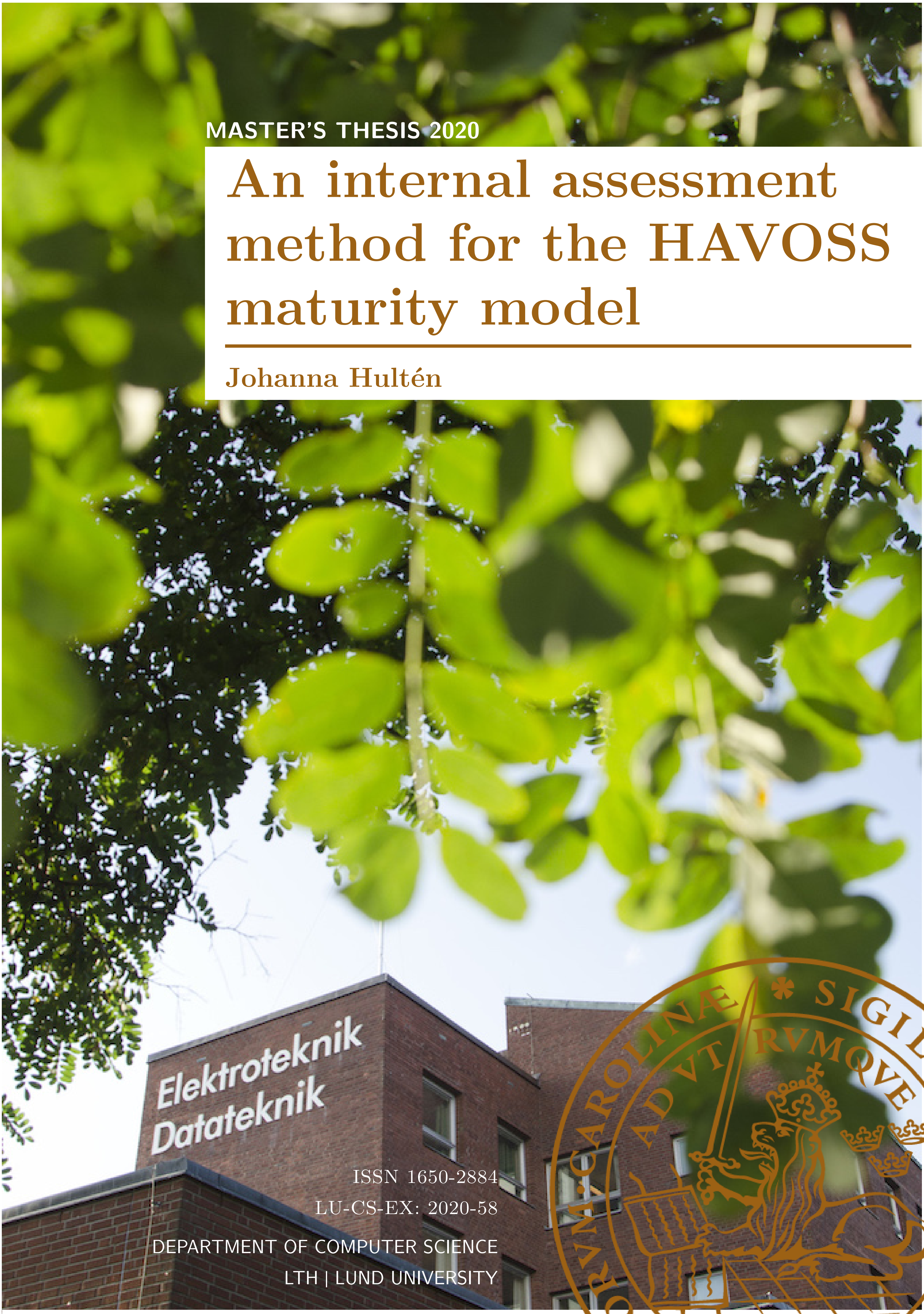
MASTER'S THESIS 2020

An internal assessment method for the HAVOSS maturity model

Johanna Hultén



ISSN 1650-2884
LU-CS-EX: 2020-58
DEPARTMENT OF COMPUTER SCIENCE
LTH | LUND UNIVERSITY

EXAMENSARBETE
Datavetenskap

LU-CS-EX: 2020-58

An internal assessment method for the HAVOSS maturity model

En intern utvärderingsmetod för HAVOSS mognadsmodell

Johanna Hultén

An internal assessment method for the HAVOSS maturity model

(A LaTeX class)

Johanna Hultén
min.mail@johannap.se

September 22, 2020

Master's thesis work carried out at
the Department of Computer Science, Lund University.

Supervisor: Martin Höst, martin.host@cs.lth.se

Examiner: Emelie Engström, emelie.engstrom@cs.lth.se

Abstract

The Handling Vulnerabilities in third party OSS (HAVOSS) maturity model does not have an accompanying assessment methodology. HAVOSS is a model developed to assess how companies handle security with OSS in IoT software. In this thesis we developed an internal assessment methodology for the model. Based on a literature study the resulting process assessment methodology is a four-phase lightweight assessment methodology. The four phases for the process assessment are a preparation phase, followed by a digital questionnaire completed by a larger number of participants, called the individual assessment. This is then followed by a workshop that is phase three where the questionnaire is completed by the workshop with the digital tool as a guide. The final phase is generating the report from the assessment. We also developed a digital tool to aid the assessment process and visualize the results. The methodology and digital tool were evaluated with staff at Lund University and two industry representatives.

Keywords: assessment methodology, digital tool, HAVOSS, internal assessment, IoT, maturity model, process assessment, self-assessment, software process assessment, software process improvement, SPI, software security

Acknowledgements

I want to start with thanking Martin Höst for his work as supervisor, supporting the work and giving valuable insights into the subject and the process of writing this master thesis in general.

Thank you to the people who gave of their valuable time to evaluate the assessment method developed in this thesis. Without their time this would not be what it is.

A big thank you to my husband Martin for supporting my work, my children for their inspiration to continue this thesis when life got crazy. Lastly, thank you to my parents and in-laws for helping out with the children so I could get away and complete this work.

Contents

Chapter 1
Introduction

No matter what industry an organization is in, the need to continuously improve exists if they want to survive. One way for software organizations to improve is to improve the processes used in the organization to produce the software, so called software process improvement (SPI). The goal of SPI is to reduce costs and increase quality of the software delivered from the organization by improving processes in the organization. To do any process improvement, the first step is to assess the organizations current processes, called a processes assessment. It gives insight to how the processes in the organization currently work.

To help assess how a company is preforming their processes, a maturity model can be used. A maturity model can be seen as a tool, defining important areas for whatever it assesses. It also ensures that no important area is missed by an organization. Using a maturity model when doing an assessment usually results in a so-called maturity level. A maturity level gives an indication of where the organization is at, based on best practices and the experience of the maturity model's authors. The maturity level also gives the organization important information on what to improve and how to prioritize those improvements to reach a higher level in the next assessment.

One such maturity model is the Handling Vulnerabilities in third party OSS (HAVOSS) model [?] developed at Lund University. The HAVOSS model is a maturity model for handling vulnerabilities in open source software (OSS) and commercial-of-the-shelf (COTS) software used in the organizations own software. The primary focus when developing the model was software for Internet of Things (IoT), but the model itself is not specific for IoT.

In this master thesis we developed a lightweight internal process assessment method to determine a maturity level based on the HAVOSS model. To go along with the assessment method, a digital tool was developed to help in the assessment and making it less work intensive.

1.1 Motivation

The HAVOSS maturity model [?] was developed to help organizations evaluate their security around their use of OSS and COTS. However, without an assessment process to accompany the model, it is hard for organizations, especially small and medium size organizations, to use the model. This is because without special knowledge of software improvement processes and process assessment in particular, the model is very hard to comprehend and use. The model could be used with one of the big assessment frameworks like the standard CMMI appraisal method for process improvement (SCAMPI) presented in Section ??, but that is labor intensive and comes with high costs.

By constructing an internal assessment methodology for the HAVOSS model that is lightweight, with as little administrative work as possible, the model becomes more accessible and the potential for practical use increases. Having the accompanying digital tool removes a lot of manual work in the form of collecting and analyzing data from paper questionnaires and generating reports. This will create an assessment process that requires as little labor as possible, resulting in a process appropriate for small and medium size organizations.

The gain for any organization using the assessment process and getting a maturity level of the HAVOSS model would be increased awareness of how they deal with their security as it pertains to OSS and COTS. It would also provide the organization with a guide to what to change and work on in order to gain a higher level in the HAVOSS maturity level and thus increasing their security management. In addition, it could be used by potential customers to indicate what organization works best with the type of problems relating to the HAVOSS model.

1.2 Objectives

The aim of this thesis was to develop and validate an internal assessment process for the HAVOSS maturity model. To accompany the assessment process a digital tool was developed as a prototype to aid in the assessment. The aim for the digital tool was to limit the manual work needed to complete the assessment for an organization.

The following research question was asked for this thesis:

RQ: How can the HAVOSS model be implemented into an internal assessment processes for an organization?

1.3 Limitations

In the process of doing this thesis some limitations had to be set in order to contain the workload to the planned time limit for the thesis. These limitations are presented in this section.

The digital tool was only developed as prototype to show and evaluate the potential of such a tool. As such it was not a production ready tool with all the needed functionality to be used out in organizations when this thesis was finished. Instead, it contained just enough functionality to verify the usefulness of such a tool and get an indication of how it can help in the assessment process.

The work of this thesis was not done in conjunction with any company but at the university. Therefore, the understanding of the problem, the demands of the process and the testing of the actual process and the digital tool was limited to be more academically geared. Evaluations were done by three employees by the university and two industry representatives. Two of the three university employees had good knowledge beforehand of the HAVOSS model where one of them helped developed it, the third university employee had good insight into the subject of evaluations. The two industry representatives work in software development but neither work with IoT. All participants in the evaluation were able to give a valuable opinion and evaluation of the work. But it is no replacement for testing at organizations and finding what actual work with the model, process and digital tool and what does not.

1.4 Report outline

In Chapter **??** background information about the key subjects of the thesis are given along with the theories that the work of the thesis is built on. This is followed by a presentation of the research methodology, along with the vulnerabilities of the research in Chapter **??**.

Chapter **??** presents the developed assessment methodology and the digital tool developed to help with the assessment. The chapter also contains the results of the evaluation of the methodology and digital tool. Chapter **??** covers the analysis and discussion about the data presented in Chapter **??** and what results can be found in the collected data. The last chapter, Chapter **??**, contains conclusions found in the thesis along with possible further work that was identified to be needed or that could bring additional interesting results connected to this thesis.

In the appendixes, a collection of best practices for internal assessment processes presented along with the interview questions used in the evaluations.

Chapter 2
Background

The need to improve the processes used in an organization exists for every organization that wants to continue to get better. Quality management is a general way for organizations to improve. For organizations dealing with software, software process improvement is a more specific way of doing quality management focused on the processes of developing software to increase quality.

2.1 Quality Management

In the beginning, quality management was only concerned with the quality of the resulting product. This has changed over time with the understanding that the quality of processes gives the quality of the resulting product. The most important feature in quality management today is that it directs attention to the improvement of production processes and not simply the characteristics of the products [?].

The current quality management philosophies have been shaped by William Edward Deming [?, ?] with his 14 principles of quality, Joseph M. Juran [?, ?] with his Juran trilogy and Philip Crosby [?, ?, ?] with his 14 quality steps. All three have multiple frameworks built upon their theories and principles. These authors and others all discuss the importance of critical factors such as leadership or management involvement, employee participation, measurement and process management to improve the quality of an organizations processes. Because of this, these factors have been defined as success factors, no matter what type of organization that the quality management is applied on. With all these frameworks and models for quality management, no single model have been established as the base model for quality management control and the theory thereof [?].

As a basis for many quality management processes is the today well known Plan-Do-Check-Act (PDCA) cycle that Deming refined from the Shewhart cycle [?, p. 132]. An illustration of this cycle is shown in Figure ??. The PDCA is an iterative

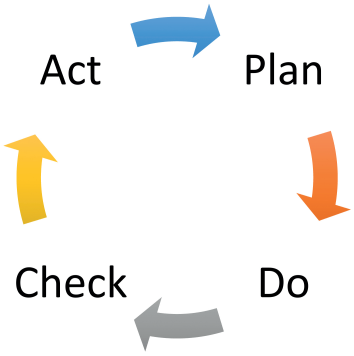


Figure 2.1: The PDCA cycle as Deming described it

four-step process for continuous improvement in organizations. The cycle starts with a planning step where the objectives of the process is established along with what processes are required to deliver the results that are desired. This is followed by "do" where the plan created in the previous step is carried out, usually in a small scale first to test and then in the entire organization. When this is done, the cycle goes into the check step where the data and results from the previous phase is evaluated to see if the changes made produced the desired results and if changes to the processes are needed for better results. This carries over into the act phase, also called the adjust phase, where the found results are used to identify issues and problems with the tested process and improvements. These are investigated to see what the root-cause is. The information found in this last step is then used in the next cycle of improvements.

2.2 Software Process Improvement

Software process improvement (SPI) is a quality management process of improving the time, cost and quality of the engineering and management processes and practices in software organizations and is not a new concept. The guided SPI methodologies were spearheaded at the Software Engineering Institute (SEI) [?] as a job for the US Department of Defense (DoD). The result was the Capability Maturity Model (CMM) which was first presented to the world in the paper "Characterizing the Software Process: A Maturity Framework" written by Watts S. Humphrey [?] in 1988. Since then the area of SPI has just grown, with a few dominating frameworks taking center stage.

2.2.1 The SPI cycle

Already in his first paper [?] Humphrey defined the five steps an organization must take to improve an organization's software capabilities. The same basic steps are still used today for software process improvement. The five steps are;

1. to understand the current status of the development processes,

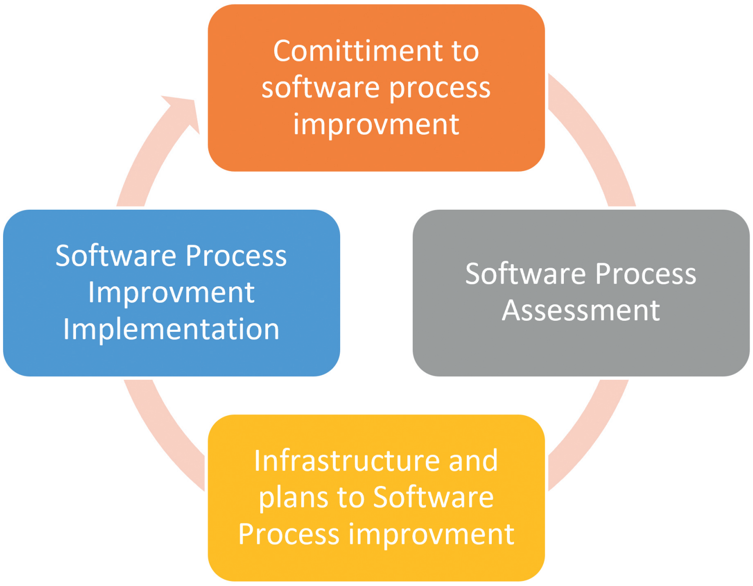

Comittiment to software process improvment
Software Process Assessment
Infrastructure and plans to Software Process improvment
Software Process Improvment Implementation

Figure 2.2: A generic software process improvement cycle, inspired from [?]

2. develop a vision of the desired process,
3. establish a list of required process improvement actions in order of priority,
4. produce a plan to accomplish these actions and
5. commit the resources to execute the plan [?].

The steps Humphrey defined can be generalized into a generic model for more or less all SPI work [?], illustrated in Figure ??. It is a four-step improvement process that can then go in cycles for continuous improvement. The cycle is as follows;

The first step is to create a commitment in management for improvement, this maps to Humphrey's fifth step of getting the resources needed to execute the plan.

The second step is a software process assessment where the objective is to understand the current status of the development processes. The step also aims find the strengths and weaknesses of the processes assessed as well as selecting what to improve. This maps to Humphrey's first step.

The third step is to create an action plan on how to improve the selected areas from the previous step. In the same step, the infrastructure needed to carry out the planned improvements are setup. This maps to the second, third and fourth step of Humphrey's plan.

The fourth step is the implementation of the planned improvements. This is sometimes first done in a pilot project followed by the rest of the organization instead of the entire organization directly. This step is not specified in Humphrey's plan, but is obviously required for continuous improvement and using his steps.

This generic model of SPI also ties back to the previously presented PDCA cycle as it maps directly onto it, except that the models have different starting points. The first step in the SPI process maps to the check step in the PDCA cycle, and then going on from there. And as such, SPI is a quality management process as any other, but with a more specific goal in mind for each step than for example the PDCA cycle.

2.2.2 Classifying SPI frameworks

There is a large number of frameworks to support SPI, all of which can be rudimentary classified into two basic categories, inductive frameworks and prescriptive frameworks [?]. This is a way to classify the different frameworks based on the basic work process used during the improvement process. Inductive frameworks use a bottom-up approach, while prescriptive frameworks use a top-down method.

Prescriptive models take a set of best practices that have provided success in other organizations and bases the improvement work on that. The set of best practices are then compared to how the organization work today and what improvements needed are decided. It is important to note that no consideration to the unique needs of the organization are taken in prescriptive models. Thus, these models have a one-size-fits-all policy that might force organizations into practices that are not strictly necessary. One of the most popular forms of doing a software process improvement is using a well-defined model, where there are two that are most often used, CMMI and ISO 15504 both of which are prescriptive.

Inductive frameworks work by finding improvements within the organization based on the situation the organization is in and not on predefined frameworks or practices. This can help create commitment from participants and management both for the assessment and subsequent improvements. As will be presented in Section **??**, this is a critical aspect for succeeding in the process improvement efforts. The main downside of these types of frameworks are that results are only produced if the organizations processes exhibits significant maturity [?].

When the issues that need improvement have been identified, the next step for the organization is to find appropriate ways to create these improvements. Prescriptive framework specifies what to improve and in what order to implement the improvements. In inductive frameworks the organization needs to come up with solutions and prioritize the improvements as best they can, allowing for solving the most critical issues first. This often makes inductive frameworks harder to follow if there is not specialized expertise in inductive software process improvement in the organization.

2.2.3 External and internal assessments

No SPI can be completed without a lot of work being done inside the organization. However, for most of the large SPI frameworks presented below, experts in the framework are needed to complete the software process assessment and sometimes the entire SPI process. These experts are often not found inside the organization but, in most cases, as experts coming in to do the assessment. This is then called an external assessment since someone from outside the organization is doing it. An external assessment can also be performed with expertise found inside the organization given that they do not have any connection to the part of the organization that is being assessed.

An internal assessment, also known as self-assessment, is often considered to be less rigorous than an external assessment. This is because it is done by the assessed organization itself and can thus not be objectively completed. With self-assessment the organization goes through the steps of the assessment without the help from outside experts to get the results of SPI. The big gain by doing the assessment as an internal assessment are the lower costs and the possibility to focus on a single area, not completing an entire framework for SPI. For small and medium size organizations this can be the only option as external complete assessments often gets very expensive.

2.3 Established SPI frameworks and their assessment processes

There exists a large number of different frameworks and maturity models for software process assessment and SPI in general. A few well known are presented in this section.

2.3.1 The Capability Maturity Model

As presented in Section **??**, the CMM framework was first presented in 1988 as the first organized way of doing SPI. The goals for CMM was to reduce the cost for software development for the US Department of Defense as well as improve software quality and maintainability. This was to be done by improving existing software development processes, and the processes related to it. [**?**]

The idea of maturity in the CMM relates to the degree to which an organization has reached formality and optimization of their processes. The CMM introduced five maturity levels, illustrated in Figure **??**. In [**?**] each maturity level have definitions of the characteristics the organization needs to have to achieve to achieve that specific maturity level.

In 1993 CMM Version 1.1, called Capability Maturity Model for Software (SW-CMM), was published [**?**] with a number of improvements made to the model based on experiences gathered since the first version was published. In the years since the introduction of the CMM and later the updated version, many more maturity models have been created based on the either of the versions of CMM, with new models still being created to this day.

An important note is that even though the models were published, how to use them was not included in either of the articles. In [**?**] there is a short note that it was used by SEI by using a questionnaire or in-depth technical reviews to gather the information needed to decide the current maturity level of the organization as the software process assessment, the second step in the SPI cycle. The result of the current maturity level was then along with the information about the current processes compared with what is needed to achieve the next maturity level in the model. This was used as a base to create a plan for what improvements to make,

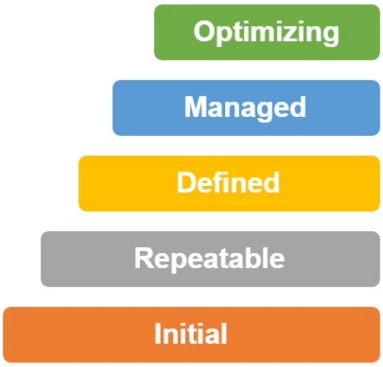


Figure 2.3: CMM maturity levels

the third step in the SPI cycle. The results of this is then implemented, completing the SPI cycle.

There is nothing specifically relating to the software development process in the maturity levels defined in [?]. The model has been used in just about every field since it was presented as a way to improve quality management, giving birth to more specific forms of CMM that was defined in later years for specific fields, including software development.

2.3.2 The capability maturity model integration

After SW-CMM was introduced and multiple new maturity models were developed based on it, there was a need to combine SW-CMM with two new ones based on SW-CMM, the result was the capability maturity model integration (CMMI). CMMI is an integration and evolution of the capability maturity model for software (SW-CMM), integrated product development capability maturity model (IPD-CMM) V0.98A and systems engineering capability model (SECM). The aim of creating the CMMI was to eliminate the need to use multiple models for improvement and evaluation.

Between 2010 when CMMI version 1.3 was published and 2018 there were no updates to the CMMI model and process. In 2018 version 2.0 of the CMMI was published, with major changes. However, since this is still so new when this thesis is written, all references to CMMI in this thesis pertains to version 1.3 if nothing else is specified. The model exists in in two versions, a staged and a continuous representation. They both have the same 22 key process areas, but these are represented differently and have two different approaches to SPI. The key process areas are defined areas of an organization that are being assessed, for example, risk management, supplier agreement management and project planning. Each assessed key process area is given a maturity level and then these are the basis for the entire organization's maturity level.

The staged representation has its aim to assessing an entire organizations maturity. The organization is evaluated against the five different maturity levels and practices are implemented to attain an overall increase in organizational maturity. Continuous representation was created for assessing individual process areas, for example requirements engineering, and improving the related practices for that area.

Even though the continuous representation allows for individual areas to be targeted, it still guides priorities and stating what practices should be improved or added and in what order. As such, it is still prescriptive in nature [?].

Any assessment methodology that is part of CMMI or using the CMMI model must fulfill the appraisal requirements for CMMI (ARC), a set of requirements put forth for CMMI for how any assessment methodology should be performed. These are defined to create a better assessment process and results. The ARCs can be used with any reference model methodology with good results. [?]

The Standard CMMI Appraisal Method for Process Improvement

The standard CMMI appraisal method for process improvement (SCAMPI) is designed to provide a well-defined set of methodologies for conducting process assessments relative to CMMI models but can also be used with other maturity models. There are three assessment versions of SCAMPI, called class A, B and C. Class A is the strictest version of the three, looking at all the ARCs. It is class A that produces a benchmark rating that allows the comparison of maturity or capability levels across organizations. Class B and C only look at a subset of the ARCs, and is less strict but also more cost effective while still giving good insight into the organization. [?]

SCAMPI contains four phases, with several essential processes belonging to each phase. The first phase is planning and preparing for the appraisal and is crucial for the success of the assessment and any improvement work based on the assessment. Requirements for the appraisal are established and then used to create a plan. An appraisal team is selected and prepared, and any information that is specific to the organization that needs to be considered is collected.

When the planning phase is completed, the assessment goes into the active phase of conducting the appraisal. SCAMPI depends on a collection of information that is collected via defined types of objective evidence. The appraisal team observes, hears or reads information that is then transformed into notes and later into model gaps and lastly into findings. The organization usually then validate the findings for it becomes formal findings. This data is later the base of the appraisal results. The third phase is to prepare a report with the findings of the appraisal, this is the followed by the fourth phase of action plan reappraisal.

2.3.3 ISO/IEC 15504

The ISO/IEC 15504 standard with the title *Information technology – Process assessment*, also goes under another, more commonly known name in research and industry, SPICE, standing for Software Process Improvement and Capability Evaluation. ISO/IEC 15504 is a set of technical standards documents for software process assessment. It is not in itself a reference model but a set of requirements on maturity models to be used in process assessment. Important to note is that it is not a single document, but a collection of documents published by ISO/IEC.

The standard was initially derived from the ISO/IEC standard for software life cycle processes called ISO/IEC/IEEE 12207 [?] and influenced by different maturity models, one of them being CMM. ISO/IEC 15504 still relies on other defined standards such as ISO/IEC 12207 [?] and ISO 15288 [?], both of which are a process lifecycle standards. In 2015 the standard was revised by ISO/IEC 33000 and because of this ISO/IEC 12207 is no longer part of ISO [?]. None the less it is still very much in use today both in practices and research.

The assessment done by SPICE is similar to SCAMPI as they both have similar requirements, where SCAMP's are defined as ARCs. A central difference between the two is that while CMMI can use both internal or external assessment group members, SPICE requires that an external assessor heads the assessment [?] and as such is only an external assessment.

2.3.4 Lightweight SPI frameworks

CMMI and SPICE are the golden standard of SPI, but there is a large need for smaller frameworks that are easier to implement. This is because in many cases it requires a big commitment in both time and money to implement the two big frameworks.

Lightweight frameworks geared towards small and medium size organizations have been based almost exclusively on prescriptive assessment models. One big reason is because these types of organizations generally lack the experts in process improvement who are dedicated to study the organization's goals, processes and problems that are needed for an inductive assessment model [?]. In general, most of the lightweight SPI frameworks are a lot more detailed in how they work and how to be implemented so to help the organizations to get through the process themselves without help from external personal. Generally small and medium sized organizations are very reactive and flexible and typically having flat structures, all that combined encourages entrepreneurship and innovation [?].

There are any number of proposed lightweight frameworks for both SPI an process assessment, including but not limited to COMPETISOFT [?], MA-MPS [?], and METvalCOMPETISOFT [?]. In common for all these frameworks and others are that even though process assessment is discussed, no detailed protocols are defined beyond mentioning a few different data collection techniques and that the collected data needs to be mapped against the maturity model used. As of yet there is no lightweight framework that have come to dominate in popularity.

2.4 Success factors and best practices for software process assessment and improvements

Success factors for quality management have at large been carried over into SPI, some of them being pointed out already in Humphrey's first publication of CMM [?]. These have then been following along with the evolution of SPI and others have been identified along the way. A few, supposedly well-known factors, have also been proven as irrelevant to the success of SPI since they first were presented. Some of the success factors presented for the improvement process at large carry into the step of the assessment process.

2.4.1 SPI success factors

Factors that affect the outcome of a software process improvement is studied in a large variety of papers, such as [?, ?, ?, ?]. Even more papers study the factors in quality management in general and most of them are applicable to SPI as well. In the following subsections, a few of the most important success factors are presented.

Business orientation

Business orientation means that the goals and actions of the SPI align with explicit and implicit goals and strategies of the organization. It is identified as one of the critical success factors, by Dybå's literature study [?], with the most influence of the results. He also comments that the requirement for this to happen is effective communication between different groups and their needs and problems.

Management commitment

Commitment from management is considered as one of the most important success factors for any quality management according to a lot of research. It is such an important factor that in the general SPI model presented above it has its own step. However, in [?] this is disproven as an important factor, in predicting the success of SPI, confirming what Abrahamson found in [?]. Instead Abrahamson speculates that it might be that someone is championing the SPI, that is, someone who is enthusiastic and goes beyond their role to make the SPI happen. In [?] it is speculated if need for management is so low could be because outside of management needing to grant time and money for the organization to complete an SPI process, they are not very much involved.

Employee commitment and participation

Employees of an organization need to take an active role in the assessment and SPI cycle at large for a successful result. Engagement of participants have two different components, participation and involvement. Participation is defined as the person participates in activities related to the SPI and involvement is the subjective psychological state reflecting the importance and personal relevance a participant puts on an SPI process. [?]

Engagement at large is found to be an important success factory in [?] as supported by other research. In [?] employee involvement is more strongly correlated to success for assessment and improvement than participation.

Measurements

Measuring data during the SPI process and to what extent that data is used to guide and assess the effects of the SPI effort is shown to have a strong correlation with the success of SPI in [?]. Having data and measurements gives the opportunity to later verify if an improvement activity has actually had the intended effect, and thus validate the SPI initiative. The data in [?] shows that the most effective way of using this data is to feed it back into the organization. This would let the employees use the data in the organization and not limit the data to be used by the managers as basis for decision making.

The human factor

In [?] they argue that one of the reasons that SPI, and the assessment process in particular, fail is because of a lack of understanding of the human in the process. As a consequence of this the process assessments and improvements become less correct and effective. It is a fact that people are an integral part of the software processes, and hence important to both the assessment and improvement of those processes [?, ?].

Process assessment is to a large degree collecting participants mental model of a process, as Humphrey puts it [?]. Four types of processes are referred to in Humphrey's book [?, p. 416], as follows;

- Perceived process, the process the person thinks he/she follows.
- Actual process, the process the person actually follows.
- Official process, the process that is written in official documents and approved by management.
- Target process, the ideal process that is to be reached.

The goal of any process assessment is to identify the actual process. The involvement of humans in the process to find that actual process is a must, but often overlooked by simplified models of the assessment process and what to do. An understanding of the different process types and taking them into account would require understanding the humans in the process.

2.4.2 Assessment processes best practices

The vast majority of research being done into success factors and best practices in the SPI field have been focused on the improvement process as a whole. Only a few papers have been found discussing the success of a process assessment process. One such paper is [?] which is a literature study into best practices for success in the assessment process. The paper resulted in a list of 38 best practices sorted into five categories. Each category is presented in short in the following subsections along with the most important practices of each category. The full list of best practices can be found in Appendix ?? where they are all shortly presented.

Assessment method best practices

In [?] 13 method best practices (MBP) were found. Two of these thirteen practices were found with a much higher frequency then the others and thus are at least the most thought of for the assessment method to succeed.

The first one of these, MBP-4, focuses on creating a flexible process with the focus of what is the highest priority for the organization. It is important, as discussed above, that the business goals and the process goals align. A good, lightweight, assessment method should allow for an organization to customize the assessment process to focus on the most acute need for their organization without too much problems.

MBP-10 defines that a questionnaire should be simple and well-structured as well as limited to 150 questions. The idea for this best practice is the need for structure and balance between thoroughness and time needed for an assessment with a questionnaire.

Supportive-tools best practices

Of the six supportive-tools best practices (SBP) defined in [?] the first two, SBP-1 and SBP-2 are the two most important according to their study. The first practice states that the tool should support the various assessment phases. That is, all phases of the assessment should be supported by the tool from collecting data and storing it, to analyzing it and then presenting it. The tool should also be able to visualize the collected data for the model in a useful way for the organization and users.

SBP-2 states that the data collected by the tool should be stored so that a historical database can be built. This is so that the data from previous iterations of the assessment can be used in new assessment iterations and providing data for what improvements have been made.

Procedure best practices

Zarour et al. defined five procedure best practices (PBP)in [?]. All five of the found practices have a relatively high frequency. With the except for the user best practices all other practice categories have a high variance in frequency of when the practice.

PBP-1 is the practice of preparing for the assessment process. It defines that there should be a preparation phase that includes all necessary steps to complete a successful assessment.

PBP-2 defines the practice of building a confidence and trust relationship with sponsors and assessment participants, and that it should be done face-to-face and not by phone or email.

PBP-3 is the practice of producing an assessment report to be delivered out into the organization, and what should be included in it as well as who should get it as a minimum.

PBP-4 defines the practice of ensuring confidentiality for all people involved in the assessment, providing data.

PBP-5 is the practice of having a feedback session after each assessment with the organization where the assessor presents the results of the assessment to the organization. This session should also give the opportunity to discuss participants comments and suggestions for the future.

Documentation best practices

The documentation best practices (DBP) in [?], are practices that tries to identify what an assessment method needs to have in ways of documentation to make it easier to complete an assessment with good results. Three of the defined eight practices have a high enough frequency to be seen as most likely relevant. These practices are defined as follows;

DBP-5 states that documents templates of the documents to be produced at the end of the assessment should be provided to reduce the effort needing to create the resulting report of the assessment.

DBP-6 defines the practice of providing guidance documents of the assessment method and the implementation in practice. Meaning that the entire assessment process must be documented, in the practice it is also included what should be included in the guiding documents.

DBP-7 details that guidance documents on how to document data collection and how that data was rated should be provided. These documents that should be produced by the assessment is then to be included in the assessment report.

User best practices

User best practices (UBP) are all the best practices concerning the humans in the assessment process. Like the procedure category, all six best practices in user best practices have a high relative frequency in the studied papers, and as such is most likely relevant. [25]

UBP-1 is the practice of defining the responsibilities of the assessment participants, for example the sponsors and interviewees as well as the assessment team.

UBP-2, sets the practice that there should be a definition of the assessment team's credentials and responsibilities. This would include what expertise the team needs to include, and what training is needed to follow the assessment method. Also the accessibility to the documents required to complete the assessment is a thing taken into consideration in this practice.

UBP-3 is to ensure the involvement of senior management and other staff members that need to be included. This would include making sure management can attend the meetings needed and helping in setting priorities.

UBP-4, defines that the sponsor must be committed to implementing the assessment method.

UBP-5 is the practice of ensuring that participants feel the benefits of the assessment. It is the assessment team's responsibility in building trust with the participants and making sure the assessment does bring value to the participants, mostly by doing due diligence during the preparation phase.

UBP-6 defines the practice of improving the credibility of both sponsors and staff who should be confident that the assessment will yield results. This is especially important to have credibility with interviewees so that they feel that their participation will help with brining good results to the organization.

2.5 The HAVOSS maturity model

The Handling Vulnerabilities in third party OSS (HAVOSS) model is a maturity model developed by three researchers at Lund University. The model aims to help “managing vulnerabilities in third party libraries and code as well as the subsequent software update activities that are required to limit a product’s exposure to attacks” [?]. The model has taken inspiration from among other things CMMI, and because of this there are obvious similarities. To create a specific maturity model for the given problem, the researchers tried to gather all the features of vulnerability handling from multiple of maturity models in security and maintenance with the focus on third party code. The HAVOSS model should because of this focus not be used in replacement of other, general maturity models but as a complement.

The model contains a total of 21 practices divided into six related capability areas. Each practice has one corresponding question in the model. The six capability areas are product knowledge, identification and monitoring of sources, evaluating vulnerabilities, remedy of vulnerabilities, delivering updates and communication.

Product knowledge is the area where the company’s understanding and knowledge of their products’ components. An increase in the maturity level in this area indicates and increase in knowledge of the components in the products. There are five practices in the area of product knowledge, thus five questions.

The area of identification and monitoring of sources includes three practices. These three practices are all linked to how vulnerabilities are found both externally and internally as well as how these sources are monitored. Evaluating vulnerabilities as an area helps organizations assess how they handle the evaluation of the severeness and relevance of identified vulnerabilities and contains two practices. The following area, remedy of vulnerabilities, includes practices for how different degrees of severity is handled in the organization. The area holds three practices to reflect the commonly used three categories of severity, urgent, needs fixing and those that does not need fixing. The fifth capability area is delivering updates. This area holds the two practices of how updates are delivered to the devices that needs updating, as it relates to security updates. The last area of capabilities is communication. These practices are the those of communicating, both internally and externally about identified and resolved vulnerabilities. In total six different practices around this are established in the HAVOSS maturity model.

These relationship between these capability areas are shown in Figure **??**. The area or product knowledge is a prerequisite for the other areas of the model, without understanding in this area the other simply cannot be understood. Then the four following areas are a sequence of steps where one leads to the other. Along these four areas are communication going out and should happen in sync with each of the areas and steps taken in them.

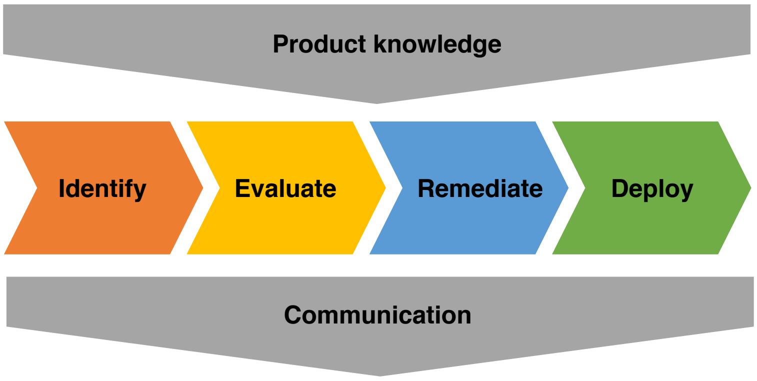


Figure 2.4: The HAVOSS capability areas, illustration inspired from [?]

As with CMM and CMMI the HAVOSS maturity model have five different levels of maturity with a higher level indicating a higher definition and standardization in the organization for security in in their components. The first level, level 0, indicates that no effort is spent on a practice, this is then followed by slowly increasing formality up to level 4, where experiences are collected using standardized procedures and those experiences are then used to constantly improve the processes. These steps, as given in the HAVOSS assessment from [?] are described in Table **??**.

Table 2.1: HAVOSS answer alternatives to the questions

Level	**Answer alternative**
0	We don't do this
1	We do this in an ad-hoc way based on individual's own initiatives
2	We know how we do this, but we do it in different ways in different teams/products
3	We have defined processes for this that are common to all teams/products
4	We collect experience and/or metrics from our approach and base improvements on that

In the work of developing the HAVOSS model, the researchers created a basic assessment sheet, converting the 21 practices into 21 questions. Each of these questions were given as question with the answer alternatives the same as the different levels of maturity, as shown in Table **??**. These questions and assessment were then used to evaluate the maturity model. The evaluations showed that the defined practices are highly relevant to organizations, giving credit to the model.

2.5.1 The Delphi Method

The Delphi method is a method for collecting opinions and reaching consensus on questions by a group of experts [?]. This is done by having an iterative process that is structured and organized with the aim to distill and correlate opinions from a group of individuals concerned with a question or problem. Often the process is done by a questionnaire that each individual in the group, then as a second iteration the answers from the last questionnaire is presented before all the participants fill in the questionnaire again. This process is then repeated a predetermined number of times, or until a predetermined percent of consensus is reached.

The Delphi method was developed during the 1950s by a US Department of Defense think thank called RAND Corporation. It was released for public use in both research and companies in 1963 and since then it has become widely popular as a research tool in many different disciplines.

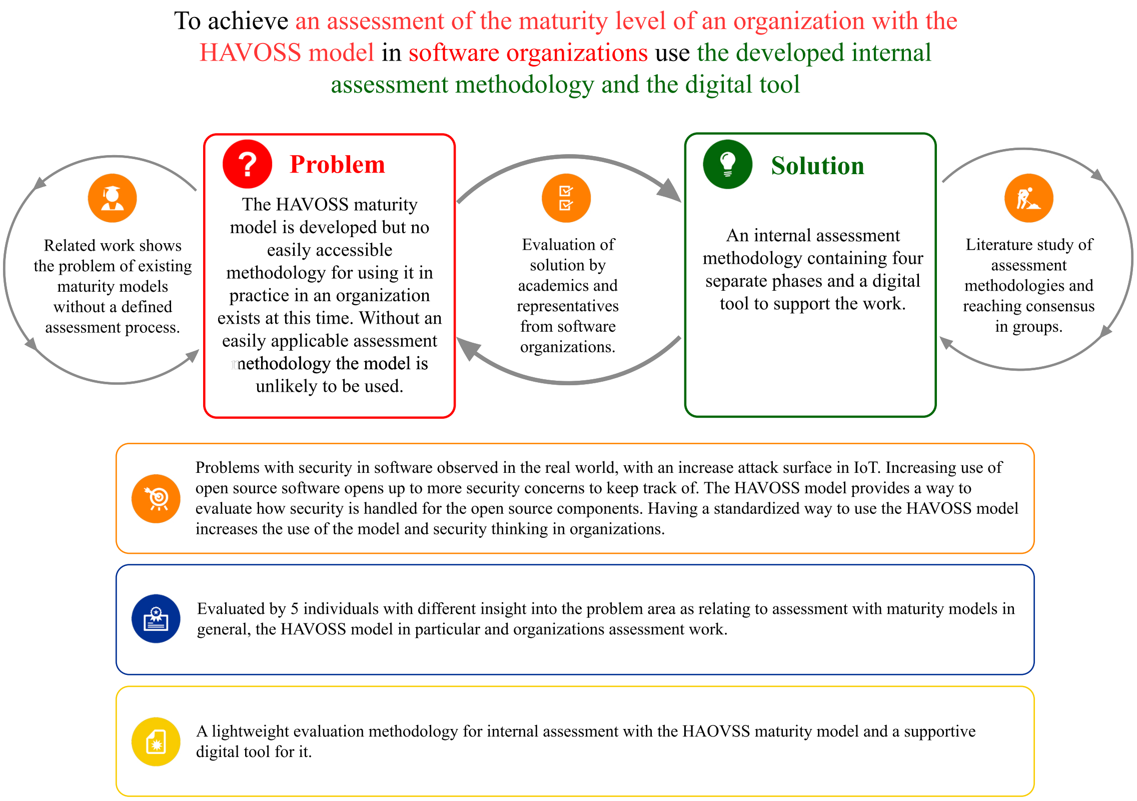
Chapter 3
Methodology

In this section of the thesis the methodology used during the work is presented. First an overview of the thesis work and guiding principles. In section **??**, each step of the work with the thesis is presented in detail. This is followed in sections **??** and **??** with more detail in how a few specific steps were taken. Finally in section **??**, the validity of the thesis is presented.

3.1 Thesis overview

For this master thesis the author looked at software process assessment methods that can be applied as an internal assessment in an organization, using the HAVOSS [**?**] maturity model. A method for the internal assessment was devised based on best practices for process improvement work in general.

This thesis studies information systems in an organization around security software and specifically of IoT devices thanks to the HAVOSS model. Information systems and its research is in the convergence of people, organizations and technology [**?**] and as such is research into all three of these areas. Hevner et al. in their article "Design Science in Information System Research" [**?**] argues that there are two paradigms in information systems research, behavioral-science and design-science. Behavioral-science sets out to develop and justify theories that explain or predict organizational and human phenomena surrounding the information systems. Design-science is a problem solving paradigm that "creates and evaluates IT artifacts intended to solve identified organizational problems" [**?**]. The work in this thesis is in the design paradigm. Design-science addresses research through the building and evaluation of artifacts designed to meet the identified business needs and has as a goal to find utility of the found results. As such, the author have tried to follow the seven guidelines outlined in the paper by Hevner et al. [**?**] to produce a robust thesis and research results.



To achieve an assessment of the maturity level of an organization with the HAVOSS model in software organizations use the developed internal assessment methodology and the digital tool

Related work shows the problem of existing maturity models without a defined assessment process.

Problem

The HAVOSS maturity model is developed but no easily accessible methodology for using it in practice in an organization exists at this time. Without an easily applicable assessment methodology the model is unlikely to be used.

Evaluation of solution by academics and representatives from software organizations.

Solution

An internal assessment methodology containing four separate phases and a digital tool to support the work.

Literature study of assessment methodologies and reaching consensus in groups.

Problems with security in software observed in the real world, with an increase attack surface in IoT. Increasing use of open source software opens up to more security concerns to keep track of. The HAVOSS model provides a way to evaluate how security is handled for the open source components. Having a standardized way to use the HAVOSS model increases the use of the model and security thinking in organizations.

Evaluated by 5 individuals with different insight into the problem area as relating to assessment with maturity models in general, the HAVOSS model in particular and organizations assessment work.

A lightweight evaluation methodology for internal assessment with the HAOVSS maturity model and a supportive digital tool for it.

Figure 3.1: Visual abstract for the thesis work, inspired by [?].

To visualize and make design science more accessible to both academic researchers and industry alike the authors of [?] suggests doing a visual abstract. The visual abstract for this thesis is shown in Figure ?? to given an overview of the entire thesis work and the process to reach the results. The text at the top of the visual abstract summarizes the problem and the solution of the thesis in one sentence. The middle part of the abstract shows the process taken during the work of the thesis. On the right there is the work to present and understand the problem of the thesis. To the left the solution and how it is reached is presented. In the middle the evaluation of the problem and solution is presented. The bottom part of the visual abstract gives in the first box a bit more of an in-depth presentation of the problem the thesis tries to solve and why it is relevant. The middle box presents how the solution developed in the thesis was evaluated to show that rigor was used to secure the results. The last box at the bottom presented the novelty the research of this thesis presents.

3.2 Thesis process

The work with this thesis was semi-sequential, as illustrated by Figure ??. Figure ?? shows the detailed steps of the process that is described in Figure Figure ?? above. The literature study as step in in Figure ??, relates to both the related works area

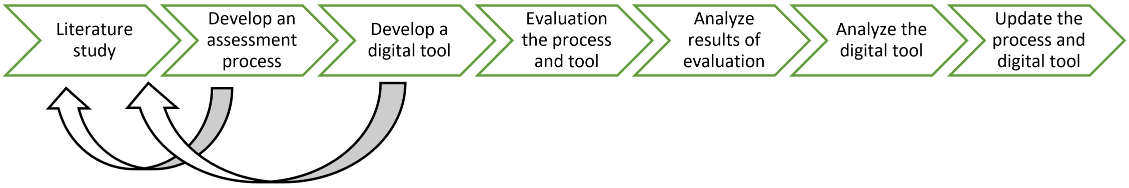

Literature study
Develop an assessment process
Develop a digital tool
Evaluation the process and tool
Analyze results of evaluation
Analyze the digital tool
Update the process and digital tool

Figure 3.2: Illustration of the process for this thesis.

and literature study shown in Figure ??. The evaluation step in the middle of both Figure ?? and Figure ?? also are the same, linking both the figures together with different visualization.

The individual steps of the thesis process were in sequence but each of the steps have in their sub steps been iterative to reach as good a result as possible. In addition, with the learning that happened at each step of the process a few times showed faults or things lacking from previous steps. This required going back to previous steps and fill the information gap or update work already completed. This happened two times, illustrated with the respective arrows in Figure ??.

In the development of the assessment process the literature study needed to be expanded. When developing the digital tool, again the literature study needed to be expanded and some of the assessment process adjusted slightly to fit the new information. This thesis report was written alongside each of the steps in the research process, resulting in a report as the work studied was completed.

3.2.1 Literature study

As a first step in this thesis work a literature study was done to understand the subject at hand as well as a get good knowledge of what research exists in the area. This was then followed up by a deeper dive into the literature to find the information needed to continue to the next step of the thesis process. This was done by finding best practices and how the previously studied processes in research worked, and how they were found to work in case studies at organizations.

Multiple online sources were used to find material, and most often the material was reviewed academic papers, but some websites were also used to get an understanding of any given subject. For the most part, two databases were used as a source for information, Lund University Library provides an online database search tool called LUBSearch with access to a wide array of academic papers. Google Scholar, Google's own tool for searching academic works and books, was also utilized as a wider tool to search for more material.

The first round of searches was for the keywords "software process assessment", "software process assessment methods" and "lightweight software process assessment methods". Based on the results from these other searches were done with less rigor to find deeper understanding of a subject or results in a paper.

For material to be used in the study, it needed to be peer reviewed articles or published books with authors established as authorities in their respective area. The material also had to be related to software quality management or process assessment methodologies and their respective common work methods. Papers presenting other maturity models without a specific assessment process were rejected. Same with papers focusing on the cost of an assessment, the improvement process instead of the assessment process, or the focus was not on small or medium size organizations.

The first few pages of results for a search were evaluated to be included in the study. Each paper included were read and results were written down in abbreviated form to be able to compare with other papers. No formal method for the literature study was used.

The result of this literature study can be found in Section ??. It was also used in the continued steps of this thesis.

3.2.2 Develop an assessment process

An internal assessment process for the HAVOSS model was developed. The process is intended to be used by organizations as a way to self-assess how their security management and communication relating to their use of OSS and COTS in their products work and find ways to improve if needed.

This was done by combining the best practices and success factors found in the literature study with multiple different existing models' assessment processes traits that fits the needs of the HAVOSS model, and the method being developed. Special care was given to create a process where a high user engagement from the entire organization can be reached.

The developed assessment process is presented in Section ??.

3.2.3 Develop a digital tool for the assessment process

The digital tool was developed as a prototyped to assist in the assessment process to avoid using pen and paper and reduce the actual workload of completing the assessment.

The main goal of the prototype was to create a tool for the individual assessments made by members in the organization and creating a summary to be used as a tool in the group discussions later in the assessment process. The results of this step are presented in Section ??.

3.2.4 Evaluation of the assessment process and the digital tool

In this step of the thesis work, evaluations of the assessment process and the digital tool from the previous steps were conducted. The work in this thesis has been done completely academic, without any involvement from any organization, but the work of developing the HAVOSS model has been done in conjuncture with the industry. Because of this the option to test with a real team in a real setting were, due to time constraints, nonexistent. Instead the evaluation was done by the faculty and PhD students at Lund University and two representatives from two large organizations in

Sweden. Not everyone got to do the exact same thing during the evaluation due to previous knowledge about the HAVOSS model and digital tool respectively. Table **??** shows an overview of the evaluators and what they knew beforehand, where they worked and if they were asked to “think out loud”.

Table 3.1: Overview of the evaluators in the evaluation and if they participated in “think out loud”

Participant	**A**	**B**	**C**	**D**	**E**
Previous knowledge of HAVOSS	Yes	Yes	No	No	No
Employment	Lund University			Company A	Company B
“Think out loud”	No	No	Yes	Yes	Yes

The evaluation was done individually with each participant. For those participants where the HAVOSS model was not previously known an introduction into the model and the purpose of an assessment was given. Then all participants were presented to the prerequisites phase and phase 1. After the presentation a short interview for the first part was conducted, using the questions in Appendix **??**, Section **??**. This interview part was ended with the open-ended question, “anything more you would like to add”.

The next part of the evaluation was introduced with the scenario that they were an employee at an organization, and had gotten an email link to the digital tool that they were now going to use to do the individual assessment, phase 2. Using a computer, all participants got to go through at least part of the assessment to test out the digital tool. Three of the evaluators were asked to “think out loud” as described in Section **??**. This resulted in observational data and notes that the authors of this thesis took as the observers. During this test, all participants were at some point asked to do the following:

1. leave a comment
2. answer no to an area question
3. go back to a previous question
4. skip ahead to the next question.

After the test of the digital tool for the individual assessment was completed, a second short interview about the experience was held, with the questions detailed in Section **??**.

This was followed by a third part of the evaluation, presenting the second step in phase 2 and then an introductory presentation of the workshop and how that should work. Then the evaluators were shown the digital tool for the workshop, and to go through parts of it. As before, some participants were asked to think out loud. During the evaluation of the digital tool, the following questions were asked as the evaluators were using the tool:

1. What answer was the most common to this question in the individual assessment?

2. At what level did the individual assessment put the organization at for this question?
3. What was the end score for the area “Product knowledge” for the workshop?
4. What is the total calculated maturity level for the organization?

All evaluators just did the first four questions of the workshop as it was the same questions as the individual assessment and nothing in the digital tool differ between the questions. After those first four questions, they were shown the summary from a full, presumed, workshop. After the questions about the digital tool were tested, a third short interview were conducted using the questions in Section **??**.

At last, phase 4 was presented before a closing interview was held, with the questions presented in Section **??**. The evaluation finished up with a final opportunity for the authors for clarifications of the evaluator and the evaluator to give thoughts and suggestions on any phase of the assessment or the assessment as a whole. The entire evaluation was recorded and then later transcribed in part. The results of these evaluations are presented in Section **??**.

3.2.5 Analyze the assessment process

The transcripts from the evaluations from the previous step were analyzed by condensing them and looking at keywords to look for similarities in the answers. The sentiment of the answers was also studied. Along with the evaluation, the process was compared to the optimal processes studied in the literature study to compare what is outside of the optimal process and what hits the mark of the optimal processes. The comments made are also taken as a whole to be considered for improvements of the process and the individual steps of it. The results of this are presented in Section **??**.

3.2.6 Analyze the digital tools part of the assessment process

The digital tool was also evaluated. The goal was to understand if the digital tool contributes value or a tool without real benefit. The following questions were asked of the tool in this analysis:

1. Does the tool reduce the time for an individual to do the individual assessment?
2. Does the tool reduce the time needed for the assessment team to complete the work of the individual assessment and prepare the workshop?
3. Does the tool fulfill the best practices defined for a supporting tool in [**?**]?
4. Does the tool support the workshop with decisions and results of the workshop?
5. Does the tool visualize the results of the assessment for the entire organization to use later as a foundation for the rest of the SPI process?

The analysis uses the data and comments from the evaluation done in the previous phase along with the authors own thoughts on the subject with the help of literature. The results of this analysis are presented in Section **??**.

3.2.7 Updating the process and digital tool

Based on the evaluations, and the analysis in the previous steps of the work of the thesis, both the developed assessment methodology and the digital tool was updated. The process was mainly updated based on the experience of the evaluators and what they had suggested. These updates are presented in Section **??**. The digital got some major updates to fix usability problems discovered during the evaluations and things found missing. The updates to the digital tool are presented in Section **??**.

3.3 Interviews

Interviews were done as a method of gathering other people's beliefs and thoughts on the subject at hand, what is called qualitive data. It is important to note that the data gathered at interviews are not first-hand or objective but subjective and might be biased. Despite this, it is a great way to get deeper understanding of the subject and more material to work on.

When preparing for the interviews the questions that needed to be answered were created. These questions were adapted during the interview as need be and new questions added, mostly as follow up questions. Most of the questions were qualitative, with a few exceptions of quantitative ones. When constructing the questions a few guiding tenants were used:

- Only ask one thing at a time
- Avoid leading questions
- Make sure the questions are clear and easy to understand.

Each interview was started with a brief overview of the topic of the interview and what was hoped to be learned from the questions.

3.4 User testing of the digital tool

The digital tool was evaluated with the goal to find answers to questions asked in the previous section. First the evaluators were asked to use the tool and then later reflecting on their experience with the tool.

The evaluation was done by giving the evaluator the digital tool for the first time. None of the evaluators had any previous experience with the tool itself, some of the evaluators did however know the HAVOSS model beforehand. All evaluators were asked to complete specific tasks during their evaluation of the digital tool, as described in Section **??**.

3.4.1 "Think out loud"

During the process of evaluating the digital tool, some evaluators were asked to try and verbalize their thinking to allow the test leader, the author of this thesis, to gain insight in their cognitive process [?]. "Think out loud" as this is called, is a usability testing protocol where the participant is asked to use the software while continuously talking out loud, that is to verbalize their thoughts as they move through the user interface to complete tasks [6]. During the tests the test leader needs to be as quiet as possible, stopping themselves from giving answers and feedback as this can bias the results of the tests. Test participants tend to stop thinking out loud during tests, just because it is unnatural or the process they are doing requires their full cognitive load. During these situations the test leader prompted the participant to "continue thinking out loud".

The think out loud protocol is considered one of the most, if not the most, popular way of testing user interfaces and user experiences today. This process has some downsides, such as it creates an unnatural situation for the test person that can become uncomfortable. To mitigate this the test leader and test person had a conversation before starting about why the test is done and that any trouble the test person might have shows problem with the software, not the person using it. Another downside the think out loud protocol is that it still filtered content from the test participants as it is simply impossible to tell everything that goes on in the participants mind. It is also quite possible that the test participants adjust their way of approaching the software simply to try and please the test leader, invalidating the results of the test. [?][?]

During the tests the test leader observed the test participants, writing down observations such as body language, and interactions with the software to have more points of data for analysis. The test sessions were recorded for audio only. These recordings were later transcribed.

3.5 Validity

Research is never perfect, and the validity of any research relies on the trustworthiness of the results. Validity was addressed in all stages of the research done for this thesis in the hopes of increasing the validity of the findings. Even though the research done in this thesis is not a case study, the validity classification scheme presented by Runeson and Höst in [?] are well founded in research and were thus used. As such, four different aspects of validity were considered during this thesis work, as presented below.

3.5.1 Construct validity

Construct validity is the aspect of validity of to what extent the operational constructs in the research are interpreted in the same way by the participants of the research and the researchers themselves and is aligned with the "real world". For the created assessment process and the digital tool we used as general language as possible, so that the tool can be applicable to as many organizations as possible. The same language rules were applied for communication in the evaluation step of the thesis work to avoid misunderstandings. However, this problem can never be entirely mitigated or avoided.

All work in this thesis was done by one author, and as such, analysis of the HAVOSS model and how it works, as well as other models could be wrong on some point. To avoid this, regular discussions and reviews of the used terminology and intended use of both the HAVOSS model and the developed assessment process were held with one of the authors of the HAVOSS model.

The data collected in this thesis is based on interviews and user tests that are all vocal based along with observations from the authors. This lends itself to misinterpretation from the authors of what the person providing the data actually meant or, in the case of user tests, felt. This risk was minimized by asking follow-up questions during the closing interview to give the authors a solid understanding of the interviewees' point.

3.5.2 Internal validity

Internal validity is mainly of interest when causal relationships are examined and the concern that the researcher is unaware of interdependencies between factors and the extent of the relationships between factors [?]. With the work on this thesis being evaluated by people connected to the HAVOSS research there is the possibility of bias towards specific ways of doing things based on other work done with the model by the evaluators. Another such problem would be that all the evaluation is done by academics and not actual software development firms and their perspectives. This is at least slightly mitigated with the fact that the HAVOSS model was developed in conjuncture with the industry and thus the researchers, who are also the evaluators in this thesis, have an understanding of how these organizations work. Another mitigating factor was that the authors of this thesis did not have previous knowledge of the HAVOSS model before the work of the thesis began, or much knowledge of process assessment and developed a first model of the assessment process from literature and not the researchers of the HAVOSS model.

3.5.3 External validity

External validity is the generalizations of the results. That is, to what extent are the research results possible to generalize, and of interest, to other people [?]. The work done, has a starting point of not being specialized for any specific organization. With a single author only one mental image of how organizations work existed. That is a potential risk, especially since the authors have little real experience of industry organizations and thus that image might be wrong or not as generalized as thought. This is mitigated in large by using many different literature sources that have results from tests in organizations.

3.5.4 Reliability

Reliability is the aspect of validity that concerns to what extent the data and the analysis are dependent on the tools chosen to collect and interpretation of data [?]. This thesis was done with a single author, and as such this is an even bigger threat. To counteract it, great care was taken to keep reliability in mind during the entire work, but especially in the analysis phase of the thesis work. Discussions were held with the thesis supervisor to make sure the analysis were correctly done.

The user testing sessions with the digital tool were especially vulnerable since the authors and developers of the tool are the same person as the test leader and thus truly can influence the results of the tests if not done correctly. To mitigate this risk a strict test protocol was developed and used, where, among other things, the test leader sat so that the test person could not see the test leader clearly all the time.

Another threat to the reliability was that there was only an evaluation of the assessment process done and no proper testing at full scale in an organization. This threat is especially real since the evaluation was done in large by people already deeply involved with the HAVOSS model. This means that it is possible that an evaluation done by independent researches with no connection to the HAVOSS model specifically could think differently and thus provide different data and results. This is at least slightly mitigated with individual evaluations done by different people that did not have contact with each other about the work done in the thesis, thus giving data independent from each other. The concern is also lessened with the evaluations done by the few evaluators that did not have any previous information about the HAVOSS model. The risk that other researchers would draw different results and conclusions from the data gathered is, as in all research, present. This is mitigated by having discussions with the thesis advisor.

Chapter 4
Results

In this section of the thesis, the assessment method and the digital tool are presented. The assessment method was developed based on the literature study, with the literature described in Section ??. That is followed in Section ?? by an evaluation of the first version of the assessment method and digital tool. After the evaluation, a few changes were made to the assessment process, as presented in Section ??. The digital tool had some major changes, which are presented in Section ??.

4.1 The assessment process

The assessment method presented below is an internal assessment methodology, to be used inside an organization without any need for external experts. It is tailored as an assessment method for the HAVOSS maturity model that assesses the security in software as it relates to OSS and COTS. To accompany this assessment method a digital tool, presented in sections ??, have also been developed.

The assessment process is inspired by SCAMPI [?] that have three distinct phase of the assessment. This assessment process has four, but that is mainly due to having two steps of gathering data that are separate from each other, while SCAMPI have all the gathering of data in the same phase. As for the mapping between the phases in SCAMPI and the developed assessment method, the tasks in the planning phase have their equivalent in the SCAMPI method. The developed methodology has one additional prerequisites phase that must be completed before the assessment process can begin. No other looked at assessment method has defined the prerequisites in their plan, in this method it is added for clarity who completes this step. The process and its phases, along with the people involved in the phases are illustrated in Figure ??.

The assessment process starts in phase one with planning, and selecting partici-

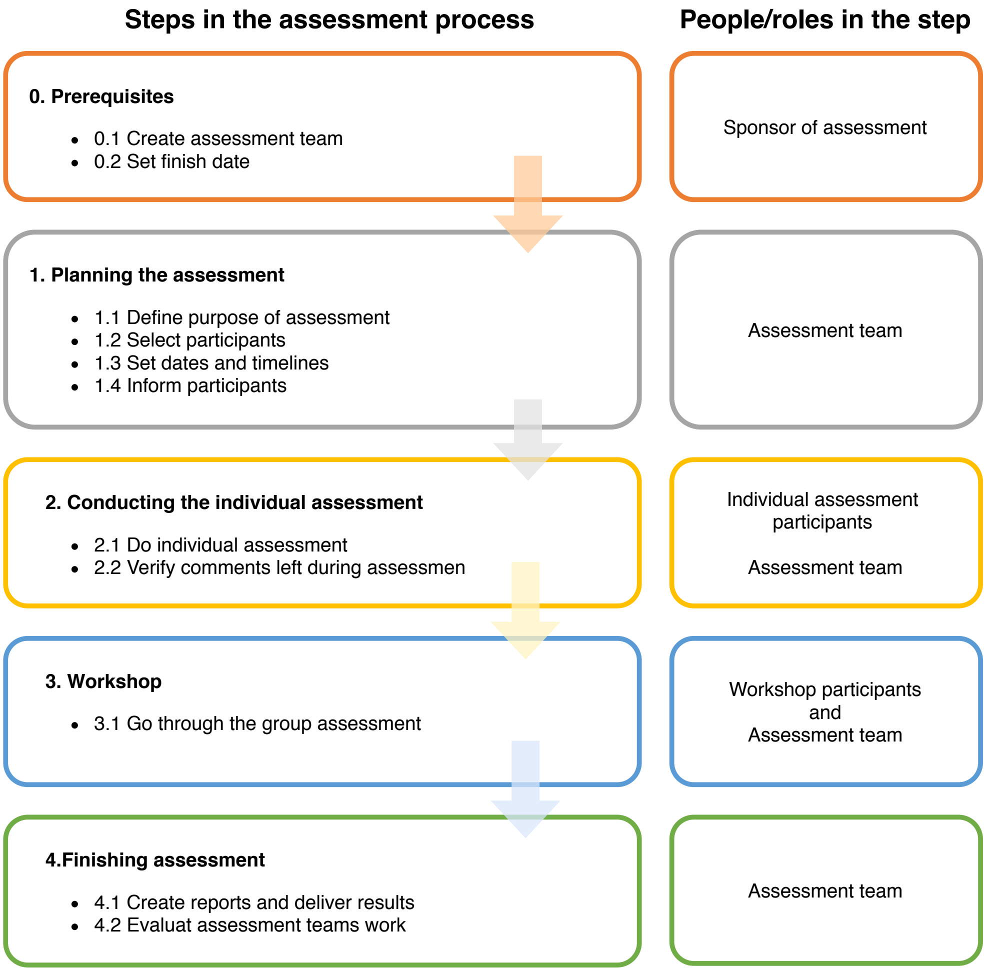

Steps in the assessment process
People/roles in the step
0. Prerequisites
• 0.1 Create assessment team
• 0.2 Set finish date
Sponsor of assessment
1. Planning the assessment
• 1.1 Define purpose of assessment
• 1.2 Select participants
• 1.3 Set dates and timelines
• 1.4 Inform participants
Assessment team
2. Conducting the individual assessment
• 2.1 Do individual assessment
• 2.2 Verify comments left during assessmen
Individual assessment participants
Assessment team
3. Workshop
• 3.1 Go through the group assessment
Workshop participants and Assessment team
4.Finishing assessment
• 4.1 Create reports and deliver results
• 4.2 Evaluat assessment teams work
Assessment team

Figure 4.1: Illustration and overview of the process assessment method developed in this thesis.

pants, all work done by the assessment team that is appointed in the prerequisites phase. Based on that many existing assessment methodologies, including the two most popular, [?, ?] relies to some extent on a questionnaire distributed to a large base of people in the organization this is used in our assessment process in phase 2, the individual assessment. The individual assessment is a questionnaire done with the digital tool where a large group of people in an organization can take part. To then increase the accuracy of the results, phase 3 consists of a workshop where a select group representing the involved parts of the organization comes together. In the workshop the group goes over the results from the individual assessment and based on that, and the groups knowledge a new answer is inputted into the digital tool as the workshop answer. This step is based on the combination theories of the Delphi Method, as presented in Section **??**, of an expert panel reaching consensus with multiple iterations over the questions, with the workshop being the second iteration and human swarms coming to an answer through a group's decision. The assessment process is ended by the assessment team completing a report and sending that off to whoever ordered the assessment.

4.1.1 Prerequisites

Before the actual assessment process can be started, there are a few prerequisites that need to be fulfilled. This would be done in the previous step of the SPI process described in Section **??**, or as preparation for the assessment, then completed by a sponsor.

A sponsor is the individual or group that wants the assessment to be conducted and makes "an order" for it in the organization. The sponsor is not directly involved in the assessment itself but invested in the results and has an agenda for the assessment. It could be to just evaluate where the organization is in the principles of security in relation to the used OSS and COTS, or as part of an improvement process for example. In most cases this would be some level of management or a single manager that wants information from lower levels of the organization.

The first prerequisite is to select an assessment team, that is the group or individual that is assigned to facilitate the assessment. This team is responsible for preparing and guiding the assessment process as well as produce the final report and deliver it to the proper person or group after the finished assessment.

The second prerequisite is to set an expected finishing date for when the entire assessment process is to be completed. In the case of a sponsor they expect the results of the assessment and decides when that needs to be delivered. For a step in the SPI circle this is not as critical but still needed for an effective assessment process.

4.1.2 Phase 1: Planning the assessment

In the planning phase of the assessment the assessment team completes the perpetration work for having a successful assessment. How large this work is depends largely on the size of the organization to be assessed since a larger organization will have more information to analyze. The steps of the planning of the assessment comes from the SCAMPI [**?**] planning phase, reinforced by the best practices presented in Section **??**.

Step 1.1 Define purpose of assessment

The assessment must have a purpose to create commitment and understanding in the organization that is being assessed. It is the job of the assessment team to define a purpose for the assessment that is easy to communicate to the assessed organization and the participants in the assessment. The purpose must also fall into line with the agenda set of the assessment sponsor.

The purpose of the assessment needs to align with the business needs, may it be that it is an important step in the SPI that align with business goals. If the alignment with business goals cannot be found, the need of the assessment should be reconsidered and not be continued until such goals are found. This is because without this alignment it will be a hard sell to get management to commit the resources needed to do the assessment.

The purpose also needs to be grounded in real needs in the organization that the employees can get behind. Pure business needs might not motivate the employees to participate in the assessment wholeheartedly since, on the surface, they do not get anything out of it. Because of this, the purpose might need to be formulated in two ways, one for management and one for the internal use with participants and the organization at large to gain better commitment.

Step 1.2 Select participants

In this step of the preparations the assessment team tries to identify the key people in the organization as pertaining to the security in the software and products. In a small organization this will be easy as everyone is likely involved and will qualify to this group, end of step. In a larger organization this step will be the most involved step of the entire assessment process for the assessment team as the organization needs to be analyzed.

This assessment process has two different types of participants, individual assessment participants and workshop participants. The latter group is by default included in the first group but not the other way around.

As a starting point, nine questions have been formulated to help identify participants, based on the question in the assessment. These questions are:

1. Who is responsible for the security of our software?
2. Who chooses, adds, updates and maintains OSS/COTS components to our software?
3. Who finds out about vulnerabilities in the products and projects of the organization?
4. Who evaluates and/or handles identified vulnerabilities?
5. Who is responsible for delivering updates and patches to customers?
6. Who is responsible for communicating with our customers about security?
7. Who is likely to get questions about security from our customers?
8. Who does the communication to customers about patches and new versions and the security related to them?

Question 8 should be given some extra consideration as this could be done by update notifications in software, and patch notes, then the question becomes who makes this happen and who writes the notes. Questions 5 through 8 are all likely to identify departments or people outside of the technical side of the organization, such as sales or customer support. It is important that these identified parts of the organization are included in the assessment.

One or more of the questions could, especially in a small organization, result in the answer of “everyone” or a group, such as “the developers”. For the individual assessment, all should be included as participants if this is the case.

An important note on selecting the participants is that they need to actually have insight into the topic at hand. It is important to remember that it is not entirely unlikely that for example the entire development team cannot be included as not all work with security and/or OSS and COTS for example. Should individuals without any insight into the topic of the assessment be selected they will not be motivated to partake and might even create an erroneous result if number of participants are low or the number of individuals that should not be included are big enough. However, it is only required that any single participant has insight into a single area of the assessment or just a few questions for them to be included in the assessment and provide valuable insight.

From the participants found for the individual assessment, a much smaller group of workshop participants must be selected. The assessment team should identify a relatively small number of participants for the workshop. The goal is that they can all be present at a single workshop for that phase of the assessment process. What limits this number will differ from organization to organization but generally a number between 5 and 10 could be seen as a good goal, given that there are at least this many employees in the organization.

Using the found participants for the individual assessment, the assessment team should look at what departments are included and chose participants for the workshop to represent that department or group of employees. It is important to not select the managers by default, as they might only work as managers and not do the technical work of the people they manage. It is of utmost importance that the participants in the workshop know what work actually gets done, not what is said on paper should get done and how.

To find the truly important participants with the greatest knowledge of how the organization works with the questions in the HAVOSS model, the assessment team might need to do interviews with members in the organization to find who are fit to be in the workshop. This would be especially true in a larger organization where everyone is not very familiar with all the employees and their functions. Another method of finding the workshop participants is to ask for suggestions of names from already identified participants of the workshop, or even just participants in the individual assessment.

Step 1.3 Set dates

In this step of the planning for the assessment team is to set a deadline for the individual assessments to be completed along with the date of the workshop in phase three.

Having the assessment continue over a long time span might invalidate the results because of changing practices, especially when flaws are discovered as part of the assessment and might be promptly fixed when seen as vital flaws for example. Because of this, a short but realistic time span should be chosen for the assessment, considering holidays and other major events that might impact the time available for the participants.

Step 1.4 Inform participants

After the previous steps are completed, the participants of the assessment need to be informed about the assessment, their expected participation in the assessment and what the timeline is. That information also needs to include how the different parts of the assessment is done. It is up to the assessment team how this is done in their organization.

4.1.3 Phase 2: Conducting the individual assessment

The second phase of the assessment is conducting the individual assessments as a digital questionnaire and gathering the data from that questionnaire to be used in the workshop in phase three.

Step 2.1 The individual assessment

The first, and major step of the second phase consists of an individual assessment that is done with the help of the digital tool. In essence the individual assessment is a questionnaire done digitally consisting of the questions from the HAVOSS model. The participants of the individual assessment go through the questionnaire with the digital tool.

The tool, as is described in greater detail in Section **??**, presents the questions of the assessment to the participants for the first time. Ideally it should be done as an individual exercise without influence from other people or discussions in groups. This is to prevent participants being influenced in how they view the organization and instead go with their knowledge of the organization and the processes in it. The tool also allows the participants to leave comments, both anonymously and named whichever they choose, to give more information to be used in the workshop when they feel that some aspect is not covered by a question or answer.

At the end of the timeframe for the individual assessment, the assessment team should remind anyone that has not completed the assessment but have been selected as a participant to do the assessment as it is vital to get as much good information as possible to create a good assessment of the situation.

Step 2.2 Verify comments left during assessment

After the completion of the individual assessment the assessment team needs to verify any additional data that came in as comments and is not obviously true or false. Depending on the type of information left as comments, the verification can be done in several ways. Examples of ways to verify are: looking at documents and/or code, interviews or conversations with people in the right place or observations of situations that the comment relates to. The results of the verification of the comments should be documented so that during the workshop they can be viewed and used during the discussions there.

The verification, no matter what method is used, should be done discreetly so that the topic is not up for discussion in the entire organization until the workshop to get fresh perspectives on it then, if at all possible. It is also important to be discrete so that no blame game starts geared towards whoever left the comment, especially in a small organization this could become a problem if it is obvious who would leave the comment, even if it was left anonymously.

4.1.4 Phase 3: The workshop

Phase three is a single step, consisting of a workshop, with the participants selected by the assessment team in phase one. The purpose of the workshop is to answer the questions of the HAVOSS model again with the help of the results from the individual assessment.

During the workshop the participants, as a group, goes through the questions of the HAVOSS model, and comes to a joint answer after a group discussion of what the correct answer for their organization would be. To their help the digital tool shows not only the question and possible answers but information from the individual assessment of how the answers were distributed. Any comments left during the individual assessment is also shown by the digital tool to the workshop. More details on how the digital supports the workshop is shown in Section **??**. The goal is to come to a consensus of the answer for the organization, or a majority vote if not all can agree.

When an answer for a question is reached, a short discussion should be held to find improvement suggestions for the question in the organization. Any improvements the group can find is written down by the assessment team. When all the questions have been answered in this manner, the workshop is finished.

During the workshop, the assessment team has the role of moderator. The team needs to make sure that the discussions are kept on topic and that not too much time is spent on any individual question. There should be an allotted time for each question along with a buffer. When the allotted time and some part of the buffer is spent on a single question, the assessment team should ask for a vote on the question if no consensus is reached and move on to find improvements. If it is during the improvement searching that the time is overspent, the question should be closed, and the workshop move on to the next question. The team should also, in the role of moderator, try to stop discussions about who is to blame if any problems are discovered within the organization.

4.1.5 Phase 4: Finishing assessment

In the last phase of the assessment, the assessment team does the final work with the assessment before it can officially be viewed as completed.

Step 4.1 Generate and deliver result reports

The assessment team creates the report requested from the sponsor using the results from the workshop and the individual assessment, possibly adding comments from other data found in the work with the assessment. This report is then delivered to the sponsor. Should no sponsor exist, as in the case of the SPI cycle, a simpler report with the results of the workshop and individual assessment can be created.

This report should also contain all the improvement suggestions found during the workshop.

The organization at large should also be given access to both the results of the assessment and the improvement sheet so that all can see that the work done for the assessment resulted in something concrete in the form of a result to the assessment and the improvement plan. The generated information from the assessment can then be acted upon as the organization sees fit to hopefully improve their work even more.

Step 4.2 Evaluate assessment

In this last step of the assessment, the assessment team conducts an evaluation of their own work to find points of improvement on what changes to make for a better assessment process next time. This evaluation is saved for the next time an assessment is made.

4.2 The digital tool

The developed tool for this thesis work is a prototype. As a prototype, it is built as a very basic version of what could be a complete assessment tool with only the most important functions implemented in the user interface.

For the purpose of this thesis work, a lot of example data was used to imitate the use of digital tool in the workshop. This was done since there was not a test company to test the tool for and thus some sort of data was needed for the individual testing cases of the tool to be made.

The idea of the digital tool is to aid in the administration of the entire assessment, provide a platform for the individual assessment and to support the workshop by presenting information. For the prototype, two main functions were selected to be implemented, with some aid functionality for each. The two functions were;

1. The function to go through the individual assessment, allowing the answering of the 21 questions from the HAVOSS model, adding comments and viewing the upstart questions created for the individual assessment to guide if the user should answer a section of questions at all.

2. Assist in the workshop by showing the results from the individual assessments, and any comments left by the participants in the individual assessment. It also allows for the workshop to enter their answer to the questions in the HAVOSS model.

The prototype is designed as a web application with an interface of a website and a theoretical server backend that was not implemented. In the following subsections the prototype of the digital tool is presented in greater detail.

In this section the first version of the digital tool is presented. This version is the one that was later evaluated, see Section **??**. Based on the evaluation the digital tool went through some updates, these updates are presented in Section **??**.

Table 4.1: Initial questions for capability areas for the individual assessment in the digital tool

Capability area	**Initial question**
Product Knowledge	Do you know how your organization tracks products, its parts and for example the environment?
Identification and monitoring of Sources	Do you know how your organization finds, handles and monitors sources of threats and vulnerabilities from different sources?
Evaluating Vulnerabilities	Do you know how your organization evaluates vulnerabilities?
Remedy of Vulnerabilities	Do you know how your organization fixes vulnerabilities in the products?
Delivering Updates	Do you know how your organization delivers updates and patches to products?
Communication	Do you know how your organization communicates outside of the organization about vulnerabilities and their fixes?

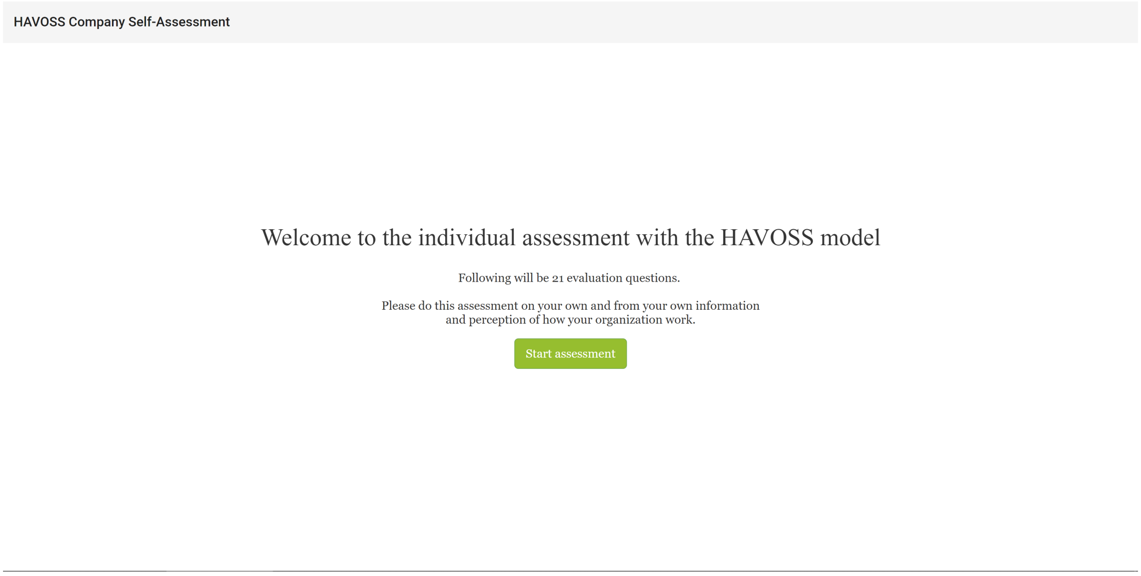
The individual assessment

For the individual assessment, the thought is that each participant has received an email with a link to the assessment, ending up on a screen like the one Figure **??** which is the start screen for the individual assessment in the prototype.

The participant chooses to start the individual assessment and the first introduction question is asked. Each capability area is started off with a question to try and figure out if the participant has the knowledge to answer the questions in the capability area. The initial question for each capability area is presented in Table **??**. The hope is to prevent participants that does not have any knowledge of an area to do the questions and feel that they are wasting their time, thus lowering their engagement in the rest of the assessment. Also preventing that same participant from guessing an answer or entering something just to answer. Figure **??** shows a screenshot from the prototype of how the initial question is shown to the participants, with the answer alternatives "Yes" and "No".

If the participant selects the "Yes" option, the next question shown will be the first question of the capability area. If "No" is selected, all the questions in the capability area are skipped and the introduction question for the next capability area is shown.

Figure **??** shows an illustration of how a question is presented in the prototype. There is a header of the capability area for orientation for the user, followed by the question in bold. Below the question are the five question alternatives shown along with a radio button each. Each of the alternatives represent one level of the HAVOSS maturity model, starting with the lowest, level 0, first. The five question



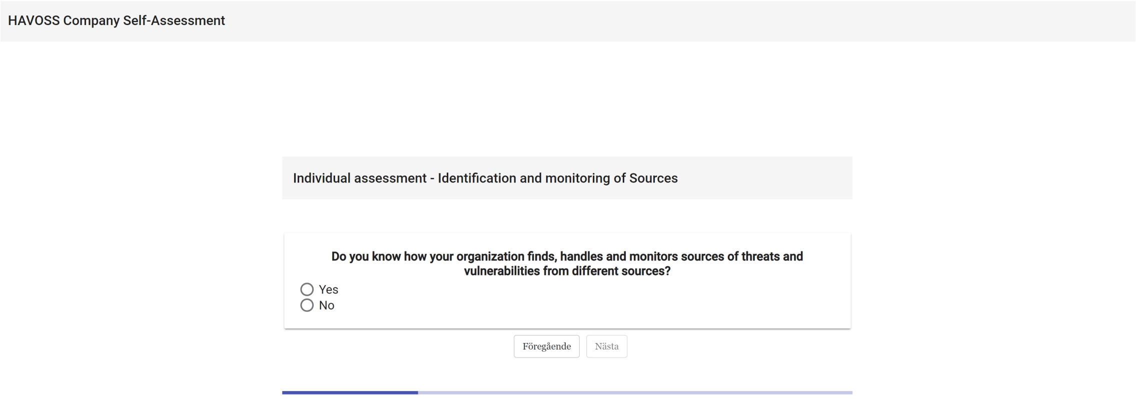
HAVOSS Company Self-Assessment

Welcome to the individual assessment with the HAVOSS model

Following will be 21 evaluation questions.

Please do this assessment on your own and from your own information
and perception of how your organization work.

Figure 4.2: Welcome screen for the individual assessment.



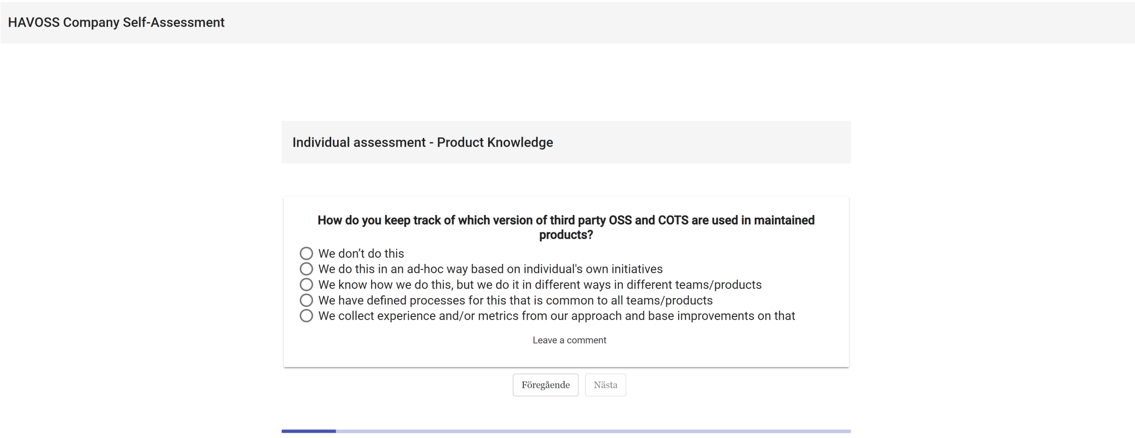
HAVOSS Company Self-Assessment

Individual assessment - Identification and monitoring of Sources

Do you know how your organization finds, handles and monitors sources of threats and vulnerabilities from different sources?

○ Yes
○ No

Figure 4.3: Introduction for a new capability area in the individual assessment.



HAVOSS Company Self-Assessment

Individual assessment - Product Knowledge

How do you keep track of which version of third party OSS and COTS are used in maintained products?

○ We don't do this
○ We do this in an ad-hoc way based on individual's own initiatives
○ We know how we do this, but we do it in different ways in different teams/products
○ We have defined processes for this that is common to all teams/products
○ We collect experience and/or metrics from our approach and base improvements on that

Figure 4.4: A standard question view of a question in the HAVOSS model.

alternatives are;

1. We don't do this

2. We do this in an ad-hoc way based on individual's own initiatives

3. We know how we do this, but we do it in different ways in different teams/products

4. We have defined processes for this that are common to all teams/products.

5. We collect experience and/or metrics from our approach and base improvements on that

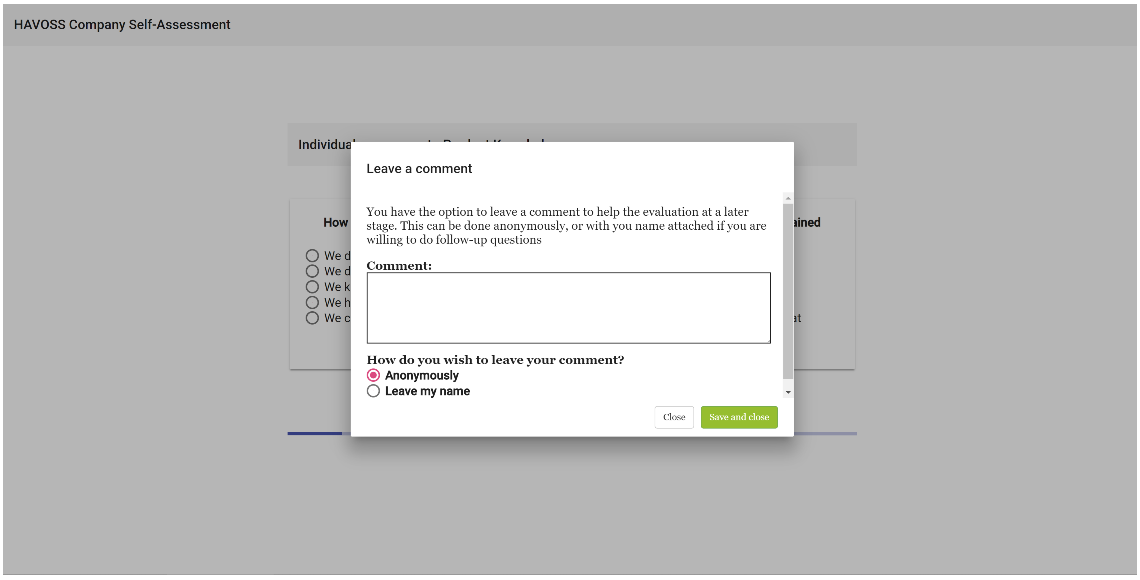
If the answer alternative selected is one of the first three, the participant can move on to the next question with the "next" button that will be green.

There is also the option for the participant to leave a comment. If this option is pressed, the comment section expands out to show the comment function, as shown in Figure **??**. Here the participants can leave their own thoughts and do so either anonymously or with their name attached to it.

When all the 21 questions of the HAVOSS model are worked through by the participant, the individual assessment is completed, and the view shown in Figure **??** is how it is shown in the prototype. With that the participant can close down the digital tool and is thus done with the individual assessment.

The workshop

For the workshop, the idea is that the digital tool is projected up on a big screen so that everyone has the same view. During the workshop someone in the assessment team then controls what step of the workshop view is shown and is responsible for inputting the results from the workshop as the workshop makes decisions.



HAVOSS Company Self-Assessment

Leave a comment

You have the option to leave a comment to help the evaluation at a later stage. This can be done anonymously, or with you name attached if you are willing to do follow-up questions

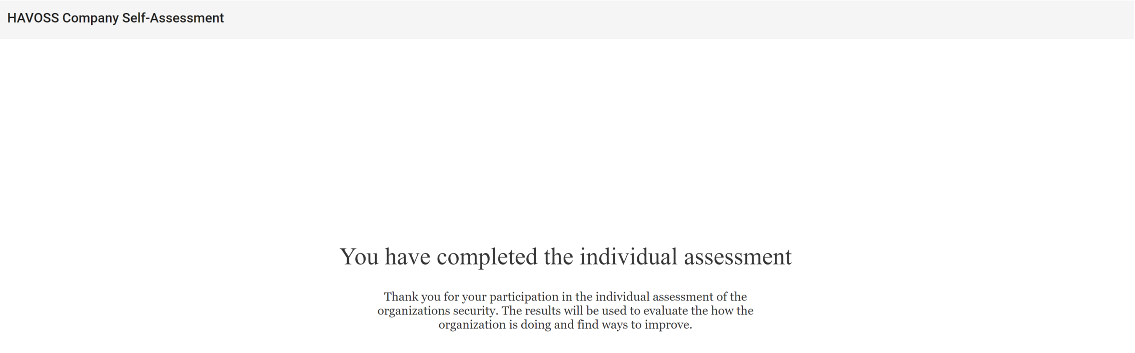
Comment:

How do you wish to leave your comment?

● Anonymously

○ Leave my name

Figure 4.5: Leave a comment view in the individual assessment.

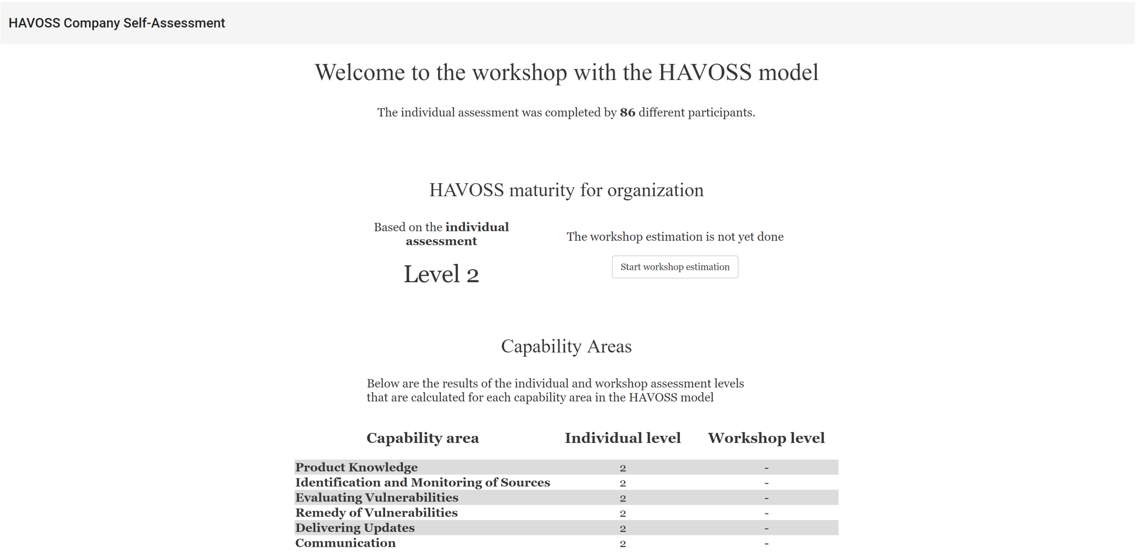


HAVOSS Company Self-Assessment

You have completed the individual assessment

Thank you for your participation in the individual assessment of the organizations security. The results will be used to evaluate the how the organization is doing and find ways to improve.

Figure 4.6: The individual assessment completed in the prototype.



HAVOSS Company Self-Assessment

Welcome to the workshop with the HAVOSS model

The individual assessment was completed by **86** different participants.

HAVOSS maturity for organization

Based on the **individual assessment**

Level 2

The workshop estimation is not yet done

Capability Areas

Below are the results of the individual and workshop assessment levels that are calculated for each capability area in the HAVOSS model

Capability area	Individual level	Workshop level
Product Knowledge	2	-
Identification and Monitoring of Sources	2	-
Evaluating Vulnerabilities	2	-
Remedy of Vulnerabilities	2	-
Delivering Updates	2	-
Communication	2	-

Figure 4.7: Start screen for the workshop in the prototype.

The workshop starts with an overview of the results from the individual assessment. Here the calculated maturity level for the organization based on the individual assessment is shown. Further down the page, the maturity level for each capability area in the model is shown, calculated on the individual assessment as shown in Figure **??**. The aim of this view is to give all the participants of the workshop an overview of what the results were and background information when moving on with the workshop.

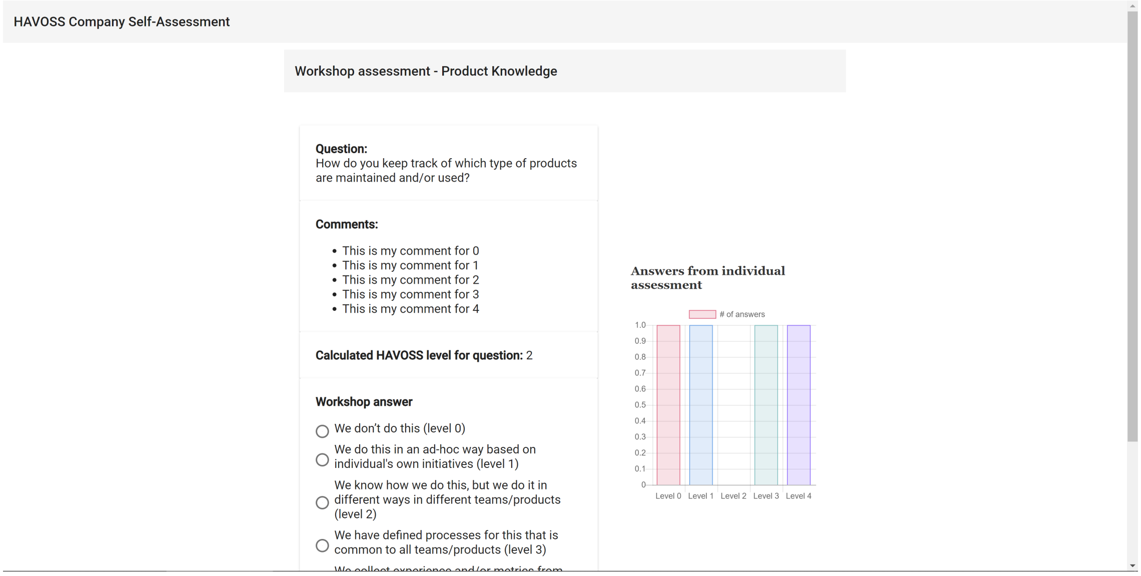
The work of the workshop then starts with clicking the button in the overview to start the workshop estimation. With this, the first question in the HAVOSS model is shown. Below it, the results from the individual assessment is summarized and shown, including the calculated HAVOSS score, any comments left by the participants and how the answers to the question distributed. The next section on the page allows the workshop to enter their answer. All this is shown in Figure **??**.

With all the questions completed by the workshop, the view is returned to the overview that the workshop started with, but now filled in with information that sums up the results of the workshop, see Figure **??**.

4.3 Evaluation of the process and tool

Staff at Lund University and representatives from industry were asked to take part in an evaluation of the assessment method and the digital tool in order to evaluate the first version developed of these. This evaluation was conducted as described in Section **??**. In total five evaluations were conducted.

The goal for the process evaluation was to establish if they thought the process would work, that is if an organization could get through it in the expected manner and produce reliable, representative results. For the digital tool the goal was to find out if the evaluators thought it was contributing to the process, collected enough



HAVOSS Company Self-Assessment

Workshop assessment - Product Knowledge

Question:
How do you keep track of which type of products are maintained and/or used?

Comments:

- This is my comment for 0
- This is my comment for 1
- This is my comment for 2
- This is my comment for 3
- This is my comment for 4

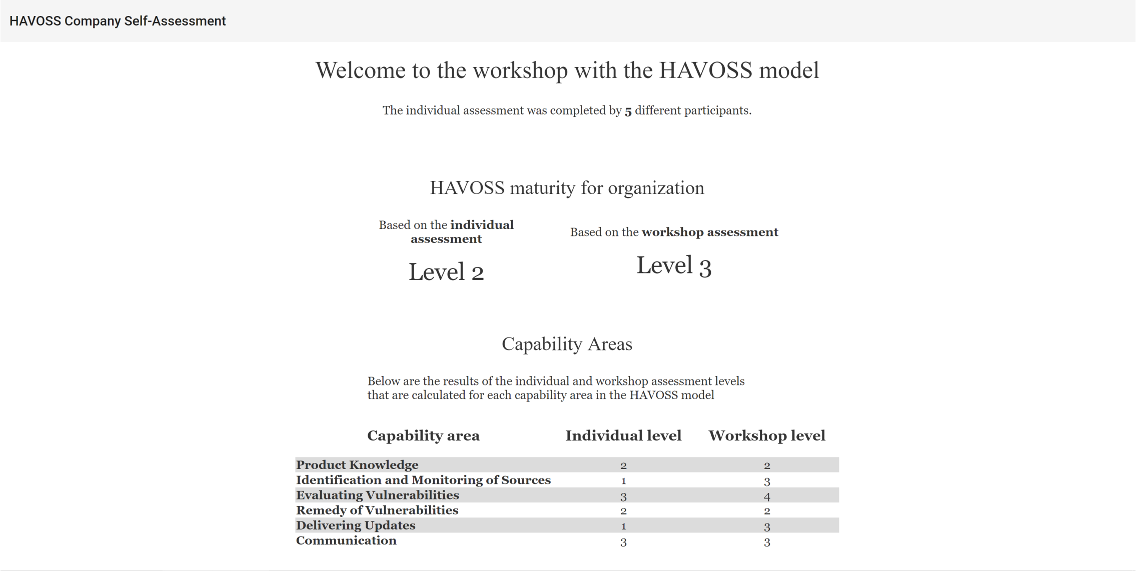
Calculated HAVOSS level for question: 2

Workshop answer

○ We don't do this (level 0)
○ We do this in an ad-hoc way based on individual's own initiatives (level 1)
○ We know how we do this, but we do it in different ways in different teams/products (level 2)
○ We have defined processes for this that is common to all teams/products (level 3)
We collect experience and/or metrics from

Answers from individual assessment

Figure 4.8: A question view with comments in the prototype for the workshop.



HAVOSS Company Self-Assessment

Welcome to the workshop with the HAVOSS model

The individual assessment was completed by **5** different participants.

HAVOSS maturity for organization

Based on the **individual assessment**
Level 2

Based on the **workshop assessment**
Level 3

Capability Areas

Below are the results of the individual and workshop assessment levels that are calculated for each capability area in the HAVOSS model

Capability area	Individual level	Workshop level
Product Knowledge	2	2
Identification and Monitoring of Sources	1	3
Evaluating Vulnerabilities	3	4
Remedy of Vulnerabilities	2	2
Delivering Updates	1	3
Communication	3	3

Figure 4.9: An overview for a completed workshop in the prototype.

information, and had the capabilities needed.

The evaluators' general improvement suggestions were also collected during this process. In the following subsections, the results of these evaluations are presented.

4.3.1 Prerequisites and phase 1

All evaluators commented on the importance of step 1.2 of the assessment process, selecting the correct people to partake in the assessment process. All participants must have insight into at least part of the problem area. The problem with people contributing data that is not correct by using personal opinions or guessing was brought up in multiple evaluations as a threat to the assessment process.

The idea of having an introductory meeting in step 1.4 of the assessment process held by the sponsor of the assessment and the assessment team for all participants was suggested. At that meeting the information about the assessment, the motivation behind it and expected outcome for the organization at large and the individuals, and or departments, are presented. The idea being to raise motivation and commitment to the assessment by the participants. At this point, a time estimate for the different parts of the assessment that the participant is expected to partake in should be given to show how much, or little, time is required to help this process.

4.3.2 The individual assessment

All evaluators stated that more information about the evaluation was needed in the individual assessment. The fact that only OSS and COTS pieces of software are evaluated with the HAVOSS model was missing in the digital tool. Information on how to think when answering the questions were brought up to improve the quality of answers, and to help the participants think of the same thing, be it one or a few projects or released products. Information if the answers were anonymized or not was not apparent. It was pointed out that most likely more explanation about the language used and what was intended with the questions was needed to make it clear to all participants. The same goes for the answering alternatives.

The fact that initial questions about the capability area to understand if the participant knows about the area that is about to start was included was seen as positive, and a way to reduce the risk of guessed answers and boredom from participants. All evaluators thought this function would likely improve the quality of the information gathered.

One evaluator made a point of the need for the assessment to be completed in as a short amount of time as possible. In contrast, another evaluator was concerned that the individual assessment would be done too quickly without afterthought. The suggestion there was to include the requirement to supply more information if the answer to a question was either “we have a process for this” or “we track this”, by defining what process was used or what metrics were tracked. The thought behind this was to remove the possibility to just answer “yes we have a process” and then not be able to clarify what that process is. This solution was discussed with a later evaluator, he identified the risk of the participant identifying that extra work was required by answering any of these two answer alternatives and choose another alternative to reduce the time or energy spent on the assessment.

Three of the five evaluators found the commenting function without prompting, showing on a possible problem with where the button was located. All evaluators gave the comment that it most likely is needed to be able to view the question and answer alternatives at the same time as writing a comment to remember what is going on.

To improve the ability to get good data from the individual assessment, it was suggested that the answers were tracked along with what department, or equivalent, the participant belonged to so as to be able to weight the answers depending on the likelihood that the person knows the actual answer to the question and not just guessing or going on opinion. It was also suggested to track the time spent on the assessment in total and individual questions, again to be able to get an indicator of how much effort was spent on the assessment and give less weight to answers where the participant obviously have rushed through the assessment.

One evaluator pointed out a risk of doing the assessment with the digital tool is the distribution of the links to the tool along with information to do the assessment. This is most likely done by sending out an email and there is a major risk that the email gets lost in a sea of email. In the end the risk is that the assessment does not get done. This risk might be higher than if a paper questionnaire was distributed instead of the digital tool.

4.3.3 The workshop

A few of the evaluators commented on the difficulty of getting a workshop to work properly and reach expected results. One evaluator pointed to the difficultly to reach conclusions on any single questions as participants easily could have the mentality that they or their department is not at fault or could not possibly be functioning at such low level. The same evaluator said that a possible solution to this problem would be to use the workshop as another way for the assessment team to collect information and contrasting viewpoints from different department and then the assessment team decide on what level is correct.

Multiple evaluators pointed out the possible difficulty for the workshop to complete all the expected work in a reasonable amount of time and that multiple sessions might be needed. To reduce the time needed for the workshop, a good moderator was seen as a critical point for success. The moderator would also reduce the previously mentioned fear of not reaching results or more debates with the "we are not that bad" mentality.

For the digital tool, multiple evaluators commented on the axis of the diagram of results from the individual assessment as confusing as well as some of the presented information such as the shown level for each question. One improvement suggestion made for the digital tool was for it to collect the improvement suggestions as well to have it all in one place and not some in the digital tool and some on paper.

4.3.4 Assessment phase 4

For phase four of the assessment, it became apparent that the output of the assessment was unclear to the evaluators, with it so loosely defined in the assessment itself. One suggestion beyond just reporting the maturity level for the organization as a whole and the individual capability areas was to include a radar chart to visualize where the weaker and stronger points of the organization are.

4.3.5 Other comments and improvement suggestions

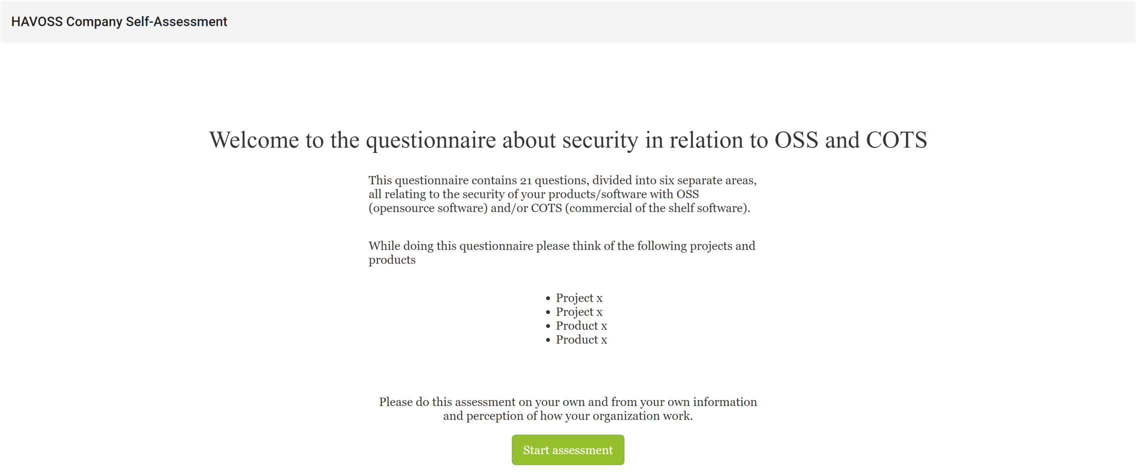
Several evaluators suggested having a short timespan for phase two and three would be a good idea, and likely raise the quality of the results and motivation of the organization to dedicate the time needed to the assessment. The suggestions given ranged from four to six weeks.

All evaluators thought the digital tool was a good idea, giving an easy way to complete both the individual assessment and workshop, even with some potential problems of visibility of the assessment in the individual phase. The fact that questions and capability areas could be skipped easily and naturally as well as the possibility to collect more data than just the answers to the questions were the biggest positives.

4.4 Updates to the assessment process

After the evaluations, it became apparent that step 1.4 of the assessment process, informing the participants, needed to be redefined and formalized. Step 1.4 is because of this turned into a formal startup meeting. Here every participant in the assessment should be present, if at all possible. The meeting is held by the sponsor of the assessment or the assessment team depending on if a sponsor exists. The meeting should start with an explanation why there is an interest in doing the assessment. This is followed with information from the assessment team about how the assessment is conducted, what dates are set, and a presentation of the purpose and objectives of the assessment. Ideally this is a short meeting to the point and allows for questions and discussions. The idea is to increase the motivation for the participants to increase the rate of participant along with the effort put into it. As the participants leave the meeting, they should all have access to the individual assessment and be prompted to do the assessment as soon as possible, preferably the same day.

At this stage, it is important to not give away the questions that the assessment. This is because it could start discussions that will later influence the results of the assessment to give the wrong image of the situation.



HAVOSS Company Self-Assessment

Welcome to the questionnaire about security in relation to OSS and COTS

This questionnaire contains 21 questions, divided into six separate areas, all relating to the security of your products/software with OSS (opensource software) and/or COTS (commercial of the shelf software).

While doing this questionnaire please think of the following projects and products

- Project x
- Project x
- Product x
- Product x

Please do this assessment on your own and from your own information and perception of how your organization work.

Figure 4.10: Welcome screen for the individual assessment

4.5 Updates to the digital assessment tool

Based on the feedback from the evaluation, the digital tool underwent a few major updates. These updates are presented in this section.

4.5.1 The individual assessment

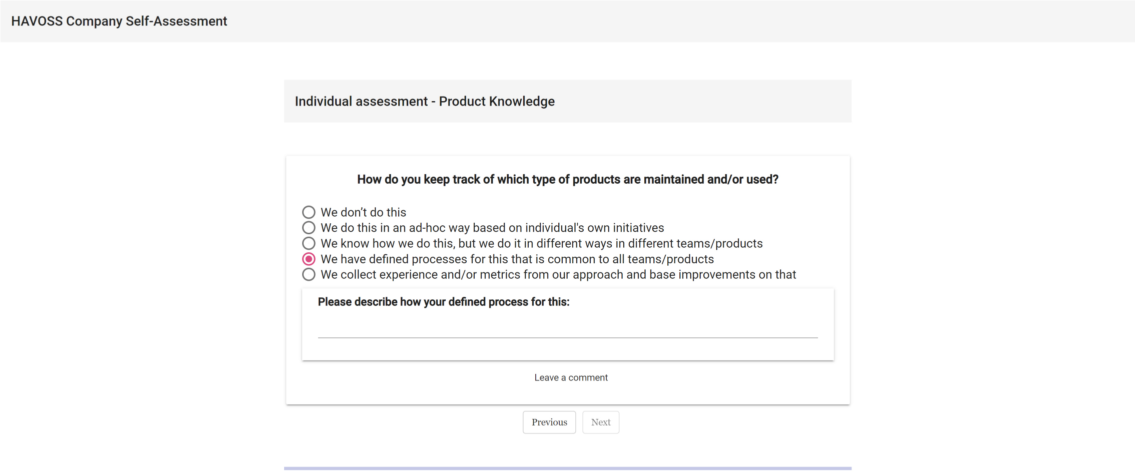
Based on the evaluation more information about the assessment process, and the individual assessment in particular the welcome screen for the individual assessment was updated. More information was added, also letting each the assessment team in each organization customize some information shown on the screen to help the participants in the evaluation. This update is illustrated in Figure **??**.

Based on a suggestion from the evaluation, whenever a participants answer either with “we have a defined process for this” or “we measure this”, there is now a requirement to enter more information about this is done in the form of free text. This must be entered before it is possible to move on to the next question, shown in . How this looks is shown in Figure **??**.

The view for when a participant wants to leave a comment changed to show the question, and the answer alternative at the same time that the leave a comment box is shown. How it looks in the updated version is shown in Figure **??**.

4.5.2 The workshop

The workshop view of any single question was changed around to be more readable and understandable. It starts with the question, then a new visible area containing the results from the individual assessment along any information left by participants as well as any comments written. Another distinct area below that one is the area for the workshop, where the answer the workshop wants to leave to the question exists along a textbox to enter improvement suggestion, shown in Figure **??**.



HAVOSS Company Self-Assessment

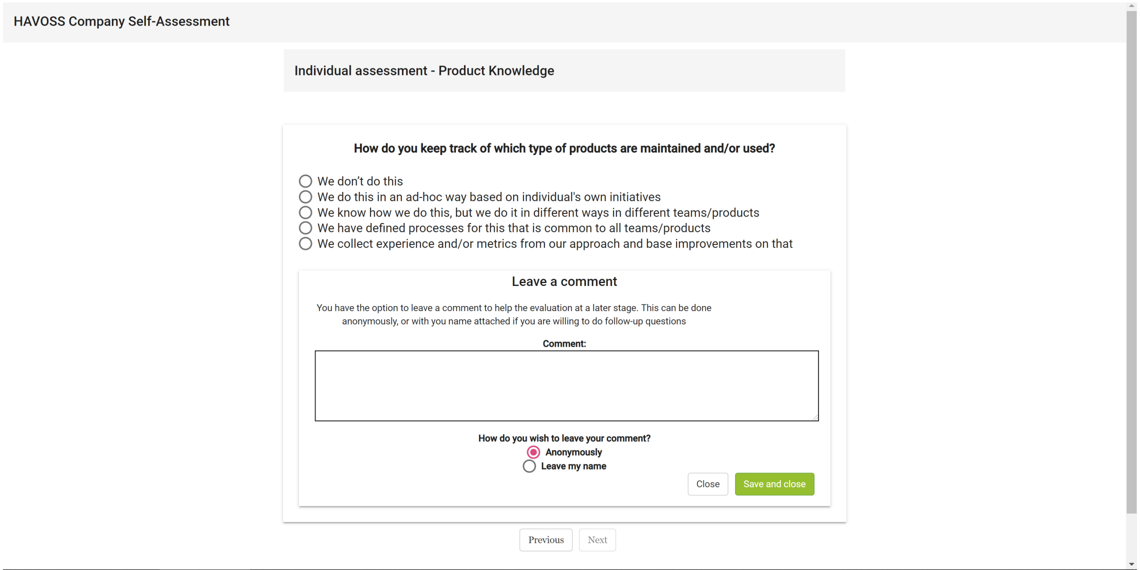
Individual assessment - Product Knowledge

How do you keep track of which type of products are maintained and/or used?

○ We don't do this
○ We do this in an ad-hoc way based on individual's own initiatives
○ We know how we do this, but we do it in different ways in different teams/products
● We have defined processes for this that is common to all teams/products
○ We collect experience and/or metrics from our approach and base improvements on that

Please describe how your defined process for this:

Figure 4.11: Please leave more information about your answer in the individual assessment



HAVOSS Company Self-Assessment

Individual assessment - Product Knowledge

How do you keep track of which type of products are maintained and/or used?

○ We don't do this
○ We do this in an ad-hoc way based on individual's own initiatives
○ We know how we do this, but we do it in different ways in different teams/products
○ We have defined processes for this that is common to all teams/products
○ We collect experience and/or metrics from our approach and base improvements on that

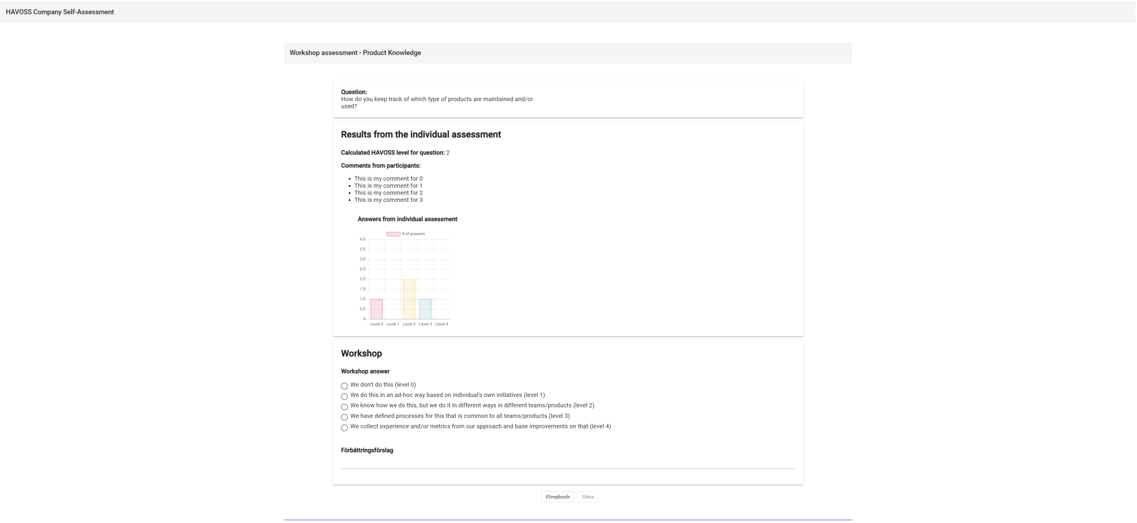
Leave a comment

You have the option to leave a comment to help the evaluation at a later stage. This can be done anonymously, or with you name attached if you are willing to do follow-up questions

Comment:

How do you wish to leave your comment?
● **Anonymously**
○ **Leave my name**

Figure 4.12: Leave a comment view in the individual assessment.



HAVOSS Company Self-Assessment

Workshop assessment - Product Knowledge

Question:
How do you keep track of which type of products are maintained and/or used?

Results from the individual assessment

Calculated HAVOSS level for question: 2

Comments from participants:

- This is my comment for 0
- This is my comment for 1
- This is my comment for 2
- This is my comment for 3

Answers from individual assessment

Workshop

Workshop answer

○ We don't do this (level 0)
○ We do this in an ad-hoc way based on individual's own initiatives (level 1)
○ We know how we do this, but we do it in different ways in different teams/products (level 2)
○ We have defined processes for this that is common to all teams/products (level 3)
○ We collect experience and/or metrics from our approach and base improvements on that (level 4)

Förbättringsförslag

Figure 4.13: A question view with comments in the MVP for the workshop.

Chapter 5
Analysis

The research question set up in Section ?? have been answered in Section ?? with the description of the assessment process. However, that is only one way of doing things, and with its own weaknesses and strengths. There certainly no single way of doing an assessment with the HAVOSS maturity model. In this section of the thesis, these strengths and weaknesses are discussed and argued for why they are left as is.

5.1 The process assessment methodology

The developed methodology puts a big emphasis of selecting the correct participants and creating an environment of high involvement. Even though the human aspect of the assessment process has been in the mind of the authors, the problems pointed out by Alberto Sampaio, Isabel B Sampaio, and Edwin Gray [?] are very much still a problem in the developed assessment method. This is because one goal is that as little time as possible should be spent on the assessment for organization as a whole. The use of a, in the human perspective, restrictive tool as it does not allow for the human perspective in the individual assessment does not help the human aspect in the assessment process. That is somewhat compensated for in the workshop where different perceptions of a process are discussed to reach a though of what the actual process is.

5.1.1 The preparations

All evaluators were in agreement that the selection of participants and communication with them about the assessment is a crucial part of the assessment. As such it is also important to the outcome of the assessment and its truthfulness. The authors agree with this, and because of that the created assessment process puts a lot of effort into identifying and selecting participants as well as communication with them.

As a first step in the communication with the participants, defining the purpose of the assessment gets major success factors for SPI in general and the assessment process in particular as discussed in Section **??**. The biggest factor being getting the goals of the assessment process and SPI aligned with the business needs, which in turn should get commitment from both management, however important or not that is, and employees in the organization as said in sections **??** and **??**. To get that commitment, the information about the assessment and its purpose needs to be of importance to everyone asked to be involved.

5.1.2 The individual assessment

The evaluators did not agree on if it was a good thing if the assessment was completed as quickly as possible or should take some time. The argument for it going quickly was for the participants to get it done quickly as a motivating factor to participate and reduce the commitment needed for individuals as well as the organization at large. This is seen as a valid point by the authors as time is most often very limited for any single employee in an organization and thus, there would be resistance to doing yet another thing that could be viewed as outside the employees' responsibilities in the organization. On the other hand, another evaluator points out the need for some afterthought while doing the individual assessment and reflection on what the actual processes in the organization are. This would greatly increase the time needed to complete the individual assessment but could, potentially, increase the correctness of the results of the individual assessment if everyone did this. It would however increase the time demand on the participants. As a compromise, the digital tool in the final version requires that additional information for the higher-level answers to the questions, as a way to give the participant a pause to think, but still keeping the time short for organizations, especially in the lower levels of maturity. An improvement to the digital tool would be to measure the time spent on individual questions and the assessment as a whole, and potentially put different importance to the answers depending on the time spent on the question.

5.1.3 The workshop

The motivation behind having a workshop was to complement the results from the individual assessment with a group evaluation of the company done with the help of the data from the individual assessment. The thinking being that the discussions that would happen in a workshop setting and the collectively larger knowledge of the organization would reach a maturity level closer to the actual process than the results from the individual assessment.

During the evaluations the evaluators pointed out some possible flaws in that thinking. One major risk identified was that the different department representatives would not want blame for a low maturity level, and thus hold a "it's not our fault, we are better than this" mentality. The discussion could then spiral out of control and not reach a productive result. Another point brought up was that if a few participants are more vocal than others their points of view would dominate the discussion and the decision making on what answer alternative to any question is the correct one. This would also reduce the idea of increasing the knowledge base for a result. To combat these problems, a strong workshop moderator would be needed to make sure discussions are held constructive and for everyone to be able to voice their opinion equally.

Another weakness with the workshop identified by the evaluators were that the time needed to answer all the questions might be larger than possible to fit into a single workshop. This is a very real possibility if a tight time schedule is not kept as suggested in the assessment process. Discussions could easily run long, again especially if feelings run high. The authors still think it is possible to complete the workshop in one sitting, getting through all the questions in a reasonable amount of time. But real setting testing would be needed to confirm it.

5.1.4 Finishing up the assessment

Phase 4 is vaguely described in this thesis by design. A final report from the assessment team is to be created, in some format. But how that should look and what it should contain would vary greatly depending on why the assessment is done in the first place. That is also why it is left with only the note that this needs to be done, not how.

The assumption however that this assessment most likely is not a one-time off thing is made, and that is why the self-assessment of the assessment team is there as a step, to improve the assessment the next time it is to be performed. This is to increase the effectiveness of the assessment team and develop the work with the assessment in the organization in the future.

5.1.5 Following best practices

The best practices that are defined by [?] are presented in general in Section **??** to **??**. In addition, all best practices are given a very short description along with the numbering in Appendix **??**. Table **??** shows the results of how the process assessment methodology developed in this thesis complies with these best practices.

The method best practices are not all implemented, in short because of how the methodology is designed, and collecting data from a few of the mentioned sources in the best practices simply is not done. The developed methodology is also highly targeted with little flexibility in how it is done, leaving another few practices to be abandoned, in favor of a very easy to follow procedure for the assessment.

All the supportive tool best practices are discussed in more depth in Section **??** below and thus left alone in this section. Procedure best practices are all followed except for one, PBP-5, which is to hold a feedback session after the assessment. This was skipped in favor of keeping the number of meetings where everyone needs to attend to a minimum and the startup meeting was deemed much more important than a follow-up one.

As for the documentation best practices, this thesis and especially section **??** is documentation of the assessment methodology. But no further guiding documents

Table 5.1: Compliance of best practices in [25] for the created assessment method

Best practice category	Fulfilled	Somewhat fulfilled	Not fulfilled
Method	MBP-5 MBP-8 MBP-10 MBP-11 MBP-12	MBP-3 MBP-6 MBP-13	MBP-1 MBP-2 MBP-4 MBP-7 MBP-9
Supportive-tool	SBP-1 SBP-2 SBP-5 SBP-6	SBP-3 SBP-4	
Procedure	PBP-1 PBP-2 PBP-3 PBP-4		PBP-5
Documentation	DBP-6	DBP-2 DBP-7	DBP-1 DBP-3 DBP-4 DBP-5 DPB-8
User	UBP-1 UBP-3 UBP-5 UBP-6		UBP-2 UBP-4

Table 5.2: Test of traceable best practices

Practice	Fulfilled	Traceability
MBP-1	No	
MBP-3	Partially	Done in part in step 3.2
MBP-5	Yes	Only option is to use the HAVOSS process model, thus completed by the process itself.
MBP-10	Yes	The HAVOSS model only consists of 21 questions and the individual assessment uses those and 5 more.

have been developed for the assessment except for what is built into the digital tool. A few of the practices could be fulfilled if the digital tool was extended. The practice pertaining to the assessment team, DBP-3, would be done in the prerequisites phase and thus not in the scope of this thesis.

Most of the user best practices are fulfilled. UBP-6 should be reached by a good definition of purpose and the startup meeting. The digital tool should also help with this, leaving the results open for the organization to view after the workshop is completed. It is questionable if UBP-1 is something that is reached with the defined process as the role of the sponsor and the selection of the assessment team is not discussed in this thesis and stuck in the prerequisite phase. This also means that UBP-2 is not reached. The responsibilities by the participants are clearly defined, depending on if they participant is going to partake in the workshop or not and that part of UBP-1 is reached. The following user best practices number three and four are not involved either as time and money is assumed to be provided for the assessment, and no further involvement from management as a group is necessary.

5.1.6 Data collection and reliability

For better or worse, the developed assessment process collects data in very few ways and does not support triangulation of information. The major part of data is collected during the individual assessment which is a questionnaire at the heart of it. The use of a questionnaire is normal for all studied assessment methodologies, including CMM and SPICE but always used in combination with other sources. Additional information sources could come up from the comments left by participants in the individual assessment, and then later investigated by the assessment team, to be used during the workshop. How large this work is depends on what comments are left and how diligent the assessment team is. During the workshop, another source of information will come in the form of verbal data from the participants but is not verified during the assessment process other than by the knowledge of the other participants in the workshop.

The other studied frameworks for process assessment methodology like SPICE focus heavily on interviews with people in the organization and studying documentation, along with questionnaires. This gives the opportunity to triangulate the information, that is get information from different independent sources and thus getting as close to the truth of the matter as possible. Even though the idea of interviews were carefully considered for a spot in the developed process, it was deemed as to work and time consuming to be included in the process when the objective was to have it as lightweight as possible. The idea is instead that with the help of wide participation in the individual assessment being able to collect enough data to provide reasonably reliable results for the workshop to work from.

One problem with using a digital questionnaire, and not personal, individual, interviews to collect data is the possibility of having someone "looking over" a participant shoulder when they are doing the individual assessment to get a view of how they think in the matter. This could then be used against the participant in question if it does not follow the onlooker's view or agenda. In an interview privacy and anonymity should be guaranteed with a good protocol, and in this case it would not be. Another problem would be if an individual with invested interest dictates to participants of the individual assessment how they should answer, or if a group does it together in a mini workshop. This would greatly reduce the reliability of the collected data and invalidate the results from the individual assessment. Again, this would be combatted with the more work intensive ways of collecting data. Preventing this can be done by having a very clear goal for the assessment that the organization and all the participants really buys into and there is no need to sabotage it. It would be important to combat any ideas of “the blame game” and territorial behavior of being “the best”.

In conclusion to the data collection and reliability questions, the aim of the developed assessment methodology is to have a process that is easy for an organization to follow and complete, with as little work as possible and get a good view of their organization. If a result with higher reliability is needed as a basis to further work in the organization, another type of assessment method is a better fit.

5.2 The digital tool

The digital tool could easily just be or become a tool used without adding any real value to the assessment process at large and the individuals involved in it. The evaluators were however in agreement that it would add value, making it easier for many parts of the assessment and hopefully increasing the engagement in the assessment process.

Does the tool reduce the time for an individual to do the individual assessment? This depends on how you look at the question and what the premises are for the individual in the individual assessment. There are two distinct cases with different answers.

In the case where the individual knows all the capability areas and thus answers all the questions in the questionnaire, no the digital tool does not have any apparent time saving qualities. The opposite is quite possible where the initial question for each capability area will take some extra time. It is also possible that for this type of individual, the fact that not all questions are viewed at the same time might slow them down when understanding the questions.

In the other case, where the individual only knows one or a few capability areas and thus only answers them, the digital tool will probably provide a quicker way to get through the individual assessment as not all questions are viewed. This should hold especially true for individuals in organizations with a low maturity level, due to the reduced need to take in extra information that the digital tool hides until it is relevant to the individual.

Does the tool reduce the time needed for the assessment team to complete the work of the individual assessment and prepare the workshop? The digital tool removes the need to input answers from a paper questionnaire that is the obvious, but not only, gain with the digital tool. The process of digitalizing the data from the questionnaire would take a lot of time for the assessment team and be prone to errors. The digital tool removes this totally, clearly reducing the time and work for the assessment team both with the individual assessment and the workshop as the data from the individual assessment is transferred over to the workshop part of the digital tool, analyzed and presented all automatically.

Does the tool fulfill the best practices defined for a supporting tool in [?]? Fulfilling the best practices defined in [?] depends on more than just the current implementation of the prototype. Many of them depends on how the supporting server software would be constructed and the communication between the user interface (UI) presented in the prototype and that server. The rest of this discussion will be based on the assumption that best practices would be used in implementing the required server software, data storage and communication between UI and server. Were it not for these assumptions, very little could be said at this time about the fulfillment of the best practices.

Of the six defined best practices, the first, SBP-1, is obviously supported at the get go, with both the individual assessment and the workshop supported. With the extension of the capability of generating reports at the end and adding a participation list in the planning phase all phases would be supported. SBP-2 is entirely based on how the server would be implemented but it is fair to assume that this would be supported by the tool and could even be used in the workshop with very little addition to the UI and prototype.

SBP-3 is, at the current state of the prototype questionable. If the end view of the workshop seen as a report of the assessment, which it could be and would provide useful data to the organization, then this is fulfilled. If on the other hand a more proper report is seen as required, then this would not be supported in the current version of the prototype. SBP-4 is also questionable, it is not bound to a timeframe at this time and as such support a change of the duration of the assessment. However, it does not provide any flexibility in how the assessment process is done.

SBP-5 again depends on the server and the communication between server and UI. But it is to be expected that it is kept private, and thus provides confidentiality for the individual assessment, if nobody is looking over a shoulder of a participant. The last one, SBP-6 is for sure fulfilled by a digital tool, as long as no data is tampered with after the fact. This would be combated with logs and security around the data.

Does the tool support the workshop with decisions and results of the workshop? This is one where the evaluators were the most in favor of the digital tool. With the digital tool, the view of the data for each question is standardized, and easy to learn and read once understood. The most likely alternative would be that the assessment team had summarized the data from the individual assessment in a PowerPoint presentation with much the same information for each question and shown on the screen. As with entering the data from the individual assessment manually, analyzing the data and creating this PowerPoint could easily introduce errors into the presentation. But this also shows that the digital tool is not needed but reduces the work for the assessment team in particular. It also puts the tool under great pressure that there is not a better way to present the data and information that the assessment team could come up with and reach better results.

In conclusion to this question, the answer is left as a maybe. It definitely has big advantages having the digital tool, but for the workshop attendees, other equally good ways could be used to do the work of the digital tool. The question remains if this is likely to be done or less information in a worse way would be presented.

Does the tool visualize the results of the assessment for the entire organization to use later as a foundation for the rest of the SPI process? In the current state of the prototype, this would be the end of the workshop view, which do present an overview of the data from the individual assessment and the workshop side by side. This is in the end a poor visualization of the results, with little visual impact and might be hard to understand without a lot of more context and understanding of the HAVOSS model. However, such a digital tool, if developed further has a great opportunity to really present the data from the entire assessment in an accessible way, with all the information needed to interpret the data for the organization or part of the organization that is wished.

5.2.1 Design decisions

Multiple design decisions had to be made during the development of the prototype. The most major ones are presented and discussed here as it could impact how the digital tool preforms but no tests on it were done.

For the individual assessment the major one would be if to show all the questions, be it the entire questionnaire or the questions of a capability area in one view, as done in most paper questionnaires, or each question individually. The latter was chosen, but no best practices were found on the subject, all suggestions being doing a test with all alternatives and see what preforms best.

For the workshop, how the individual questions were presented were majorly changed between the first version, presented in Section **??**, that was evaluated and the changes to the tool are presented in Section **??**. During the evaluation it became apparent that the information from the individual assessment needed to be clearly presented as such to the workshop and somehow separated from the work of the workshop, as done in the second version. This separation could be done multiple ways, such as showing the data from the individual assessment first and then hidden, to show the workshops work or as the second version, on top of each other, or it could have been side by side.

The chosen design has some flaws is deemed to have the possibility of impacting the workshop negatively, that in not the entire view of the question view can be seen at the same time on a laptop or being projected from one on a big screen. Different alternatives however were not seen as better by the authors of the thesis when privately tested during development. The need to scroll up and down could cause irritation in a workshop were each individual cannot see the view they wish to during discussions.

Chapter 6
Conclusions

Software process improvement probably exists in just about every software organization around. The name SPI has been around since the late 1980's in a format that is very similar to todays, but today it has gotten more refined and with a wider variety of frameworks to use as a guide on how to complete the process. The general SPI process can be seen as having four process steps that iterate over and over, where one of the steps is the software process assessment step which has been the focus of this thesis.

It was found that not all that much research has been done into how the software process assessment should be completed, with CMM just glossing over it. CMMI have its own process assessment SCAMPI, that is in itself well defined but hard to apply for an organization without special knowledge or taking in external consultants to do the evaluation.

Based on different approaches of software process assessment, and heavily inspired by SPICE, a process assessment methodology was developed for the HAVOSS maturity model. The aim of the assessment process was to be an internal assessment methodology and as lightweight as possible so that organizations of any size would be able to complete it with minimal expert knowledge. The developed method, presented in detail in Section **??**, consists of four distinct phases with one prerequisites phase, not included in that count. To help the assessment and reduce the required administrative effort for the assessment a minimal viable product for a digital assessment tool was developed. The digital tool was to support the process and collect the data for the assessment and then use it to generate the results from the assessment.

We found from an evaluation of the assessment methodology that the developed methodology was thought to work, and the tool a good addition even though some question marks remains, especially on if the workshop part would work as thought or need a redesign. The developed methodology takes into account many of the found best practices in [**?**] as well as success factors of SPI in general to hopefully generate results that align with what is needed for an organization to build on the results from the assessment. At the same time, some best practices were ignored during the work of this thesis or seen as outside of that same work, leaving the question if they would improve the assessment methodology.

6.1 Future work

There are plenty of more work that could be done based of this thesis, starting with testing of the developed assessment methodology. This should be done with and without the digital tool, giving the opportunity to verify to what degree the digital tool helps the process.

The digital tool in itself could also be developed and tested further, optimizing the user interface by doing user testing of different designs to gain more understanding and engagement from the participants would likely provide very valuable insights into how such a tool should be structured and what it can help with. Along the same path, testing to see what impact different visualization of data and the results of a process assessment done with the digital tool could also bring very interesting insights into how assessment work could be improved and used out in organizations.

It would also be interesting to see other uses of the HAVOSS model with more ridged assessment methodologies. Especially with more triangulation of information to increase the reliability of the results and have options for organizations to use from with it comes to methodology based on what the results of the software process assessment is going to be used for.

References

[1] Pekka Abrahamsson. Is management commitment a necessity after all in software process improvement? In *Proceedings of the 26th Euromicro Conference. EUROMICRO 2000. Informatics: Inventing the Future*, volume 2, pages 246–253, 2000.

[2] Dennis Ahern, Richard Turner, and Aaron Clouse. *CMMI Distilled: A Practical Introduction to Integrated Process Improvement.* Addison-Wesley, Boston, 2 edition, 2003.

[3] Henri Barki and Jon Hartwick. Measuring User Participation, User Involvement, and User Attitude. *MIS Quarterly*, 18(1):59, 1994.

[4] Jose A. Calvo-Manzano Villalón, Gonzalo Cuevas Agustín, Tomás San Feliu Gilabert, Antonio De Amescua Seco, Luis García Sánchez, and Manuel Pérez Cota. Experiences in the Application of Software Process Improvement in SMES. *Software Quality Journal*, 10(3):261, 2002.

[5] Fabiano Cattaneo, Alfonso Fuggetta, and Donatella Sciuto. Pursuing coherence in software process assessment and improvement. *Software Process: Improvement and Practice*, 6(1):3–22, 2001.

[6] CMMI Intitute. Standard CMMI Appraisal Method for Process Improvement (SCAMPI) Version 1.3b. Technical report, CMMI Insititue, Pittsburgh, 2014.

[7] Philip B Crosby. *Quality is free : the art of making quality certain.* McGraw-Hill, New York, 1979.

[8] Philip B Crosby. *Quality without tears : the art of hassle-free management.* McGraw-Hill, New York, 1985.

[9] Philip B Crosby. *Quality is still free : making quality certain in uncertain times.* McGraw-Hill, New York, 1996.

[10] Norman Dalkey and Olaf Helmer. An experimental application of the Delphi method to the use of experts. *Management science*, 1963.

[11] Gordon B Davis and Margrethe H Olson. *Management Information Systems: Conceptual Foundations, Structure and Development.* McGraw-Hill, New York, 2nd editio edition, 1985.

[12] Tore Dybå. An empirical investigation of the key factors for success in software process improvement. *IEEE Transactions on Software Engineering*, 31(5):410–424, 2005.

[13] William Edwards Deming. *Quality, productivity and competitive Position.* Massachusetts Institute of Technology, Cambridge, 1982.

[14] William Edwards Deming. *Out of the Crisis.* MIT Press, Cambridge, 1986.

[15] William Edwards Deming. *The new economics for industry, government, education.* Massachusetts Institute of Technology, Cambridge, 1993.

[16] Khaled El Emam, Jean-Normand Drouin, and Walcélio Melo. *SPICE: The theory and Practice of Software Process Improvement and Capability Determination.* Software Engineering Insitute, Washington DC, 1997.

[17] Emelie Engström, Margaret-Anne Storey, Per Runeson, Martin Höst, and Maria Teresa Baldassarre. How software engineering research aligns with design science: a review. *Empirical Software Engineering*, 2020.

[18] Tony Gorschek. *Requirements engineering supporting technical product management.* PhD thesis, Blekinge Institute of Technology, Karlskrona, 2006.

[19] Alan R Hevner, Salvatore T March, Jinsoo Park, and Sudha Ram. Design science in the information systems discipline: An introduction to the special issue on design science research. *MIS Quarterly: Management Information Systems*, 28(1):75–105, 2004.

[20] Watts S. Humphrey. Characterizing the Software Process: A Maturity Framework. *IEEE Software*, 5(2):73–79, 1988.

[21] Watts S. Humphrey. *Managing the Software Process.* Addison-Wesley, 1989.

[22] Watts S. Humphrey. *A discipline for software engineering.* Addison-Wesley, 1995.

[23] Robin Hunter and Richard H Thayer. *Software process improvement.* IEEE Computer Society, Los Alamitos Calif., 2001.

[24] Mark I Hwang and Ron G Thorn. The effect of user engagement on system success: A meta-analytical integration of research findings. *Information & Management*, 35(4):229–236, 1999.

[25] ISO. ISO/IEC 15504-1:2004, 2015.

[26] ISO/IEC. ISO/IEC 12207:1995 Information Technology – Software life cycle processes, 1995.

[27] ISO/IEC. ISO/IEC 15288:2002 Systems Engineering – System life cycle processes, 2002.

[28] Joseph M. Juran. *Juran on quality by design : the new steps for planning quality into goods and services*. Free Press, New York, 1992.

[29] Joseph M. Juran and Frank M Gryna. *Juran's quality control handbook*. McGraw-Hill, New York, 1988.

[30] Tom Knoll. *The think-aloud protocol*. Oxford University Press, Oxford, 2018.

[31] Elia Kouzari, Vassilis C. Gerogiannis, Ioannis Stamelos, and George Kakarontzas. Critical Success Factors and Barriers for Lightweight Software Process Improvement in Agile Development - A Literature Review. In *2015 10th International Joint Conference on Software Technologies (ICSOFT), Software Technologies (ICSOFT), 2015 10th International Joint Conference on VO - 1*, pages 151–159. SCITEPRESS, 2015.

[32] Jakob Nielsen. Thinking Aloud: The #1 Usability Tool, 2012.

[33] Pegah Nikbakht Bideh, Martin Höst, and Martin Hell. HAVOSS: A maturity model for handling vulnerabilities in third party OSS components. In *International Conference of Product Focused Software Development and Process Improvement (PROFES)*, Wolfsburg, Germany, 2018.

[34] Hanna Oktaba, Félix García, Mario Piattini, Francisco Ruiz, Francisco J. Pino, and Claudia Alquuicira. Software process improvement: The competisoft project. *Computer*, 40(10):21–28, 2007.

[35] Mark C. Paulk, Bill Curtis, Mary Beth Chrissis, and Charles V. Weber. Capability maturity model, version 1.1. *IEEE Software*, 10(4):18–27, 1993.

[36] Dewayne E. Perry, Lawrence G. Votta, and Nancy A. Staudenmayer. People, Oroganizations, and Process Improvement. *IEEE Software*, 11(4):36–45, 1994.

[37] Francisco J. Pino, Oscar Pedreira, Félix García, Miguel Rodríguez Luaces, and Mario Piattini. Using Scrum to guide the execution of software process improvement in small organizations. *Journal of Systems and Software*, 83(10):1662–1677, oct 2010.

[38] Ita Richardson and Christiane Gresse von Wangeheim. Why Are Small Software Organizations Different? *IEEE Software*, 24(1):18–22, jan 2007.

[39] Per Runeson and Martin Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14(2):131–164, 2009.

[40] Alberto Sampaio, Isabel B Sampaio, and Edwin Gray. The need of a person oriented approach to software process assessment. In *2013 6th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*, pages 145–148, 2013.

[41] Muhammad Sulayman, Emilia Mendes, Cathy Urquhart, Mehwish Riaz, and Ewan Tempero. Towards a theoretical framework of SPI success factors for small and medium web companies. *Information and Software Technology*, 56(7):807–820, jul 2014.

[42] SCAMPI Upgrade Team. Appraisal Requirements for CMMI Version 1.3 (ARC, V1.3). Technical report, Software Engineering Institute, 2011.

[43] Sidney G Winter. Organizing for continuous improvement: evolutionary theory meets the quality revolution. *Evolutionary dynamics of organizations*, 1994.

[44] Noppachai Wongsai, Veeraporn Siddoo, and Rattana Wetprasit. Factors of influence in software process improvement: An ISO/IEC 29110 for very-small entities, 2015.

[45] Pınar Yolum, Tunga Güngör, Fikret Gürgen, Can Özturan, Kival C Weber, Eratóstenes E R Araújo, Ana Regina C Rocha, Cristina A F Machado, Danilo Scalet, and Clênio F Salviano. Brazilian Software Process Reference Model and Assessment Method. In *Computer & Information Sciences - ISCIS 2005*, page 402. jan 2005.

[46] Mohammad Zarour, Alain Abran, Jean-Marc Desharnais, and Abdulrahman Alarifi. An investigation into the best practices for the successful design and implementation of lightweight software process assessment methods: A systematic literature review. *The Journal of Systems & Software*, 101:180–192, 2015.

Appendices

Appendix A
Best practices summary

This appendix contains an overview of the best practices found in [?] that was used in the work of this thesis. All the best practices are included with a short description of them and the percentage frequency they were found in [?] out of the 29 papers they studied in their study.

Table A.1: Method best practices

Number	Practice	Frequency
MBP-1	Collect data from interviews	10.3
MBP-2	Collect data from documents	6.9
MBP-3	Check the accuracy of the assessment's findings(data collected)	6.9
MBP-4	Provide flexible and customizable method focusing on principal high-priority processes	34.5
MBP-5	Identify the process reference model used to select processes	3.4
MBP-6	Identify strengths, weaknesses, improvement opportunities and threats	10.3
MBP-7	Suggest a feasible improvement action plan ,which addresses the special needs of the company	10.3
MBP-8	Provide a usable assessment method for on-site assessment and self-assessment	6.9
MBP-9	Provide compliance with a formal assessment method	6.9
MBP-10	Build a simple, well-structured questionnaire with no more than 150 questions	27.6
MBP-11	Design the assessment to last for a reasonable length of time	10.3
MBP-12	Ensure the reliability of the assessment result	17.2
MBP-13	Ensure completeness	3.4

Table A.2: The supportive tools best practices

Number	Practice	Frequency
SBP-1	Support various assessment phases	37.9
SBP-2	Build and use a database of historical SPA data	13.8
SBP-3	Generate assessment reports automatically	3.4
SBP-4	Provide a flexible support tool	6.9
SBP-5	Maintain assessment confidentiality	6.9
SBP-6	Ensure repeatability of the results	3.4

Table A.3: Procedure best practices

Number	Practice	Frequency
PBP-1	Prepare the assessment process	37.9
PBP-2	Build confidence and trust relationships with participants	13.8
PBP-3	Produce an assessment report to be delivered to the organization	10.3
PBP-4	Ensure confidentiality	20.7
PBP-5	Hold a feedback session after each assessment	20.7

Table A.4: Documentation best practices

Number	Practice	Frequency
DBP-1	Provide guidance for identifying assessment purpose, objectives, and logistics	3.4
DBP-2	Provide guidance for identifying an organizational unit	3.4
DBP-3	Provide guidance for the assessment team	6.9
DBP-4	Provide guidance for ensuring confidentiality	6.9
DBP-5	Provide document templates	13.8
DBP-6	Provide guidance to document the assessment method and its implementation in practice	31.0
DBP-7	Provide guidance to document data collection and rating results	10.3
DBP-8	Provide guidance for the follow-up assessors	3.4

Table A.5: User best practices

Number	Practice	Frequency
UBP-1	Define assessment participant responsibilities	24.1
UBP-2	Define assessment team credentials and responsibilities	13.8
UBP-3	Ensure the involvement of senior management and other staff members	17.2
UBP-4	Ensure sponsors' commitment	24.1
UBP-5	Ensure that participants feel the benefits of the assessment	13.8
UBP-6	Improve the credibility of both sponsors and staff, who should believe that the assessment would yield a result	13.8

Appendix B
Interview questions

In this appendix, the interview questions used during the evaluations are presented. Each subsection contains the different parts of the interview that were completed at different points in the interview.

B.1 Interview questions for first two phases

1. From your experience, is anything missing in setting up for an assessment process?

2. What do you believe the most important part of the preparations will be?

3. From your experience, do you believe that all relevant participants would be identified with the questions presented in step 1.2?

4. From your view, is there anything that would be hard to do in reality?

5. Anything you would add, change, or improve in the presented method?

B.2 Interview questions after individual assessment tested

1. Was there something during the individual assessment you couldn't understand?

2. Was there something you particularly liked during the individual assessment?

3. Was there something you didn't like during the individual assessment?

4. What problems can you see with the individual assessment in its current form?
5. Is there anything you would improve or change in the individual assessment?

B.3 Interview questions after workshop test

1. What did you like about the workshop as a format for the assessment?
2. What problems do you see with having the workshop with its current format and goal?
3. Do you think the digital tool will help during a workshop? Why?
4. Do you have anything to comment on the digital tool?

B.4 Closing interview questions

1. Is there anything you would add to the last phase?
2. As a whole assessment methodology, do you think it will work to produce a reliable result?
3. As a whole, do you see any problems with the methodology?
4. What is the weakest spot as you see it, for the methodology?
5. What is the strongest point as you see it for the methodology?

INSTITUTIONEN FÖR DATAVETENSKAP | LUNDS TEKNISKA HÖGSKOLA | PRESENTERAD 2020-06-11

EXAMENSARBETE An internal assessment method for the HAVOSS maturity model
STUDENT Johanna Hultén
HANDLEDARE Martin Höst (LTH)
EXAMINATOR Emelie Engström (LTH)

En metod för tillverkarna av smarta hem produkter att utvärdera säkerheten i programvaran

POPULÄRVETENSKAPLIG SAMMANFATTNING **Johanna Hultén**

Fler och fler produkter så kallade “smarta produkter” flyttar in i våra hem. Smarta produkter är generellt uppkopplade mot Internet vilket också skapar nya möjligheter för hackare att ta sig in i våra liv. Vi har studerat hur tillverkare av smarta produkter kan utvärdera säkerheten i programvaran i produkterna.

I ett samhälle där i princip alla äger en smartphone vill fler och fler kunna kontrollera mer och mer av sin vardag från telefonen. Ett vanligt exempel på smarta produkter idag är lampor som går att styra från telefonen. Den ökande efterfrågan på smarta produkter gör också att det finns allt fler olika typer av smarta produkter att köpa, från både små och stora företag, exempelvis Ikea.

Smarta produkter skapar nya säkerhetsrisker, exempelvis skulle en hackare kunna se videon från en smart kamera i barnens rum. Därför är det viktigt att företagen som utvecklar smarta produkter jobbar aktivt med säkerheten. Idag är det vanligt att använda kod i sina produkter som finns tillgängligt för alla på Internet. Det medför möjligheter men också problem och risker.

För att hjälpa företag utvärdera hur de väl de arbetar med säkerheten i koden som tas från Internet till smarta produkter utvecklade tre forskare på Lund Tekniska Högskola en modell som kallas HAVOSS. Den innehåller 21 punkter att titta på, alla relaterade till säkerheten. I detta examensarbete skapade vi en metod för att använda HAVOSS modellen för att företag så enkelt som möjligt kunna utvärdera sin verksamhet.

Resultatet av examensarbetet är en metodik uppdelat på fyra separata steg. Steg ett är förberedelser som görs av de personers om organiserar utvärderingen. Steg två är en enkät där de 21 punkterna från HAVOSS modellen uttrycks i form av varsin fråga. Enkäten görs elektroniskt av alla inom företaget som tros kunna bidra med relevant kunskap. Steg tre är ett gruppmöte med en mindre grupp från företaget som får ta del av resultaten från den tidigare enkäten. Gruppen ska komma fram till gemensamma svar för samma frågor som i enkäten. Det sista steget är för gruppen som organiserar utvärderingen. De ska skriva en rapport innehållande resultatet från den gjorda utvärderingen.

Som hjälp för att genomföra utvärderingen skapades också en prototyp av ett digitalt verktyg. I det digitala verktyget genomförs enkäten i steg två. För steg tre finns det stöd i prototypen för att visa upp frågorna, resultateten från enkäteten och att sammanställa och visualisera resultaten från gruppmötet.

Fem personer fick under arbetet utvärdera den framtagna metoden och det digitala verktyget. Utifrån det som framkom under utvärderingarna förändrades processen och det digitala verktyget till att fungera ännu bättre.