2.10 Valter

Valter is the term bank of the Finnish Government. It is a multilingual term bank that contains currently 15 glossaries compiled by several Government agencies. Glossaries include for example

public buildings in Finland, Government agencies and other public bodies, court terms and budget glossary but not all the glossaries have the same structure and classification of terms. Technical solution of Valter is based on MOT language service and it is available online. The Prime Minister's Office updates Valter glossaries whenever resources are available for it. (80.)

5.3 Compliancy Analysis

Compliancy analysis is done with the help of a simple table including all the previously modified requirements and the previously presented solutions. The compliancy analysis of existing solutions is provided in appendix 1 of this report. In the analysis each requirement is evaluated either compliant "x", partially compliant "(x)" or not compliant "-" in the solution. The analysis is done based on the information available in the solution for each requirement. For example, the requirement of including global terms is checked through what word collections the solution includes. In this case, including European Union terms or the Medical Subject Headings (MeSH) is a clear indication of including international terms.

Some of the solutions' descriptions and help files do not provide enough information to determine whether the solution complies to the requirement. In these cases, example words are searched for in the solution to determine compliancy. Words like "asumistuki" (housing allowance), "adoptioneuvonta" (adoption councelling), "päihdetyö" (social work with substance abusers) and "aikuiskoulutustuki" (financial aid for adult students) are used to evaluate whether the special language of social welfare is included in the solution. Another example set of words used in this study is words for "koti" (home) and "syöminen" (eating) for general language.

Finally, the number of compliant and partially compliant requirements are calculated for each solution. In this case and for this project purpose it was adequate to weigh each requirement similarly even though the requirements are not equally important. The compliancy analysis table provided in appendix 1 shows significant differences between the solutions. When counting the sum of partially compliant and compliant requirements for each solution, the sums very between 17 and 27 of the maximum 30. Table 3 lists the solutions in compliancy order with the numbers of compliant and partially compliant requirements as well as the sum.

TABLE 3. Results of Compliancy Analysis of Existing Solutions

Solution	Number of Compliant Requirements	Number of Partially Compliant Requirements	Sum
1. MOT Language Services	19	8	27
2. TEPA Term Bank	18	7	25
3. Finto	13	12	25
4. Duodecim Dictionaries	14	7	21
5. Kielitoimiston sanakirja	13	8	21
6. Oxford Dictionaries	11	10	21
7. Code Service	16	3	19
8. Valter	15	3	18
9. Hoidokki	14	3	17

The most compliant solution is MOT Language Services with an appropriate dictionary set that influences the results significantly. This is the only solution of the selection anyhow that would require procurement and configuration procedures in the case company. Valter and Hoidokki that are least compliant are also very limited and specialized in their content. They are surely useful if in need of terminology in their specialized content but for general usage they are not evaluated very useful.

6 IMPLEMENTATION

This chapter first describes the plan and process for the implementation phase of the solution. After that, three main parts of the development work are presented in their own sections. The development is divided into three distinct levels or parts: the development of the technical solution, the processes of using the technical solution for different purposes and finally the larger scale process, or rather the management of the smaller scale processes.

6.1 Implementation Plan and Process

The end-user requirements strongly indicated that a solution is needed as soon as possible, and it must be taken into use immediately. This was the main issue influencing the implementation plan and process. Another delimiting issue and guiding principle in planning was the operational methodology used at the case company, namely Scaled Agile Framework (SAFe®). The implementation plan and process were decided upon in March 2018 after it was clear that the project could be started from resources' perspective.

Schedule and duration of the main implementation phase of the project were defined by the study leave of the sole full-time person in the project. The main implementation phase was delimited to April and May of 2018 with some implementation work done already in March. After the main implementation phase, a working technical solution would have to be in use and the processes related to it defined. It was decided in March that the development would happen in one-week iterations and the progress and results would be shown and discussed in a demo event once a week until the participating end-users would agree that the solution is acceptable. Usually in SAFe® and other agile methodologies as well, a sprint duration of two weeks is used, but for this project size and people involved it was more suited to use one-week iterations. The demo event functioned also as sprint planning for the next iteration.

The implementation plan was followed quite accurately and only the first planned demo was missed due to illness. Figure 9 describes the realized implementation schedule of the three main development levels on a weekly basis with the weekly demo dates. The first weeks were focused in developing the technical solution, and then the focus shifted to development of the processes

and the management process. In the fifth demo on 11.5. it was decided that no further development is needed for the technical solution or the processes before more extensive user experience is gained. No changes had been made to the solution or the processes during the previous week and no new feedback was given regarding those. It was the last demo event of the project. The 23.5. event was an organization meeting for a glossary CoP (Community of Practice) following the SAFe®. This meeting ended the development of the management process in the scope of this project and handed the ownership to the Glossary CoP. The CoP and what lead to it is explained further in the chapter "6.4 Development of Management Process".

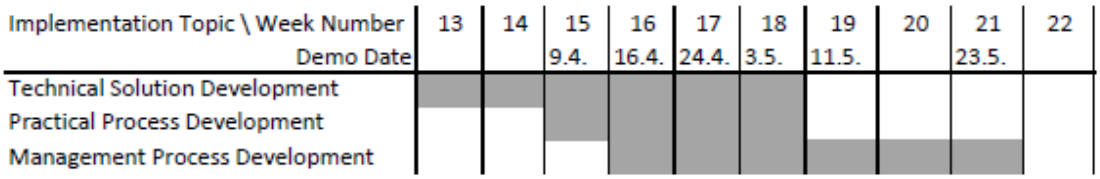


FIGURE 9. Realized Implementation Phase Schedule

The researcher invited and hosted the events in a meeting room at the case company facilities as well as online for remote participants in several locations. The demo invitations were sent to an ever-growing group of people except for the last invitation. The group first consisted of people from one product team, but before the last demo there were people from nearly all the product teams in the targeted scope. After each event a short memo was sent to all the people in the invitation list regardless of their participation in the event to keep them up-to-date on the project even if they happened to have an over-lapping appointment.

For a small project like this, it was decided that no project management software or task tool would be needed since using such would bring more overhead to the project than bring benefits. Records of the project contents and progress were kept in e-mail and naturally in the technical solution itself. The demo events had varying agendas depending on the focus of the recent development and the feedback from the participants. Further requirements and details were discussed in these events as well. The first three demo events lasted for an hour and the last two demo events were half-an-hour long. This reflects also the amount of changes and further change requests discussed in the events.

6.2 Development of Technical Solution

Development of the technical solution in the case company does not start from zero in this project because of the previously created company internal vocabularies. A study regarding the existing internal solutions is presented briefly in the chapter “5.2.1 Company Internal Vocabularies”. The study revealed that there were several slightly different types of word collections already created in the case company but no specific software for terminology work or vocabularies. The technical solutions varied from a table on a wiki-page and a spreadsheet attached to a wiki-page to term specific objects in the product development management system. Also, the online service of Duodecim Dictionaries was accessible in the case company even though not known by many.

6.2.1 Fundamental Decision

The interviewees and the people participating in the demo events were not aware of all the existing solutions but knew about one or two of them each. It became clear to the researcher already during the preliminary phase of the project that rather than developing a highly technical solution or procuring a commercial solution, the end-users would benefit more from a consolidation project that would aim at using one common solution and one process for terminology work in all the products and projects. A consolidation project would also result in a better understanding of the needs for a possible terminology management software in future. This rationale was also backed up by some of the terminology management related requirements listed in chapter “4.5 Requirements from Literature Study”:

- Terminology work benefits should be maximized by having only one source for terminology.
- Terminology work benefits should be maximized by as wide as possible usage.
- Terminology work benefits should be maximized with a thorough preliminary study before big investments.

The end-users required that the solution should not introduce new software that requires procurement procedures, configuration and end-user training before it could be taken into use. A design decision was made to select the best suited existing technical solution and modify it according to the requirements as well as include the terminology from the other solutions into the selected one. This decision enabled continuous use of a solution i.e. the end-user requirement of

having a solution immediately in use got fulfilled. It also freed resources to the process development and consolidation efforts that are described later in this report.

6.2.2 Name and Location

The solution should be a vocabulary according to the various word collections' definitions mentioned in the chapter "3.1.2 Collections of Words" and because the solution would include several special languages and general language words. According to the mentioned definitions, a glossary is often considered less formal than a vocabulary. It anyhow became clear during the literature study and the study of existing solutions that the term glossary is quite freely used in these kinds of contexts in Finland. As an example, Valter that is described previously in this report, includes 15 glossaries, not vocabularies, even though it is a publicly available national term bank. Glossary was selected as the name of the solution in this project to make it easy to approach for everyone.

There were several discussions in the interviews as well as the first demo events regarding the locations of the existing in-house solutions as well as the ideal locations. The product development management system was evaluated more difficult to reach and use than wiki. Not all the end-user groups use the product development management system and they would have to start using it for the glossary only. Wiki was evaluated sufficient for usability in the solution that had a simple table of words on the wiki-page. The simple table -solution was evaluated impossible to use anyhow with a greater number of words. The embedded spreadsheet -solution was considered far too complex for maintenance. It required making a local copy of the spreadsheet and then replacing the old one on the wiki-page after an update and it could not keep track of concurrent updates. Other feasible locations were evaluated, and the document management system got selected because everyone in the department had to use it occasionally or daily and everyone must have access rights to it. It also would take care of versioning and concurrent updates. A file could be opened as read-only, and editing could be enabled only if changes were needed.

6.2.3 File Format

It was suggested in some of the interviews that a spreadsheet solution would be quite flexible and user-friendly. The only downside of the existing solution done with a spreadsheet was the location

that made the usability and maintainability poor. Spreadsheet applications generally support various import and export options, so it was evaluated that it could be rather easy to integrate the various existing solutions into one spreadsheet and then in future to export some sections of the solution according to needs. It was also evaluated by the researcher that if a commercial solution of some kind is decided upon later, a spreadsheet application provides some chances for easy integration in that situation as well.

The researcher decided that a spreadsheet would be used at least as the first solution when she was planning the implementation phase. The file format issue was discussed in the first demo event and it was agreed that the presented spreadsheet solution is acceptable for this project because it enables fulfilling the end-user requirements. It is a basic principle in agile methodologies to create the minimum viable solution first. This project had also other requirements that the file format alone could not solve, so focus was transferred to those and eventually the file format was not questioned even by the new participants in the later demo events.

6.2.4 Sheets

An existing spreadsheet solution was taken as baseline for the development of the glossary. The baseline glossary had only one sheet with the glossary contents on it and no description or other information about the glossary at all. Wiki-page solutions had a brief description on the top of the page before the content table, but the solutions were so simple otherwise that they were evaluated quite unhelpful in the development phase. The researcher concluded that to keep up with the process and keep the users informed about the solution, some helpful sheets would be needed in the spreadsheet. This would also guarantee that the glossary file is not separated from the process or the other needed and related information. The created sheets were bilingual each, all the information was provided both in Finnish and in English side by side. The glossary spreadsheet had four sheets starting from the first demo event until the end of the implementation phase:

- Description-sheet that provided a general description of the glossary, objectives of it and brief information about the Glossary CoP.
- Guide-sheet that described how to use and maintain the glossary.
- Search-sheet that provided links to selected online services to search for the terms before including them in the glossary.
- Glossary-sheet that contained the glossary contents.

The Description-sheet provided information on who the glossary was meant for and how everyone could expand, update and maintain it. The sheet explained the larger scale objectives of the glossary work that are described in further detail in chapter "6.4 Development of Management Process" of this report. The sheet also had contact information of the Glossary CoP if the user would have any questions or ideas or comments regarding the glossary. The Guide-sheet took care of describing the glossary related processes, so it is described in this report in chapter "6.3 Development of Practical Processes".

The interviewees nor the demo event participants knew about the various company specific solutions created earlier. The researcher concluded that the people most probably did not know about other terminology related services either. Hence, a Search-sheet was created to help the glossary users find correct translations and definitions to terms they were looking for even in the situations that the term was not yet included in the company internal glossary. The online services that got over 20 as the total number of compliant and partially compliant requirements in the compliancy analysis were added to the Search-sheet as links. These services were presented to the demo participants in the first and second demo event and they were very happily received.

6.2.5 Data Fields

The largest existing spreadsheet solution by number of words was taken as baseline for the development of the glossary. The glossary had fields:

- Finnish Term
- Finnish Status
- Finnish Note
- English Term
- English Status
- English Note

The Note-field was used for all information regarding the term: definition, source and domain. Most of the terms did not anyhow have any notes and those that had notes included a varied combination of the information. The Status-field had mostly values "Approved" but there were also some with values "To be checked" and "Rejected". Over 250 rows did not have a status defined.

The existing data fields did not fulfill the minimum entry information requirement mentioned in chapter “4.5 Requirements from Literature Study”: “The solution must support minimum entry information: Term, definition or example textual context, domain and subdomain, source, date, responsibility.” The existing solution did not have any information about the when and by whom the items were added, accepted or rejected in the glossary. The Note-field contents would have to be divided into several separate fields to make the glossary more readable and the missing information more visible. New fields were needed for the domain and source information and the Note-field would need to be renamed. After adding the new fields to the glossary, all the entries with any information in the Note-field would need to be browsed through and divided into the new ones accordingly.

A lot of work was done before the first demo event with the structure of the spreadsheet, and in the first demo event, the glossary had entry information fields:

- Finnish Term
- Finnish Synonyms & Definition
- English Term
- English Synonyms & Definition
- Status
- Term Group
- Context
- Notes
- Source

These fields and their purpose were presented more closely to the participants in the second demo event because the first one stayed on a more general level presenting the project and the fundamental decisions made. The term, definition and source fields were clear to the participants, but all the other fields required some explaining. Some of the participants first thought that no context or term group information would be needed but it was easy to justify their existence with the several product teams and several user groups that would be using the glossary. People generally struggled in the event in understanding the difference between the Context-field and the Term Group -field. Some existing example terms were presented as well as some examples thrown by the participants, and the fields started to make sense. Context-field denotes the product or a combination of products or even more generic level of context where the term is used. Term Group

-field on the other hand denotes the subject field within that context or the special language the term is valid in. A general language word would be marked with “General” in both fields.

The Context-field and the Term Group -field were developed further by implementing drop-down lists for them with fixed value options. A macro was developed for the Context-field, so that it supports multiple selection. The same functionality was not developed for the Term Group-field because it was discussed in a demo event and the decision was that it is unnecessary for term group, but much needed for the context.

There were some suggestions and feedback about the value options in the demo events mainly on 16.4. and 24.4. One correction request was done in the 3.5. demo event as well. During the implementation phase of the project, the researcher made these corrections and implemented the development suggestions. It was decided in the demo events that after the implementation phase, the Glossary CoP that was to be founded would be responsible for the maintenance and changes in the value options. More information about the Glossary CoP can be found in chapter “6.4 Development of Management Process”.

In the third demo event on 24.4. the data structure of the spreadsheet had reached its maturity with the last changes presented regarding the entry information fields:

- Finnish Term
- Finnish Synonyms & Definition & Comments
- English Term
- English Synonyms & Definition & Comments
- Reliability
- Term Group
- Context
- Change History
- Source

The Term Group and Context -fields were still further developed with pop-up instruction functionality for the last demo event on 3.5. Figure 10 shows the pop-up instruction for the Term Group -field that becomes visible when the field is selected.

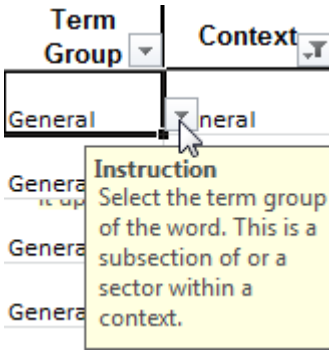


FIGURE 10. Term Group -Field Pop-Up Instruction

The Status-field and its purpose was also presented with details in the second demo event on 16.4. and the name of the field had been changed to Reliability by the third demo event on 24.4. This field is so closely related to the work process of the glossary that the field contents and the development of it is described in more detail in chapter "6.3 Development of Practical Processes". Another bigger change was changing the previously called Notes-field into Change History. This field is also closely related to the process of each term entry in the glossary, so details of the development are presented in the process development section of the report.

6.2.6 Content Development

The baseline glossary contained 2 055 rows of words and phrases in the beginning of the implementation phase. The glossary was clearly meant for translation purposes with words in Finnish and nearly all of them translated to English. The glossary had a Note-field for each language, and some of the terms had a definition type of description in it either in Finnish or both in Finnish and English. Also, some of the terms had a source mentioned in the Note-field for the definition but not all of them. Some of the terms had information about the context of the term in the Note-field rather than about the term itself.

The interviewees had reported that the seemingly large glossary, even though inviting with its size, had problems that had resulted in it not being used by most of the interviewees any more. The glossary was reported having the following problems:

- The glossary did not contain needed terms.
- The glossary did contain useless words like company names and web addresses.
- The terms could have a status "approved" but the user did not agree with the translation.

The researcher studied the glossary contents briefly when evaluating the best suited baseline solution before the implementation phase of the project. The quality of some of the glossary contents was a problem, but the solution still was the best option for a baseline solution. It was decided that a heavy maintenance and cleaning operation would be needed before developing the glossary further. Nearly 400 rows were removed before the first demo event. The removed rows included for example exact duplicates, various inflectional forms in Finnish, web addresses and company names.

One of the other existing solutions was integrated to the baseline solution before the first demo event to make the solution contain more relevant terms for the participating end-user groups. This made the previously created Context-field relevant as well. The glossary contained roughly 1 700 terms in the first demo event, and after a brief presentation of the contents the demo event participants agreed that the worst quality problems were fixed, and the solution is going to the right direction. During the implementation phase further maintenance type of cleaning was done in the glossary and some relevant terms were added also to test the created processes. The glossary contained about 1770 terms at the end of the implementation phase. Appendix 2 of this report contains a screenshot of few selected entries in the glossary at the end of the implementation phase. Only a few words in the general context are selected.

6.3 Development of Practical Processes

6.3.1 First Process Version

The glossary process development started after the first demo event on 9.4. The first version of the process was presented after one iteration in the second demo on 16.4. The first process version had entry Status-field values:

- New
- Proposal
- Approved
- Removed

The principle in this process was that the person who added a word in the glossary would select the status “New” and add as much information as available into the various fields for the word. This

person would need to add their name and date in the Notes-field. The next person in need of the same word in the same context and term group would check if the word information was correct and adequate in their opinion. If not, they would have to contact the person mentioned in the Notes-field and try to get to an agreement about the definition and translation mainly. If agreement could not be reached, the second person would create a duplicate entry with their own information and select the status “New”. It the two people reached agreement and possibly added or fixed the information, the status would be changed to “Proposal” and the second person would also add their name and date to the Notes-field in addition to the previous one. The third person in need of the same word would then again contact the previous two for adding more information if available. The third person would then change the status to “Approved” and add their name and date third in Notes-field for that entry. The “Removed”-status could be selected by anyone if there was some notification from authorities for example for changing a national term. The reason for changing an entry status to this one would have to be stated in the definition, and again name and date would be added in the Notes-field.

The first process description was discussed in the demo event for quite some time. The participants thought it seemed rather heavy and they first assumed they would not be allowed to use the words until they had status “Approved”. It was also suggested that the word “New” should be replaced with “Draft”. The glossary spreadsheet had a Guide-sheet describing the process to the users and the participants requested that a specific instruction should be added allowing usage of terms in “New”-status. There was a discussion about what the different status values mean and the conclusion was that they are rather an indication of reliability of the information than a permit to use a word. The words with “Removed”-status raised some questions as well regarding whether to keep them in the glossary at all. It was agreed to keep them in the glossary because they could be filtered away if needed and they would help prevent adding the same outdated or wrong words several times. It was clear to the researcher after the second demo event that the process required some more work to make it easier to understand and lighter to follow.

6.3.2 Approved Process

After some consideration the researcher concluded that most of the terms that would be added to the glossary would be copied from some source for easier access and for those terms a three-stage acceptance process is heavy. The medical and social welfare terminology should be provided

nationally to guarantee the nationally required system integration. Programming related as well as software development related terminology could possibly be found online or in some course material if needed. The Source-field would be important for these entries. If the translations and definitions would have to be created in-house for the most part, the situation would be different. Some terms surely are created also in-house but even then, they rarely are the handiwork of just one person. For the purposes of terminology work in the case company a simpler process would be adequate at least in the beginning of the consolidation and harmonization effort and the Glossary CoP would follow up on the process later.

For the third demo event on 24.4. the glossary process had changed quite radically. Instead of a Status-field, the glossary had Reliability-field, and instead of Notes-field there was a Change History -field. The Change History -field operated the same way still as before. It was a manual field that the user would have to fill in whenever modifying an entry. In the demo event it was agreed that the user ID for the company system was good for this purpose because it would be easy to find the person for more information if needed. The user ID was individual and adequately short. The Reliability-field had a drop-down selection with predefined five-level star-rating and a "Do Not Use"-option in addition. This can be seen in Figure 11 like it was presented in the demo event on 24.4.

FIGURE 11. Star-Rating for Reliability

The new lighter process was considered good and the participants only wanted to have the pop-up instruction for this field as well like it was already done for another more challenging field. People especially thought that five levels were good because it would make the selection easier when adding or modifying an entry. The pop-up instruction functionality was implemented and presented in the last demo event on 3.5. This pop-up instruction is shown in Figure 12.

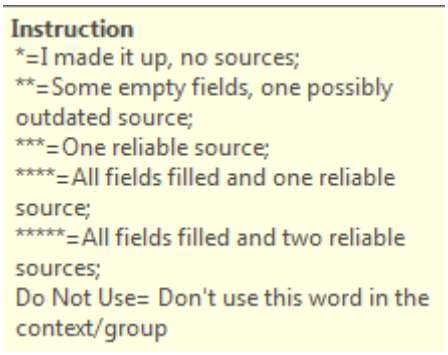


FIGURE 12. Pop-Up Instruction for the Reliability-Field

The Guide-sheet took care of describing the process with more details than the pop-up instructions. The guide was based on use-case type of instructions for various situations in the glossary including adding items and getting them out of use as well as using the glossary and improving the reliability of the entries. The instructions were provided both in Finnish and English and updated during the project each time there was changes into the data fields and the process. Table 4 presents an English example like it was in the end of the implementation phase. The selected example process is for adding an item into the glossary which is referenced from many of the other processes as well. Company specific information is garbled in the example.

TABLE 4. The Process of Adding a Word into the Glossary

Action	I need to add a word into the glossary.
Preconditions	1. The word is not in the glossary. 2. I need the word several times and cannot remember the definition or translation easily.
Instructions	1. I add the word into the glossary with the information I have found. 2. If needed, I ask a colleague for a hint, so that I can fill in at least two of the fields A-D according to my usage needs. 3. Term Group and Context according to my usage needs. 4. Reliability: * = I filled in the information from my own head or feel otherwise unsure of the correctness; **= Missed a couple fields but the word is defined in one language in a 20 year-old paper dictionary; ***= One of the languages has an appropriate definition matching the context online and I added the link, but the other language is missing or I translated it myself; ****= The word is defined in both languages but with inadequate references; *****= The

	word has both languages with context matching definitions available online; Do Not Use= This word is not to be used. 5. I add my xxx id with a date to Change History-field. 6. I save an updated version of the glossary.

6.4 Development of Management Process

There were two objectives defined in the preliminary phase of the project for the glossary and the processes:

1. To ease any translation work needed in the development of healthcare and social welfare solutions in the case company.
2. To harmonize term usage both in Finnish and in English in all documentation and source code.

These two objectives were objectives for the whole terminology management in this project and after the project they are the objectives of the Glossary CoP. These objectives were mentioned also on the glossary Description-sheet already in the first demo event, so that everyone participating the project would know them.

The management process was just briefly mentioned in the first demo event on 9.4. because it was only known at that point that such a thing would be needed as well for the solution to be viable for a longer time, but the researcher had not begun planning it yet. In the second demo event on 16.4. a few more issues were mentioned about the management process when the practical process was described, and that demo event can be considered as the starting point for developing the management process. A this point the group running the management process was still called a SIG (Special Interest Group) which is a rather known term in computer industry for any organized group advancing a specific technical or non-technical issue. This term stayed until the last demo event on 11.5. when the group got renamed CoP (Community of Practice) to match the SAFe® and enable fitting the process easily to other processes in the case company. The Glossary CoP organization meeting on 23.5. ended the development of the management process in the scope of this project and handed the ownership of the processes and the glossary to the Glossary CoP.

The first task of the Glossary CoP mentioned briefly in the first demo event was the maintenance of the glossary. In the following demo events the responsibilities and tasks of the Glossary CoP were clarified further and before the end of the implementation phase the tasks included:

- Management of the glossary for example by making changes to the data fields or the drop-down selection options.
- Maintenance of the glossary for example by combining duplicate terms.
- Process development when needed.
- Evaluation of the need for a commercial solution.
- Term recommendations in the glossary in future to harmonize the used terminology.
- Follow up on the national and other terminology efforts outside the case company and making that knowledge available in the case company.
- Spreading knowledge about terminology work and the glossary in the case company.

Other aspects of the Glossary CoP were covered mainly in the last two demo events, like for example what kind of people would be needed in the group and how big the group should be and how often it should meet. It was estimated that a group of about ten people would enable adequate representation of various roles and product development teams. The first plan was to have the group meet once a month. Most importantly, people in the group would have to be interested in this type of work and the group members could change according to changing situations and projects. The people participating in the demo events started asking around in their other meetings for possible candidates and some roles were so well represented in the demo events that they had negotiations of who would participate in the Glossary CoP. In the organization meeting on 23.5. there were eight people present and it was agreed that a couple more could be asked still.

7 RESULTS

The results of this thesis work are examined through a compliancy analysis of the implementation against the previously set requirements. Appendix 3 presents a complete table of end-user requirements and requirements extracted from the literature study as well as researcher's evaluation of compliancy to each of them reached during the implementation phase of the project. Only some of the most significant requirements and their implementation compliancy are presented in this chapter and further details should be studied in the appendix.

The implementation is fully compliant with all the nine end-user requirements, so it is not a surprise that the end-users were very pleased in the demo events. The ambiguous requirements of "Must be simple to use." and "Maintenance must be simple." were discussed in the demo events and the implementation was compliant after some development iterations. The requirements extracted from the literature study were divided into three separate groups:

- eight tool related requirements
- two process related requirements
- eight terminology management related requirements

All the tool related requirements were fulfilled during the implementation phase according to the analysis in Appendix 3. A spreadsheet solution with macros is very flexible.

There were two process related requirements extracted from literature study. One of the requirements stated that the process must define the terminology workflow. The implementation of this requirement was analyzed to be compliant even though the workflow is shorter and simpler in the developed agile terminology management than the workflow of professional terminology organizations described in the literature study. The developed agile terminology management process does not need to focus on the beginning part of the workflow like the process of the separate terminology organizations needs to. This could be achieved because the terminology work happens within the core processes in the developed process for agile terminology management. The developed processes and the tool do enable also following the terminology workflow described in the literature study, so a separated terminology organization and process can easily be extended in future from the agile process developed now. The other process related requirement stated that the process should ensure limited number of synonyms for each term. This implementation was

analyzed not compliant because the developed glossary structure or the processes do not limit the number in any manner. This quality aspect is left for Glossary CoP to follow up with term recommendations and maintenance. The number of synonyms is not a problem when starting organized terminology management, but it can become a problem later.

Eight requirements were extracted from the literature study regarding terminology management. Most of these requirements are such that they must be re-evaluated regularly. In the compliancy analysis the evaluation is done with the current operation right after implementation phase. The new management process has been in use for such a brief time at the time of evaluation that for example the requirement of as wide as possible usage of the glossary cannot be fully compliant yet. The glossary and the processes are meant for everyone's use in the case department. By allowing people to add their own terminology in the glossary, they will consider the common glossary their own which will increase the usage. The project and the Glossary CoP in future make terminology work and the glossary more known in the case company, which will also increase the usage, but this is an ongoing effort.

Another requirement for terminology management was that there should be only one source for terminology. The implementation of this requirement was partially compliant at the end of the implementation phase. People were requested to integrate their own personal word lists into the common glossary and use it in future. Also, there were no more company internal glossaries in the company wiki at the end of the implementation phase. The various solutions in the product development management system were under further analysis by the Glossary CoP at the end of the implementation phase before integrating them. It was already clear that technically the integration would be easy, but the content of the solutions was still being analyzed.

The other six terminology management related requirements were analyzed fully implemented at the end of the implementation phase, but the just established Glossary CoP must follow up on all the terminology management related requirements to make sure the tool and the processes stay up to date and meet the needs of the users also in future. The results of the thesis project are very promising because the newly developed and defined agile terminology processes and tool were happily received and taken into use already. The project exceeded the end-user expectations. Fulfilling the requirements extracted from the literature study created a solution that is viable for a long time in the case company.

8 CONCLUSION

This thesis project aimed at creating and developing an agile terminology management solution for a case company and succeeded in it quite well. This thesis was grounding work in the case company because there was no terminology management process in use in the case department before. There were several different word collections in use in the department and people had also their own personal glossaries to cope with the very information centric work. No agile terminology management process was discovered during the literature study of this project, so one had to be created and developed for the purposes of the agile case company. Also, a tool for maintaining the agile case company glossary was developed in this project.

The developed solution containing the processes and the tool was for the most part compliant with the requirements collected from end-users and terminology work related literature. There were in total 27 requirements collected in the project and at the end of the implementation phase 24 were fully implemented, two were partially implemented due to the brief usage period and one was not implemented. The few requirements that were not fully compliant yet can be developed further by the Glossary Community of Practice (CoP) that was established in the case company at the end of the project to manage the solution.

The creation of the Glossary CoP also gives a strong signal that the work is really needed, and it will be continued in the case company. The Glossary CoP has responsibilities with the solution; the processes and the tool, but also with making the terminology work and the glossary known in the company. Some methods that the CoP could use to increase people's knowledge of the used terminology are for example a monthly terminology newsletter or a term-of-the-week posting in the company internal news. Also, the CoP could perform the questionnaire mentioned in chapter 3.3.1 of this report to make Return on Investment (ROI) estimations for the terminology work in the case company.

There is a lot of changes currently happening at the national level of healthcare and social welfare and it also has implications to both technical and functional requirements of software operating in those fields. The existing national terminology is expanding and changing quickly, so the existing terminological resources require updating. The agile process developed in this project provides a suitable solution for the case company to cope with the constantly changing situation. The

traditional terminology work processes described in the literature are slow for the quickly changing environment of software development.

REFERENCES

1. Wright, Sue Ellen and Budin, Gerhard. *Handbook of Terminology Management : Volume 1: Basic Aspects of Terminology Management.* Amsterdam/Philadelphia : John Benjamins Publishing Company, 1997. ISBN 90 272 2154 5.
2. CGI Suomi Oy. Historia. *CGI.* [Online] CGI Suomi Oy. [Cited: April 24, 2018.] https://www.cgi.fi/historia-suomessa.
3. —. Hyvinvointi. *CGI.* [Online] CGI Suomi Oy. [Cited: April 24, 2018.] https://www.cgi.fi/hyvinvointi.
4. —. Potilastietojärjestelmät. *CGI.* [Online] CGI Suomi Oy. [Cited: April 24, 2018.] https://www.cgi.fi/hyvinvointi/potilastietojarjestelmat.
5. Laki kuntien ja kuntayhtymien eräiden oikeustoimien väliaikaisesta rajoittamisesta sosiaali- ja terveydenhuollossa . *[Act on Temporarily Restricting Specific Legal Transactions in Health and Social Welfare of Municipalities and Joint Municipal Authorities].* Helsinki : Ministry of Social Affairs and Health, June 29, 2016. Act 548/2016.
6. Laki kuntien ja kuntayhtymien eräiden oikeustoimien väliaikaisesta rajoittamisesta sosiaali- ja terveydenhuollossa annetun lain muuttamisesta. *[Act on Amending the Act on Temporarily Restricting Specific Legal Transactions in Health and Social Welfare of Municipalities and Joint Municipal Authorities].* Helsinki : Ministry of Social Affairs and Health, December 28, 2017. Act 1057/2017.
7. Government and Ministries. Timetable. *Regional Government, Health and Social Services Reform.* [Online] March 2018. [Cited: April 26, 2018.] http://alueuudistus.fi/en/timetable.
8. —. Financing of healthcare and social services. *Regional Government, Health and Social Services Reform.* [Online] 2018. [Cited: April 26, 2018.] http://alueuudistus.fi/en/financing/social-and-health-services.
9. Timo T. Seppälä, Markku Pekurinen. Sosiaali- ja terveydenhuollon keskeiset rahavirrat. [Online] 2014. [Cited: April 26, 2018.] http://urn.fi/URN:ISBN:978-952-302-236-2.
10. Ministry of Social Affairs and Health. Terveydenhuolto Suomessa, Sosiaali- ja terveysministeriön esitteitä 2. [Online] 2013. [Cited: April 25, 2018.] http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/69928/URN_ISBN_978-952-00-3389-7.pdf.

11. National Institute for Health and Welfare. SOCIAL WELFARE AND HEALTH CARE REFORM IN FINLAND. *THL.fi.* [Online] 2018. [Cited: April 24, 2018.] https://thl.fi/en/web/social-welfare-and-health-care-reform.
12. the European Parliament and the Council. DIRECTIVE 2007/47/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Strasbourg : the European Parliament and the Council, September 5, 2007. Directive 93/42/EEC.
13. Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä . *[Act on Electronic Handling of Customer Data in Health and Welfare].* Helsinki : Ministry of Social Affairs and Health, July 1, 2007. Act 159/2007.
14. Finlex Data bank. *Finlex.* [Online] [Cited: April 21, 2018.] http://finlex.fi/en/.
15. Kansallinen Terveysarkisto (Kanta). Definitions for Patient Records Archive. *Kanta.* [Online] 2017. [Cited: May 6, 2018.] http://www.kanta.fi/en/web/ammattilaisille/potilastiedon-arkiston-maarittelyt.
16. P. Doupi, E. Renko, P. Hämäläinen et al. eHealth Strategies, Country Brief: Finland. [Online] 2010. [Cited: April 21, 2018.] http://www.ehealth-strategies.eu/database/documents/Finland_CountryBrief_eHStrategies.pdf.
17. World Health Organization. International Classification of Diseases. *World Health Organization.* [Online] WHO, February 23, 2018. [Cited: May 4, 2018.] http://www.who.int/classifications/icd/en/.
18. Komulainen, Jorma et al. *Suomalainen tautien kirjaamisen ohjekirja [Finnish Guide for Documenting Diseases].* Tampere : Terveyden ja hyvinvoinnin laitos, 2012. ISSN 1798-0100.
19. World Health Organization. International Classification of Diseases > ICD-11 Revision. *World Health Organization.* [Online] WHO, October 31, 2017. [Cited: May 4, 2018.] http://www.who.int/classifications/icd/revision/betaexpectations/en/.
20. Health Level Seven International. About HL7. *Health Level Seven International.* [Online] Health Level Seven International. [Cited: May 4, 2018.] http://www.hl7.org/about/index.cfm.
21. HL7 Finland. Briefly in English. *HL7 Finland.* [Online] [Cited: May 4, 2018.] http://www.hl7.fi/briefly-in-english/.
22. Kansallinen Terveysarkisto (Kanta). Services. *Kanta.* [Online] 2017. [Cited: May 4, 2018.] http://www.kanta.fi/en/palvelut.
23. SNOMED International. About. *SNOMED International.* [Online] [Cited: May 4, 2018.] https://www.snomed.org/about.
24. Tolppanen, Esa-Matti. THL valmistelee SNOMED CT käyttöönottoa Suomessa [National Institute for Health and Welfare Prepares Adoption of SNOMED CT in Finland]. *Terveyden ja*

hyvinvoinnin laitos. [Online] THL, March 1, 2018. [Cited: May 4, 2018.] https://thl.fi/fi/web/tiedonhallinta-sosiaali-ja-terveysalalla/-/thl-valmistelee-snomed-ct-kayttoonottoa-suomessa.

25. Ojala, Matti. Ajankohtaista sosiaali- ja terveyssektorin termityössä. *Terminfo.* 2004, Vol. 25., 3.
26. Holmér, Åsa. Translating Snomed CT. *Terminfo.* 2010, Vol. 30, 4.
27. SNOMED International. SNOMED CT Worldwide. *SNOMED International.* [Online] [Cited: May 4, 2018.] https://www.snomed.org/snomed-ct/snomed-ct-worldwide.
28. English in computing. *Wikipedia.* [Online] October 24, 2017. [Cited: April 17, 2018.] https://en.wikipedia.org/wiki/English_in_computing.
29. *Is this code written in English? A study of the natural language of comments and identifiers in practice.* Timo Pawelka, Elmar Juergens. Bremen : 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME), 2015. pp. 401-410.
30. C# Coding Standard. *Philips Healthcare - C# Coding Standard.* [Online] [Cited: April 18, 2018.] http://tics.tiobe.com/viewerCS/index.php?ID=52&CSTD=Rule.
31. GNU Coding Standards. *GNU Operating System.* [Online] [Cited: April 18, 2018.] https://www.gnu.org/prep/standards/standards.html#Comments.
32. Coding standards. *Drupal.* [Online] [Cited: April 18, 2018.] https://www.drupal.org/docs/develop/standards/coding-standards.
33. World language. *Wikipedia.* [Online] [Cited: April 18, 2018.] https://en.wikipedia.org/wiki/World_language.
34. *"Pitkät lomat ja lyhyet työpäivät" - Ulkomaalaiset työntekijät yllättyivät suomalaisesta työelämästä.* Korpimies, Annika. Helsinki : Alma Talent Oy, April 16, 2018, Kauppalehti.
35. Foreign-language speakers. *Statistics Finland.* [Online] [Cited: April 21, 2018.] https://www.stat.fi/tup/maahanmuutto/maahanmuuttajat-vaestossa/vieraskieliset.html.
36. Statistics Finland / Population structure. Foreign-language speakers. *Statistics Finland.* [Online] [Cited: April 21, 2018.] http://vertinet2.stat.fi/verti/graph/Graphserver.aspx?ifile=quicktables/Maahanmuuttajat/kieli_1&lang=1&rind=6&gskey=2&mimetype=png.
37. Iivonen, Sari. MUSLIMIPERHE TERVEYDENHUOLLON ASIAKKAANA- TERVEYSNETTI. [Online] 2013. [Cited: April 21, 2018.] https://www.theseus.fi/bitstream/handle/10024/69886/Iivonen_Sari.pdf?sequence=1&isAllowed=y.

38. Lukkarinen, Nora. Multiculturalism in Health Care: Learning Package for Public Health Nurse Students. [Online] 2013. [Cited: April 21, 2018.] https://www.theseus.fi/bitstream/handle/10024/61021/Lukkarinen_Nora.pdf?sequence=1.
39. Hennariina Kolli, Natalia Kettunen. Over The Language Barrier to Consensus. [Online] 2010. [Cited: April 21, 2018.] https://www.theseus.fi/bitstream/handle/10024/14018/Pori_Kettunen_2010.pdf?sequence=1.
40. Business Finland, Visit Finland. Visit Finland Visitor Survey 2017. *Visit Finland.* [Online] 2018. [Cited: April 22, 2018.] http://www.visitfinland.fi/wp-content/uploads/2018/04/T2018-Visit-Finland-Visitor-Survey-2017.pdf?dl.
41. Koski, Paula. Ruotsalaispotilaiden määrä Suomessa lisääntyy ja nyt halutaan muihinkin kuin syöpähoitoihin - ”Potilaan kohtaaminen parempaa”. *Iltalehti.* April 21, 2018.
42. Mika Gissler, Timo A. Tanninen, Pia Blomqvist et al. *Statistics on Patient Mobility in the Nordic Countries - Pre-Study.* Copenhagen : Nordic Medico-Statistical Committee, 2017.
43. DIRECTIVE 2011/24/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 9 March 2011 on the application of patients' rights in cross-border healthcare. *Official Journal of the European Union.* Strasbourg : s.n., March 9, 2011. Directive 2011/24/EU.
44. Act on Cross-Border Health Care. Helsinki : Ministry of Social Affairs and Health, December 30, 2013. Act 1201/2013.
45. European Commission. *Commission report on the operation of Directive 2011/24/EU on the application of patients' rights in cross-border healthcare.* Brussels : European Commission, 2015. Report From The Commission To The European Parliament And The Council.
46. Health Care Act. *Section 6.* Helsinki : Ministry of Social Affairs and Health, December 30, 2010. Act 1326/2010.
47. Language Act. *Section 5.* Helsinki : Ministry of Justice, June 6, 2003. Act 423/2003.
48. Association of Finnish Local and Regional Authorities. *Ruotsin- ja kaksikieliset kunnat.* 2017. p. 2.
49. Laki terveydenhuollon ammattihenkilöistä annetun lain muuttamisesta. *[Amendment to Health Care Professionals Act].* December 30, 2015. Act 1659/2015.
50. Health Care Professionals Act. Naantali : Ministry of Social Affairs and Health, June 28, 1994. Act 559/1994.
51. Lääkäriliitto. *Lääkärit 2016.* 2016. p. 6.
52. Finnish Medical Association. *Physicians 2015.* 2015. p. 6.
53. Scaled Agile, Inc. Provider of SAFe. *Scaled Agile.* [Online] Scaled Agile, Inc. [Cited: May 23, 2018.] https://www.scaledagile.com/.

54. *SPECIAL LANGUAGE AND GENERAL LANGUAGE: LINGUISTIC AND DIDACTIC ASPECTS.* Varantola, Krista. 2 (23), s.l. : LSP Centre, CBS, Denmark, December 1986, Unesco Alsed-LSP Newsletter, Vol. 9. ISSN 0106-0341.
55. International Organisation for Standardization. Terminology work - Principles and methods. Geneva : International Organisation for Standardization, November 1, 2009. ISO 704:2009.
56. Rouse, Margaret. Definition Ontology. *TechTarget WhatIs.* [Online] TechTarget WhatIs, September 2005. [Cited: May 17, 2018.] https://whatis.techtarget.com/definition/ontology.
57. Natalya F. Noy, Deborah L. McGuinness. Ontology Development 101: A Guide to Creating Your First Ontology. [Online] Stanford University. [Cited: May 17, 2018.] https://perso.liris.cnrs.fr/alain.mille/enseignements/Ecole_Centrale/What%20is%20an%20ontology%20and%20why%20we%20need%20it.htm.
58. *There Is No Knowledge Without Terminology. How Terminological Methods and Tools Can Help to Manage Monolingual and Multilingual Knowledge and Communication.* Sauberer, Gabriele. 2, Vienna : TermNet - International Network for Terminology, 2011, SYSTEMICS, CYBERNETICS AND INFORMATICS, Vol. 9, pp. 56-60. ISSN: 1690-4524.
59. *Seija Suonuutille SESKOn tunnustuspalkinto 2017.* Hieta-Wilkman, Sinikka. 3, 2017, Terminfo.
60. University of Tampere. OPETUSOHJELMA 2018–2019: Viestintätieteiden tiedekunta - Teknisen viestinnän erikoistumisopinnot. *University of Tampere.* [Online] University of Tampere. [Cited: May 20, 2018.] https://www10.uta.fi/opas/opetusohjelma/opetusohjelma.htm?lvv=2018&ops=164&ots=278&kieli=fi&display_long=true.
61. *Sanastotyön alkutaipale: perusteet hyvälle sanastotyön aloitukselle.* Suonuuti, Seija. 1, s.l. : Terminfo, 2013, Terminfo.
62. *Summaries.* Suhonen, Sirpa. 1, s.l. : Terminfo, 2013, Terminfo.
63. *Sanastotyön asema yrityksen prosesseissa.* Suonuuti, Seija. 4, Helsinki : Tekniikan Sanastokeskus ry , 1998, Terminfo, Vol. 18, pp. 12-15. ISSN 0358-751 7.
64. *Terminology workflows in theory and practice.* Chiocchetti, Elena, Ralli, Natascia and Wissik, Tanja. [ed.] G. Budin and V. Lušicky. Vienna : University of Vienna, 2014. Languages for Special Purposes in a Multilingual, Transcultural World, Proceedings of the 19th European Symposium on Languages for Special Purposes, 8-10 July 2013. pp. 525-535. ISBN: 978-3-200-03674-1.
65. K. Beck, M. Beedle, A. van Bennekum et al. Manifesto for Agile Software Development. [Online] 2001. [Cited: May 5, 2018.] http://agilemanifesto.org/.

66. Mary Poppendieck, Tom Poppendieck. *Lean Software Development: An Agile Toolkit.* s.l. : Addison-Wesley Professional, 2003. ISBN 978-0-321-15078-3.
67. Scaled Agile, Inc. What Is SAFe? *Scaled Agile.* [Online] Scaled Agile, Inc., October 7, 2017. [Cited: May 4, 2018.] https://www.scaledagileframework.com/what-is-safe/.
68. —. SAFe 4.5 Introduction, Overview of the Scaled Agile Framework for Lean Enterprises. [Online] 2017. [Cited: May 5, 2018.] http://editor.scaledagile.com/wp-content/uploads/delightful-downloads/2018/01/White_Paper_SAFe-4.5.pdf.
69. Järvinen, Pertti. *On Research Methods.* Tampere : Opinpajan kirja, 2012.
70. *Analysis and optimization of software requirements prioritization techniques.* Muhammad Aasem, Muhammad Ramzan, Arfan Jaffar. Karachi : IEEE, June 2010, Information and Emerging Technologies (ICIET), 2010 International Conference on.
71. McIntyre, John. MoSCoW or Kano Models - how do you prioritize? *HotPMO!* [Online] October 20, 2016. [Cited: May 3, 2018.] https://www.hotpmo.com/management-models/moscow-kano-prioritize.
72. National Institute of Health and Welfare. Code Service. *THL.* [Online] National Institute of Health and Welfare, October 26, 2016. [Cited: May 26, 2018.] https://thl.fi/en/web/information-management-in-social-welfare-and-health-care/standardisation-of-data-and-requirements/code-service.
73. Kustannus Oy Duodecim. Duodecimin sanakirjat. *Duodecim Terveysportti.* [Online] Kustannus Oy Duodecim, 2017. [Cited: May 26, 2018.] http://www.terveysportti.fi/sovellukset/sanakirjat/#/.
74. National Library of Finland, Ministry of Finance, Ministry of Education and Culture. About. *Finto Finnish Thesaurus and Ontology Service.* [Online] National Library of Finland, Ministry of Finance, Ministry of Education and Culture. [Cited: May 26, 2018.] http://finto.fi/en/about.
75. Finnish Foundation of Nursing Education. Hoidokki. *Hoidokki.* [Online] Finnish Foundation of Nursing Education, 2010. [Cited: May 25, 2018.] http://www.hoidokki.fi/index_eng.html.
76. Institute for the Languages of Finland. Welcome to Kielitoimiston sanakirja. *Kielitoimiston sanakirja.* [Online] Institute for the Languages of Finland, February 28, 2017. [Cited: May 26, 2018.] https://www.kielitoimistonsanakirja.fi/.
77. Kielikone. Kielikone Ltd. *Motstore.* [Online] Kielikone, 2017. [Cited: May 25, 2018.] https://www.motstore.fi/kielikone-eng.
78. Oxford University Press. Our story. *Oxford Dictionaries.* [Online] Oxford University Press, 2018. [Cited: May 26, 2018.] https://www.oxforddictionaries.com/our-story.
79. Finnish Terminology Centre TSK. Welcome to TEPA Term Bank. *TEPA Term Bank.* [Online] Finnish Terminology Centre TSK. [Cited: May 25, 2018.] http://www.tsk.fi/tepa/en/.

80. Prime Minister's Office. About Valter. *VALTER Government Termbank.* [Online] Kielikone Ltd., 2018. [Cited: May 25, 2018.] https://mot.kielikone.fi/mot/valter/netmot.exe.

Requirements \ Solution Names	Code Service	Duodecim Dictionaries	Finto	Hoidokki	Kielitoimiston sanakirja	MOT language services	Oxford Dictionaries	TEPA Term Bank	Valter
1. The solution must enable controlling user access.	-	x	-	-	-	x	-	-	-
2. The solution must include global terms.	-	x	x	-	-	x	x	x	-
3. The solution must include national terms.	x	x	x	x	x	x	-	x	x
4. The solution must enable including in-house terms.	-	-	-	-	-	x	-	-	-
5. The solution must include source code terms.	-	-	(x)	-	(x)	x	(x)	(x)	-
6. The solution must include product user interface terms.	x	x	x	x	x	x	-	x	-
7. The solution must include special language of medicine.	x	x	x	x	(x)	x	(x)	(x)	-
8. The solution must include special language of social welfare.	x	(x)	(x)	-	(x)	x	(x)	x	-
9. The solution must include special language of client administration.	x	(x)	(x)	(x)	(x)	(x)	(x)	(x)	-
10. The solution must include special language of software product development process.	-	-	-	-	(x)	x	(x)	x	-
11. The solution must include special language of programming.	-	-	-	-	(x)	(x)	(x)	x	-
12. The solution must include general language.	(x)	(x)	(x)	-	x	x	x	(x)	-

Requirements \ Solution Names	Code Service	Duodecim Dictionaries	Finto	Hoidokki	Kielitoimiston sanakirja	MOT language services	Oxford Dictionaries	TEPA Term Bank	Valter
13. The solution must include Finnish language.	x	x	x	x	x	x	-	x	x
14. The solution must include English language.	-	x	(x)	x	-	x	x	x	x
15. The solution must include descriptions for terms.	x	(x)	(x)	x	x	(x)	x	(x)	(x)
16. The solution must include descriptive terminology entries.	(x)	x	x	x	x	x	x	x	x
17. The solution must include prescriptive terminology entries.	x	(x)	(x)	(x)	x	(x)	(x)	x	x
18. The solution must include appellations.	x	x	x	-	(x)	x	(x)	x	x
19. The solution must include entry domain and subdomain information.	x	(x)	x	x	x	x	x	x	x
20. The solution must include entry source information.	x	-	x	-	-	(x)	(x)	x	x
21. The solution must include entry date information.	x	-	(x)	x	-	-	-	(x)	(x)
22. The solution must include entry responsibility information.	x	(x)	(x)	-	(x)	(x)	(x)	(x)	(x)
23. The solution must enable process development.	x	x	x	x	x	x	x	x	x
24. The solution must be simple to use.	-	x	x	x	x	x	x	x	x
25. The solution must enable maintenance and extending.	-	-	-	-	-	x	-	-	-

Requirements \ Solution Names	Code Service	Duodecim Dictionaries	Finto	Hoidokki	Kielitoimiston sanakirja	MOT language services	Oxford Dictionaries	TEPA Term Bank	Valter
26. The solution must be usable immediately.	x	x	x	x	x	-	x	x	x
27. The solution could include symbols.	-	-	(x)	-	-	(x)	-	-	x
28. The solution could include concept systems and diagrams.	(x)	-	(x)	(x)	-	(x)	-	-	x
29. The solution should not introduce new software.	x	x	x	x	x	-	x	x	x
30. The solution should not be machine readable i.e. an ontology.	-	x	(x)	x	x	x	x	x	x
Total	16 x, 3 (x)	14 x, 7 (x)	13 x, 12 (x)	14 x, 3 (x)	13 x, 8 (x)	19 x, 8 (x)	11 x, 10 (x)	18 x, 7 (x)	15 x, 3 (x)

EXAMPLE ENTRIES IN THE CASE COMPANY GLOSSARY

APPENDIX 2

Finnish Term	Finnish Synonyms & Definition & Comments	English Term	English Synonyms & Definition & Comments	Reliabilit…	Term Group	Context	Change Histor…	Source
aikataulu		schedule		•	General	General	marti… 20180419	
CDA	tulee sanoista Clinical Document Architecture. CDA on XML-muotoinen merkintästandardi, jota käytetään potilasasiakirjojen välityksessä. Standardi määrittelee enkoodauksen, rakenteen ja semantiikan. HL7-organisaatio julkaisi version 1.0 (CDA R1) lokakuussa 2000 ja version 2.0 (CDA	CDA	The HL7 Clinical Document Architecture (CDA) is an XML-based markup standard intended to specify the encoding, structure and semantics of clinical documents for exchange. In November 2000, HL7 published Release 1.0. The organization published Release 2.0 with its "2005 Normative Edition."	•••	Healthcare	General	marti… 20180419	https://en.wikipedia.org/wiki/Clinical_Document_Architecture
Duodecim	Suomalainen Lääkäriseura Duodecim on vuonna 1881 perustettu tieteellinen yhdistys. Se kehittää lääkärin ammattitaitoa ja käytännön työtä täydennyskoulutuksen, julkaisujen ja apurahojen avulla. Seuraan kuuluu yli 21 000 lääkäriä ja lääketieteen opiskelijaa sekä noin	Duodecim	The Finnish Medical Society Duodecim is a scientific association established in 1881 to develop the professional skills and clinical practice of doctors through continuing education, publications and grants. The Society's membership comprises more than 22,000 doctors and medical students and almost 100 member associations.	•••••	Organization Name	General	marti… 20180419	https://www.duodecim.fi/duodecim/ ; https://www.duodecim.fi/english/duodecim/
ERAS	tulee sanoista erillinen asiakirja; Näkymä, jota käytetään niiden potilaan hoidon kannalta tarpeellisten tietojen kirjaamisessa, jotka liittyvät toisen henkilön itsestään tai omasta elämäntilanteestaan kertomiin, yksityiskohtaisiin arkaluonteisiin asioihin Huom. Toisen henkilön tiedot kirjataan potilaan elämäntilanteen kartoittamisen tai muun vastaavan syyn vuoksi. Merkinnöistä tulee ilmetä, ketä yksilöityä henkilöä tiedot koskevat (vrt. Potilasasiakirja-asetus 298/2009, 7 § ja 2. momentti)	ERAS	Use the acronym. It comes from Finnish words for Separate Document. A view that is used to record the essential information for treatment of a patient that is told by another person and is about this other person. The recorded information is detailed and sensitive information about the health or other life situation of that other person that needs to be recorded to evaluate the patients life situation or for some other similar reason. The created records must identify the other person.	••••	Healthcare	General	martin… 20180419	http://koodistopalvelu.kanta.fi/codeserver/pages/classification-view-page.xhtml?classificationKey=226&versionKey=301
toimipiste	Terveydenhuollon toimintayksikön osa, jolla on tietyt palvelutehtävät ja näiden tehtävien hoitoon koulutettu ammattihenkilökunta. Esimerkkejä terveydenhuollon toimipisteistä ovat terveysasemat ja sairaalan tai terveyskeskuksen vuodeosastot ja poliklinikat. Erityisesti terveydenhuollon tietojärjestelmissä sanaa toimipiste käytetään kuvaamaan organisaatiohierarkian alinta tasoa.	outpatient or inpatient department	point-of-service; local unit;	••••	Healthcare	General	marti… 20180419	http://www.tsk.fi/tepa/fi/haku/toimipiste

IMPLEMENTATION COMPLIANCY ANALYSIS AGAINST REQUIREMENTS APPENDIX 3/1

Requirements		Compliancy Evaluation
	End-user Requirements	
1.	Must be accessible only within health and social welfare department.	Like stated in chapter 6.2.2: "…the document management system got selected because everyone in the department had to use it occasionally or daily and everyone must have access rights to it." The implementation is compliant.
2.	Must include global, national and company terms with descriptions.	The several types of context and term group options enable fulfilling this requirement, and the baseline glossary already included all these mentioned types of terms. The implementation is compliant.
3.	Must include terms used in source code and product user interfaces with descriptions.	The term group options enable including terms also in these groups. The implementation is compliant.
4.	Must support two languages but enable also single language usage with definitions.	The glossary is bilingual, but the process enables and allows also monolingual usage. The implementation is compliant.
5.	Must be simple to use.	This requirement was fulfilled according to demo event feedback. The implementation is compliant.
6.	The solution should not introduce new software.	The implementation is compliant.
7.	Must use existing user access controls.	The document management system takes care of the user access control. The implementation is compliant.
8.	Maintenance must be simple.	This requirement was fulfilled according to demo event feedback. The implementation is compliant.

IMPLEMENTATION COMPLIANCY ANALYSIS AGAINST REQUIREMENTS APPENDIX 3/2

Requirements		Compliancy Evaluation
	End-user Requirements	
9.	Maintenance must not depend on one person.	Everyone in the department can and should do maintenance of the glossary when using it and the Glossary CoP is responsible of it as a group. The implementation is compliant.
	Tool Requirements	
10.	The solution must enable entries of several special languages.	The term group options enable including terms also in these groups. The implementation is compliant.
11.	The solution must enable entries of general language.	The term group and context options enable including also these terms. The implementation is compliant.
12.	The solution must enable both descriptive and prescriptive terminology entries.	The structure of the glossary supports both uses but this information is not forced in the structure. It can be evaluated later if a structure is needed for this. The implementation is compliant.
13.	The solution must support terms and appellations.	The implementation is compliant.
14.	The solution must support minimum entry information: Term, definition or example textual context, domain and subdomain, source, date, responsibility.	The implementation is compliant.
15.	The solution could support concept systems and diagrams or in minimum not prevent extending to this direction later.	Spreadsheet applications support drawing diagrams, so they could be attached to the glossary itself or just links to them could be included. The implementation is compliant.
16.	The solution could support symbols.	Inserting symbols is supported by the solution. The implementation is compliant.

IMPLEMENTATION COMPLIANCY ANALYSIS AGAINST REQUIREMENTS APPENDIX 3/3

Requirements	Compliancy Evaluation
Tool Requirements	
17. The solution should not be machine readable i.e. an ontology.	The solution is meant for human use only. The implementation is compliant.
Process Requirements	
18. The process must define the terminology workflow, for example: needs analysis, source documentation, term extraction and selection, term entry elaboration, quality check, dissemination, maintenance.	The workflow is described by the processes, but the developed agile terminology management is not focused so much on the beginning part of the example workflow because the work happens within the core processes of the case company. The implementation is compliant.
19. The process should ensure limited number of synonyms for / in entries.	The structure of the glossary or the glossary process does not limit the number of synonyms. This quality aspect is left for Glossary CoP to follow up with term recommendations and maintenance. The implementation is not compliant.
Management Requirements	
20. Terminology work must have defined objectives.	The implementation is compliant.
21. Terminology work benefits should be maximized by having only one source for terminology.	People were requested to integrate their own personal word lists into the glossary and use it in future. Also, there were no more glossaries in the company wiki in the end of the implementation phase. The various solutions in the product development management system were under further analysis and study by the Glossary CoP at the end of the implementation phase before integrating them. The implementation is partially compliant.

IMPLEMENTATION COMPLIANCY ANALYSIS AGAINST REQUIREMENTS APPENDIX 3/4

Requirements	Compliancy Evaluation
Management Requirements	
22. Terminology work benefits should be maximized by as wide as possible usage.	The glossary is meant for everyone's use in the department and the process supports it. By allowing people to add their own terminology in the glossary, they will consider the common glossary their own. The project and Glossary CoP in future make terminology work and the glossary more known in the case company, which will increase the usage. After this brief time using the common glossary and the new processes related to it, the implementation is partially compliant.
23. Terminology work benefits should be maximized with a thorough preliminary study before big investments.	This project provides for a basis of the study and the Glossary CoP continues following up on the terminology work needs. In addition, no big investments were made during the project. The implementation is compliant.
24. Terminology work and process should be developed and evaluated continuously.	The implementation is compliant.
25. The core processes supported by the terminology work should be identified.	The implementation is compliant.
26. Terminology work should be fitted and scheduled to the supported core processes in the company.	The agile terminology management process created in this project supports the core processes in the case company well. The Glossary CoP keeps following up that the terminology processes stay supportive to the core processes. The implementation is compliant.

IMPLEMENTATION COMPLIANCY ANALYSIS AGAINST REQUIREMENTS APPENDIX 3/5

Requirements	Compliancy Evaluation
Management Requirements	
27. Terminology work related roles should be in minimum roughly defined.	Various roles that should participate in the Glossary CoP were defined in the project. The implementation is compliant.