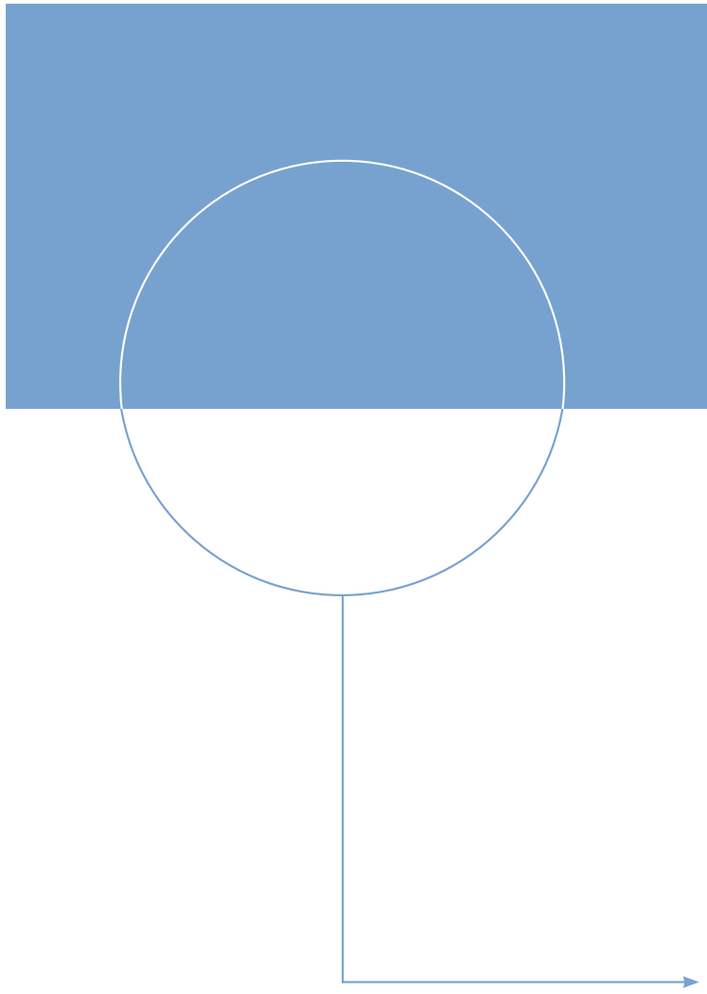


Master's thesis

2019

Vilde Innset Hurum

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Information Security and Communication Technology

Vilde Innset Hurum

Implementation of Data Loss Prevention Mechanisms in a Knowledge Organization: A Balance Between User Experience and Security

May 2019

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of Science and Technology

NTNU

Norwegian University of
Science and Technology

Implementation of Data Loss Prevention Mechanisms in a Knowledge Organization: A Balance Between User Experience and Security

Communication Technology
Submission date: May 2019
Supervisor: Maria Bartnes
Co-supervisor: Roy Thomas Selbæk Myhre

Norwegian University of Science and Technology
Department of Information Security and Communication Technology

Problem description:

In 2014, half of the data loss incidents were discovered in the business sector. As employees create and distribute documents and emails on a daily basis, there is a risk of data loss incidents. It is therefore important that organizations have security procedures and classification schemes in place to minimize the risk.

A Data Loss Prevention (DLP) solution can be implemented to protect an organization's critical data, such as intellectual property, records about the employees and customer data. The solution prevents information loss by utilizing a number of techniques. These techniques include classification of information, policy enforcement for information processing and data transfer, monitoring of the information's location and detection of whether information is sent to unauthorized users. It can serve as an assisting device to help the employees classify and handle data correctly. However, employees might perceive a DLP solution as a barrier since it may involve additional work. It is therefore crucial that the user experience is seamless and intuitive and also that the solution is used properly.

This master thesis explores how a DLP solution can be implemented in an organization with focus on user experience and the effect on the employees' classification routines. Through interviews, usability testing and questionnaire, potential barriers and desirable features in a DLP solution is investigated in order to ensure a good user experience. In addition, the effect the solution will have on the employees' classification routines is estimated as some of the features are user-driven. The case organization is a knowledge organization consisting of 2000 employees. The organization has decided to implement the cloud-based DLP solution Azure Information Protection (AIP).

The main tasks in the thesis include:

- Discuss the importance of user experience when implementing a DLP solution in an organization with focus on potential barriers in the technology and why it is crucial that employees are aware of the consequences of classifying documents and use the solution correctly.

- Plan and conduct a study in order to investigate to what extent the employees' classification routines in the organization can be affected by implementing a DLP solution and identify potential barriers in the solution and find out how to minimize these.

- Qualitative and quantitative analysis of semi-structured interviews, usability tests and questionnaire.

Abstract

Organizations often have policies regarding how to protect and classify their data. Despite this there are, unfortunately, many data loss incidents happening in the business sector. There is also a risk that employees are not aware of the policies or that the policies are not practiced correctly. To ensure that documents and email are protected, a Data Loss Prevention (DLP) solution can be implemented. However, it is crucial that employees accept the solution, use it properly and do not perceive it as a barrier in their daily work life. This master thesis project aims to explore how a DLP solution can be implemented in a knowledge organization with focus on user experience and how it affects the employees' classification routines.

Limited research has been conducted on the balance between security and user experience with DLP solutions, and it is therefore highly relevant. In this master thesis project, usability tests of the DLP solution Azure Information Protection (AIP) was conducted, together with interviews, in order to explore a proper balance between security and user experience. This was followed by a survey in order to gather information that could provide more insight. The case organization is a knowledge organization consisting of 2000 employees.

The findings indicate that it is important that the employees perceive the features as useful, and that they retain control over the classification functions and understand how they work. Furthermore, potential barriers should be introduced with care to ensure they do not interrupt the employees' workflow. It was found that barriers are only acceptable in cases where the value of the function is clearly recognized. The results also showed that employees with management roles are only slightly more aware of classification than other employees. Despite an existing security culture, it was revealed that the solution will both enable employees to more easily practice the organization's classification policy, improve classification routines and help better protect customer data. In addition, the DLP solution will make them more aware of the organization's classification policies. However, a successful implementation of a DLP solution demands for actions by the organization, such as providing information, enforcement and tutorials.

Sammendrag

De fleste organisasjoner har retningslinjer for hvordan man skal beskytte og klassifisere data. Til tross for dette, skjer det dessverre mange hendelser hvor data går tapt eller lekkes i næringslivet. Det er også en risiko for at ansatte ikke er bevisste på organisasjonens retningslinjer eller at de ikke blir fulgt på riktig måte. For å sikre at dokumenter og e-post er beskyttet kan en Data Loss Prevention (DLP)-løsning implementeres. Samtidig er det viktig at de ansatte aksepterer løsningen, bruker den riktig og ikke oppfatter den som en barriere i sitt daglige arbeid. Denne masteroppgaven undersøker hvordan en DLP-løsning kan implementeres i en kunnskapsorganisasjon med fokus på brukeropplevelse og hvordan den kan påvirke de ansattes klassifiseringsrutiner.

Det er begrenset med forskning på balansen mellom sikkerhet og brukeropplevelse i DLP løsninger, og det er derfor et relevant tema. Denne masteroppgaven gjennomførte brukertesting av DLP-løsningen Azure Information Protection (AIP), i tillegg til intervjuer, for å finne en akseptabel balanse mellom sikkerhet og brukeropplevelse. Dette ble etterfulgt av en spørreundersøkelse for a samle inn informasjon som kunne gi mer innsikt. Organisasjonen som ble studert er en kunnskapsorganisasjon med 2000 ansatte.

Resultatene indikerer at det er viktig at de ansatte opplever DLP funksjonene som nyttige, og at de har kontroll på dem og forstår hvordan de fungerer. I tillegg bør man være bevisst på å introdusere potensielle barrierer i løsningen for å sikre at de ikke forstyrrer de ansattes arbeidsflyt. Det ble i tillegg funnet at barrierer kun er akseptable når funksjonens verdi er tydelig. Resultatene indikerer også at ansatte med lederansvar er noe mer bevisste på klassifisering enn andre ansatte. Til tross for en eksisterende sikkerhetskultur, viser det seg at løsningen vil gjøre det lettere for de ansatte å følge organisasjonens klassifiseringspolicy, forbedre klassifiseringsrutinene og beskytte kundedata bedre. I tillegg vil DLP-løsningen gjøre dem mer bevisste på organisasjonens klassifiseringspolicier. For å lykkes i implementeringen av en DLP-løsning kreves det også tiltak fra organisasjonen, for eksempel i form av informasjon, håndhevelse og opplæringsprogrammer.

Acknowledgement

First, I would like to thank my supervisors Maria Bartnes and Roy Thomas Selbæk Myhre for their time and guidance throughout this process. I would also like to thank my father for proofreading my master thesis.

Contents

List of Figures

List of Tables

List of Acronyms

AIP Azure Information Protection.

ARM Azure Rights Management.

DLP Data Loss Prevention.

HCI Human Computer Interaction.

HCI-SEC Human Computer Interaction - Security.

IDS Intrusion Detection System.

IS Information System.

ISO International Organization for Standardization.

IT Information Technology.

NIST The National Institute of Standards and Technology.

NSD Norwegian Center for Research Data.

TAM Technology Acceptance Model.

VPN Virtual Private Network.

Chapter 1
Introduction

Like everything else in our world, the workplace has become digital. Employees are introduced to new technologies in their daily work practices and data and information are digitally distributed and shared with co-workers and customers. In addition, most organizations use both internal and external electronic communication channels and allow employees to connect their personal mobile devices to company networks. The digital workplace revolution, despite all its benefits, introduces vulnerabilities that can be exploited by malicious actors. Businesses and organizations are potential targets for industrial espionage with intent to steal trade secrets and critical information. Consequently, there is a great need for protection of an organization's sensitive information, such as intellectual property, customer data and personal records about employees.

About half of all data loss incidents registered in 2014 was discovered in the business sector [ASM16]. For an organization, the consequences of data loss can be crucial, as they may risk reputation, capital and competitiveness [LK10]. To prevent data loss, companies and organizations must establish policies and procedures regarding access to information and data. Furthermore, the employees must contribute to the implementation in order to ensure the best possible protection. It is not possible to achieve complete protection of sensitive information in an organization just by securing components, networks and clients. Even though there exists many detection and prevention schemes, such as Intrusion Detection System (IDS), firewall, and Virtual Private Network (VPN), these are only effective in cases where rules are well defined [PS17]. However, as employees make use of various accessible communication channels, such as email and instant messaging, these rules can easily be violated. Human error is a common cause of unintentional data loss in an organization. Thus, securing against human factors, such as how employees use, store, and send information and their information security behaviour in general, is equally important to include in security management. Implementing a Data Loss Prevention (DLP) solution can help reducing the risks associated with human factors. DLP is a technical security measure that enforces policies for information processing

and transfer and supports classification of information, protecting and monitoring of critical and sensitive data [LK10]. In addition, it can be used to increase the awareness among users in an organization about classification and processing of information that needs a certain level of protection.

1.1 Motivation

Organizations often have established procedures and classification schemes to enforce what sensitivity level and access permission to be applied to certain documents. By implementing a DLP solution, employees are allowed to classify and control corporate documents and monitor them through the network. The solution provides flexibility for the user and its actions since it includes user-driver classification features [Cha11]. For instance, when an internal email is about to be sent, the user can select between removing external recipients or delete sensitive information from the text or attached documents. According to security specialist Graham Titterington [Cha11], a DLP solution could thus educate the employees about the importance and practice of information security. The solution will influence the employees' classification routines and work as a policy reminder [Cha11]. By improving the employees' security practices, the security of the whole organization will benefit as well.

Since the employees have the responsibility to apply the correct classification level according to the organization or customer's classification schemes, there is a risk of false negatives and false positives. False negatives occur when sensitive data is marked as non-sensitive, while false positives occur when non-sensitive data is marked as sensitive [Marnd]. For instance, to be sure not to leak any information, an employee may classify most documents as "confidential". Such false positives can make it difficult to manage the document later as a result of the unnecessary restrictions and the high cost. In the case of false negatives, there is a possibility for compromise or loss of data. Thus, it is important that the correct classification level is applied to get the correct restriction and avoid undesirable events. If used correctly, a DLP solution can contribute as an assisting device to ensure that policies are maintained.

A challenge when implementing complex procedures involving several steps, such as in a DLP solution, is that it might prevent employees from doing their daily work efficiently [LK10]. If technical solutions are perceived as barriers or practiced incorrectly violations may occur. For example, if an employee feels that sending a confidential email in a DLP solution requires execution of many steps, he or she might be tempted to send the email from their personal email account. It is therefore important to find a balance between security and user experience when selecting and implementing a DLP solution. Thus, it is interesting to investigate further what factors that influence employees' perception of barriers.

Research on data loss prevention techniques is increasing, but there is lack of research on the detection of data loss from a user behaviour perspective [PS17]. Previous studies have focused on different factors, for instance how personality and the social environment affects the security practices of employees. One study explored the effect on employees' security behaviour based on who had read and who had not read the security policies in an organization [Vei16]. In a study related to technology adoption, the majority who said they would use the security program turned out in later time not to do so [SWS15]. Thus, it is clear that certain barriers exist. This research focuses on the balance between security and user experience when introducing a DLP solution in an organizational context and also on how the implementation may affect the employees' classification routines. There is a lack of studies in this area and by conducting this study, the results can be useful for a wide range of organizations enforcing classification schemes and implementing Data Loss Prevention solutions.

1.2 Research questions

To narrow down the scope, this master thesis project explores the balance between security and user experience when introducing a DLP solution in an organizational context and the predicted effect on the employees' classification routines. Based on the challenges addressed above, the following research questions will be investigated:

- **RQ1:** To what extent can DLP features be introduced before they are perceived as barriers and reduce the user experience?

- **RQ2:** How does a DLP solution affect the employees' classification routines?

The hypotheses H1-H3 will be investigated to support the exploration of RQ2. H1 and H2 were established to check if the implementation of the DLP solution has an effect on the employees' awareness of the organization's classification policies and potential false positives.

- **H1:** Employees tend to classify documents as confidential by default, which sometimes might result in information being stricter classified than required.

- **H2:** A DLP solution will make employees more aware of an organization's classification policies.

The results from the research conducted by Stanton et al. [SMSJ04] reveal that the employees' security behavior is affected by their manager's security behavior. If this is the case, initiatives targeting classification routines of managers may be the most effective. H3 is formulated to investigate this.

- **H3:** Employees with management roles are more aware of classification than other employees.

1.3 Outline

In Chapter 2 relevant literature and related research studies are presented, while the research method is described in Chapter 3. Chapter 4 presents features of the DLP solution AIP that was investigated. Then, the results are presented in Chapter 5 and further discussed in Chapter 6. Finally, Chapter 7 concludes the research.

Chapter 2
Background and Related Work

This chapter presents relevant terms and definitions and a literature review of related research in sections where appropriate.

2.1 Data loss prevention

A DLP solution is concerned with protecting information, while other security measures aim to protect networks, communications, etc. Even though organizations have implemented security measures, such as firewalls, antivirus and encryption, data leakages may occur. A DLP solution differs from other data protection technologies in that it can detect and prevent unauthorized users from accessing certain data and protect sensitive data from being shared accidentally [TS14]. While other technologies focus on protection of data from access by outsiders, DLP focuses on protection of intentional and unintentional data leakages both within the organization and to the outside world [TS14]. However, it is important to have security in several layers and a DLP solution is a contribution to this. Tahboub et al.[TS14] states that DLP solutions have a centralized approach compared to other security solutions that have an ad-hoc approach.

DLP is defined as the process of monitoring and protecting content from misuse [PS17]. The goal is to protect data at rest, data at the endpoint and data in motion and thereby maintain the confidentiality of the data [LK10].

- *Data at rest:* data that resides in file systems and databases.
- *Data at the endpoint:* data on laptops and external drives.
- *Data in motion:* data that moves through and outside the corporate network. Examples are emails and instant messages.

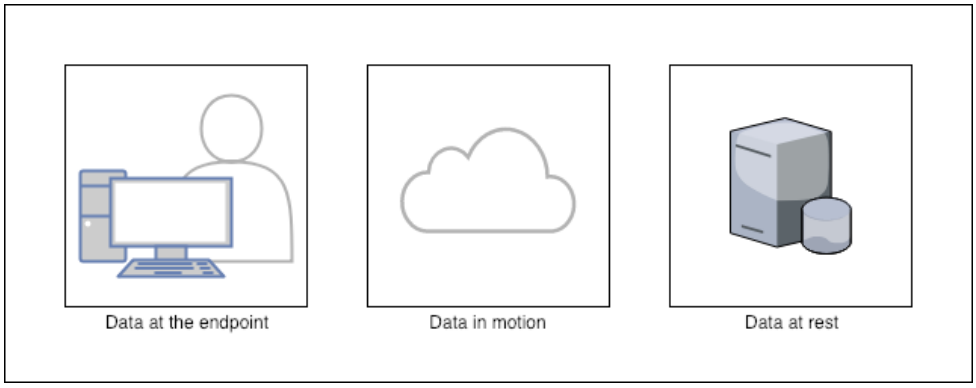


Figure 2.1: Data forms in DLP.

Data at rest is often protected with strict access privileges and encryption, while data at the endpoint is often not encrypted and can therefore be vulnerable. Since people make decisions regarding data distribution at the endpoint, this area is the most exposed to data leakages [ASM16]. There are three main phases in the DLP process; the data collection phase, the analysis phase and the remedial action phase [PS17]. Data from Internet and intranet logs are collected and further investigated in the data analysis phase. Here, a rule matching and policy, content and context verification process is performed. As a result of this process, the correct action is applied. Typical actions are blocking, alerting or allowing the user to perform further actions based on the security policy [PS17].

In the market today, there are several vendors offering Data Loss Prevention solutions. McAfee, Symantec, Trend Micro, Microsoft and Forcepoint are popular suppliers. Even though their solutions all offer data loss prevention, detection and monitoring features, there exists differences when it comes to what analysis techniques are used and also which remedial actions are included [ASM16].

As with most Information Technology (IT) solutions there are challenges with DLP. One of the challenges is concerned with reading encrypted data and data hidden in images, videos and audio [TS14]. Encryption may prevent authorized users from getting access to the data, but applying this technique also makes it difficult to analyze by a DLP solution [PS17]. Another challenge is that the solution only protects data in known channels. If data is sent from applications that are not included in the DLP solution, protection becomes difficult. In addition, proper policies and appropriate user access rights must be in place for the system to operate optimally. If this is not done carefully, inaccuracies may appear. More and more businesses and organizations are moving their data assets to the cloud. As a consequence, data

protection becomes even more important and a DLP solution that supports cloud storage is an essential part of this. All of these aspects must be taken into account when selecting and implementing a DLP solution.

2.2 Classification schemes and access control

Most organizations make use of some form of classification schemes to protect their data. Classification of documents has historically been used for a long time and started in military and governmental institutions [LHM01]. Shaika et al. [SS15] defines *Data Classification* as:

> *The process of defining various data levels and deciding a level of sensitivity to it.*

How crucial it is to protect the data depends on the business and service delivery models [SS15]. The data classifications reflect to what extent data must be protected and its level of importance in the organization. How the data is classified depends on different aspects, such as risk associated with disclosure of the data. In order words, the scheme consists of security levels customized for the organization, ranging from the most sensitive level, for instance "top secret", to the least sensitive level, "unclassified". By applying the scheme to documents it is clear what data is confidential and what is not [ASM16]. Access rights are crucial in a DLP solution as correct admission rights ensure that data is not lost or compromised [ASM16]. How the DLP solution examined in this research manages classification labels and access right is described in Chapter 4.

2.3 The human factor and user experience

A successful implementation and introduction of a new software system in an organization depends on a number of factors. Among these are individual characteristics, such as previous experience, knowledge and the degree of involvement in the implementation [Lec15]. According to the paper written by Dourish et al. [DGDdlFJ04] it is crucial that the user understands the solution in order to have the maximum effect of it. If the employees do not understand, accept or use the solution, it is waste of resources to implement it.

The human factor is essential in a successful DLP solution implementation. Not least when it comes to document classification because it is the end user's responsibility to apply the correct sensitivity label to a document. This implies that the user must know the organization's classification policies and identify the data in the document correct in order to label accordingly. In many cases this can be a

challenge. Even though humans make mistakes, analyst and security expert, Graham Titterington, states that the user-driven classification of documents most likely will have a notable positive effect [Cha11]. He also states that if classification is done by an automated system rather than by the user who wrote the document, the chance is bigger that wrong labelling may occur.

There are several definitions of user experience. As a good user experience depends on the individual's perception of usefulness, functionality and efficiency, it can be hard to define [Kun03]. The author Elizabeth Rosenzweig [NM18] states that:

> *The goal of user experience is to design products that are less prone to human error.*

Another example is Nguyen et al. [NM18] who states that a good user experience involves an interface that is simple to navigate and that one should be able to operate without being concerned about potential threats. Mistakes and misunderstandings can be a result of limited experience or performing actions that are unintended [NM18]. International Organization for Standardization (ISO) [ISO18] defines user experience as:

> *User's perceptions and responses that result from the use and/or anticipated use of a system, product or service.*

The ISO definition is used as the main reference in this research.

A common issue with a software solution that may result in a reduced user experience is if it requires the users to perform additional tasks compared to their normal activities. Thus, an important challenge when implementing a DLP solution is the additional security mechanisms introduced and imposed on the user [DGDdlFJ04]. If the new features are perceived as barriers the user might circumvent them, for instance by sending a screenshot of a confidential document and send it to users with insufficient access privileges [ASM16].

Some strengths related to user experience that characterizes a good DLP solution were addressed in the Gartner report [RK17]. Among these were the ability to configure the classification levels according to the organization's policy, monitoring and reporting, intuitive navigation within the solution, and clear identification of where and what rule is applied.

To examine and assess which solution that will be suitable for a given organization, there is a need to establish a responsible committee [RM10]. The committee must

create a plan on how to implement the solution and inform the organization about the process. As the joint venture Websense [RM10] states: *The biggest mistake is not to prepare the organization.* Internal testing is important in order to accomplish this.

2.4 Information security awareness in organizations

Organizations want to protect their information's confidentiality, integrity and availability. As humans are regarded the weakest link when it comes to securing systems and networks, the attention regarding information security improvements should be on them [WH03]. According to Niekerk et al. [NS05] education of employees is a key factor for establishing a security culture. However, in order to change the culture, the employees have to understand why the current solution is not good enough. The National Institute of Standards and Technology (NIST) [WH03] states that an awareness and training program is essential for the employees to understand their responsibilities related to IT security in the organization, and how to use and protect IT resources. NIST also emphasizes that CIOs, program officials and IT security program managers should act as promoters to enable continuous improvement and that they by doing this are success indicators of the program.

There exist different definitions of the term Information Security Awareness. One definition by Wolf et al. [WHP11] is that security awareness in general consists of two equally important parts; knowledge of policies and potential threats and how to practice the policies. Each part is inefficient without the other [WHP11]. According to Furnell et al. [FGD02], employees might be aware that there exist risks but not what the risks are. The paper also points out that the most common reasons for lack of security awareness in organizations are limited security expertise and financial resources to train staff, lack of knowledge of potential risks or more focus on other business priorities.

The DLP solution may contribute to an active policy persevering process as the user-driven classification can work as a reminder for the employees. If the employees are aware of the consequences of classifying wrong, the potential perceived barriers in the solution might be ignored.

Related research: There have been conducted several studies regarding which factors affect the security culture in an organization and the employees' security behaviour. The studies were based on both the employees and managers awareness.

Safa et. al [SSS+15] base their study on a model to minimize risks related to users' behaviour in organizations. Results show that threat, subjective norms and awareness have a positive impact on the security behavior. However, the findings revealed that the users' perceptions of control did not correspond with how they

actually behave. This finding is important to keep in mind when analyzing the results in this master thesis project.

Veiga [Vei16] conducted a study on how reading the organization's security policies affect the employees' security behaviour. The study concludes that reading the policy had a positive effect on the security culture within the organization. The author suggests that in order to minimize risks, incidents and error related to human factors an organization should ensure that the information security policy is read by all the employees. One of the goals in this master thesis project is to also investigate how using the DLP solution affects the awareness of organizations classification policies.

Stanton et al. [SMSJ04] find that security behaviour in an organization is affected by a number of factors. These include job role, job satisfaction and organizational commitment. In addition, the organization type has an effect. For instance, there is more daily focus on information security in military and financial institutions. Hence, in this master thesis project the results are discussed with these factors in mind in order to reveal to what extent they apply to organizations similar to the one examined in this master thesis project.

Other factors that have been investigated in previous research is how managers affect the security behaviour in the organization. A study by Knapp et al. [KMRF06] finds that the top management support has a positive impact on both the security policy in the organization and how the organization's security culture. This is also supported in a study by Chan et al. [CWK05]. Furthermore, the study finds that co-worker socialization has a positive impact. In addition, Strand [Str18], in a research study conducted with the same organization as in this master thesis project, found that the level of responsibility in the organization reflected how concerned the employees were about security routines. Consequently, the research in this master thesis project investigates if this also applies to classification routines and whether enforcement from top level behavior is required to ensure that the employees will successfully adopt the new solution. Demographic data, such as job position in hierarchy, was therefore valuable information to include in this master thesis project. Strand also found that the employees had different understanding of what information security meant in the organization. Therefore, it most likely exists different expectations to a DLP solution and different classification routines and this is also investigated in this master thesis project.

2.5 Technology acceptance model

Several theories have been developed to predict how users accept and use new technology. One relevant theory for this study is the TAM model [FD86] shown in Figure 2.2. The focus of the model is how individual factors affect a user's acceptance of

the technology. Among these factors are *External Variables*, such as age, gender and experience, *Perceived Usefulness*, *Perceived Ease of Use*, *Attitude* and *Intention to Use*. *Perceived Usefulness*, *Perceived Ease of Use* and *Attitude* are considered motivational factors and are described in more detail in Table 2.1. In later time, the model has been developed further by extending the number of factors that may affect the actual use of the new technology.

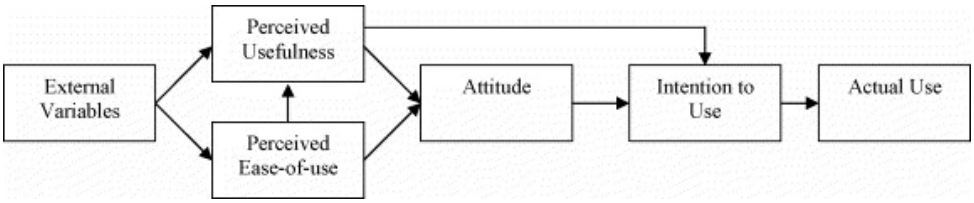


Figure 2.2: TAM [BJH06].

Factor	**Description**
Perceived Usefulness	The degree to which an individual believes that using a particular system would enhance his or her job performance.
Perceived Ease of Use	The degree to which an individual believes that using a particular system would be free of physical and mental effort.
Attitude	The degree of evaluative affect that an individual associates with using the target system in his or her job.

Table 2.1: Definitions of motivational factors in TAM [FD86].

Davis [FD86] states that *Perceived Ease of Use* affects *Perceived Usefulness* since a system that is easy to use will result in increased job performance and thus greater usefulness for the user. He also connects the variables to different responses; *External Variables* respond to design features in the system, while the motivational factors relate to responses as follows; *Perceived Ease of Use* and *Perceived Usefulness* are related to cognitive response, *Attitude* is related to affective response, while *Actual Use* is related to behavioral response to the system. The cognitive factor is associated with the user's needs and perception of the technology. The affective factor relates to feelings and emotions towards the technology and the behavioral factor is concerned with how the user interacts with the technology [FD86].

Related research: When introducing security tools in an organization it is important that they are efficient and that human errors are minimized. The TAM model described has been used as a framework in previous studies in order to reveal how the employees perceive the introduction of a security tool and how motivated they are to use it. When Shropshire et al. [SWS15], used the TAM model in their research they extended it with the two additional personality factors; *conscientiousness* and *agreeableness*. They argued that these factors could better explain variances in user behavior. Even though it turned out that the factors contributed to a positive effect on intention to use the security system, the majority of the participants who said they were going to use the system did not. More precisely, only a quarter of those who said they were going to use the system actually did. This shows that the actual usage of a solution is affected by other factors than personality and intention [SWS15]. A potential reason could be the users' limited experience with the system and barriers that were not revealed at the time when the participants answered the survey. Consequently, this master thesis project performed a usability test prior to the implementation of the solution and engaged actual users to minimize the amount of perceived barriers.

Another factor that may be related to usage is habits. Burton-Jones et al. [BJH06] state that the actual usage is often more affected by habits than intent. Burton-Jones et al. refers to a study where the only significant predictor of later use of an Information System (IS) was prior use. This shows that factors related to habits or willingness to change should be addressed in the master thesis project. This argues for extending the TAM model with additional factors.

2.6 Potential barriers in technology

When introducing a new technology in an organization there is no guarantee that the employees will be satisfied. As there exists individual differences both in technical experience and attitude towards change, the technology can be perceived as a barrier in the employees' daily work. MacKay [Mac91] introduces several factors that can be perceived as barriers when a new software technology is introduced. These barriers can be grouped into features of the software, individual factors, external factors, as well as a combination of these. Such barriers are important to consider when implementing a new software solution in an organization. Some potential barriers that are relevant to consider in this master thesis project are shown in Figure 2.1.

Related research: There have been conducted studies regarding barriers and influences in adoption of technology related to learning and teaching. In a study by Beggs [Beg00] faculties at a university were investigated. The participants ranked different barriers related to use of technology; improved learning, clear advantages over traditional, equipment availability, technology ease of use and time to learn

Figure 2.3: Potential barriers in technology [Mac91].

technology. The study found that time to learn and training in the technology were essential factors for the adoption of the technology [Beg00]. Thus, it is of interest to investigate what the employees expect from the organization before implementing the solution.

A study by Ng. et al. [NKX09] found that a user's security behavior related to email attachments is determined by perceived susceptibility, perceived benefits and self-efficacy. The study also reveals that the users did not perceive barriers related to practicing safe email behavior. However, the participants all had an IT background, which might have affected the result. Participants in this master thesis project therefore have different backgrounds in order to get results that are representative for organizations with employees with different backgrounds.

2.7 Security and usability

There is a general belief that when it comes to software technology, usability and security are often in conflict [JDK06]. The term Human Computer Interaction (HCI) does not consider potential threats and vulnerabilities that can arise in a system or application [KFR10]. When there is also a focus on security the term can be extended to Human Computer Interaction - Security (HCI-SEC). The main challenge when implementing a system that includes security features and especially when adding security features to an existing software system is not to degrade the usability.

A threat model that includes both usability and security can be illustrated as shown in Figure 2.4 [KFR10]. In the model, the focus is on legitimate users' mistakes and not on malicious attackers. The legitimate user does not intend to break the system. The model shows that the factors *Memorability* and *Knowledge/Skill* are applicable both in the *Usability* and *Security* section.

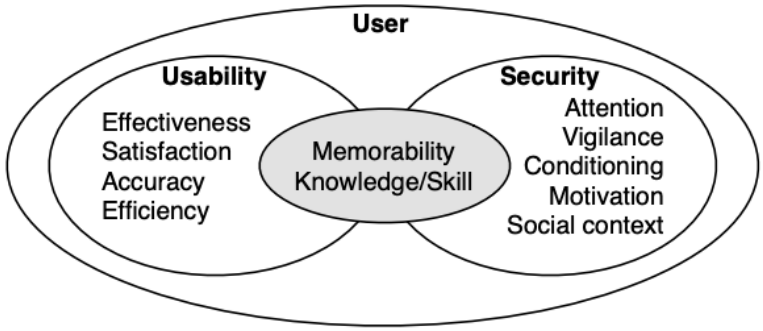


Figure 2.4: Security-usability threat model [KFR10].

Memorability, when related to usability is concerned with to what extent the user has to remember and recall something, e.g. a password for login. The usability is affected when the user has to remember a large number of passwords, making the task hard to accomplish. In addition, a password policy or forgetfulness may require frequent password resets. When related to security, memorability may degrade it since a user may write passwords down in order to remember them or use the same password many places.

The Knowledge/Skill factor, when related to usability, refers to how easy it is for a user to learn and operate a system. When related to security it refers to the extent to which a user knows when something is secure and not. For instance, a user can struggle with distinguishing between a secure and an insecure website.

Related research: The usability and security aspect of software systems has been topic for research since the 2000s. In studies reviewed by Kaida et al. [KFR10] it is found that the users' focus when using a system is on the parts they consider important and security tasks are often not among these. Furthermore, some studies found that training had little effect on the usage of the security features as these were not a part of the users' goals [KFR10].

In a paper by Fidas et al. [FVA10] the contradicting design requirements faced by developers when designing for usable security where examined. The paper concludes that the priority must be for the user and not the system. A bad user interface experience due to security features will have negative effects on the security. In addition, whether the primarily focus of the system is security or just an additional feature is an important factor when designing the user interface [FVA10]. Dhillon et al. [DOSC16] present a design guidance for software developers and engineers in

the light of security and usability objectives. They provide a set of objectives as a decision basis in order to balance security and usability. Four main factors are identified: maximize standardization and integration, maximize ease of use, enhance system related communication, and maximize system capability [DOSC16].

A usability study of the software tool Polaris with focus on security aspects was conducted by Dewitt et al. [JDK06]. The study found that despite the fact that the security features in Polaris were designed with usability in mind, the participants still had problems making security related decisions. Getting work done fast had a higher priority than knowingly compromising security. An important finding in the study, that is also considered in this master thesis project, was the additional difficulties introduced when the security features are added to an existing system rather than being integrated from the start. Overall, the study identifies three usability problems that must be addressed; *reducing the burden on the user to make security related decisions, counteracting user's apathy by ensuring that the fast way of doing things is the secure way and integrating security software with the operating system throughout development* [JDK06].

A user study by Zurko et al. [ZKSB02] explored the effect of changing a default security value on an active content protection mechanism from *Open* to *Secure* in the client-server software platform Lotus Notes by IBM. It was found that users would still allow unsigned active content to run since there was no change in how the choice to proceed was presented in their workflow. It was also found that the security culture or security-related user interfaces must be changed if warning users by making them click boxes to proceed with their work should have an effect. Zurko et al. recommend change in terms of education and appropriate information from the software. A software that can easily distinguish between safe and unsafe conditions, educate the users to choose the safe option and review and audit the unsafe option is preferable [ZKSB02]. The findings in the above studies are important as they provide useful insights for the investigation into finding a proper balance between security and user experience.

This chapter has presented relevant terms and factors that can affect the use of new technology in an organization, such as perceived barriers, users' information security awareness and the managers' engagement and influence on the employees. In addition, the TAM model was presented as a framework for designing an information system study. Several studies on security behavior in the organizational context have been conducted. Yet, there is a limited research done on how technology can affect employees' classification routines and awareness and to what extent potential barriers affect the adoption of technology. Given this background information, details about how this master thesis project was conducted and how the data analysis was performed will be presented in the next chapter.

Chapter 3
Methodology

This chapter presents the methodology used in this research. Arguments are given to explain why the different methods were chosen to provide accurate answers to the research questions defined in Chapter 1. In addition, the methods' strengths and limitations are discussed.

Since many of the aspects being investigated in this project concern social science, such as the employees' awareness of classification routines, willingness to adopt new technology, organizational culture and management influence, it is natural to consider methodology applied in social science research. In addition, the project includes both qualitative and quantitative data collections. Based on this, the books Real World Research by Robson [Rob11] and Qualitative Research as Stepwise-Deductive Induction by Tjora [Tjo18] are chosen as the main references for the research approach.

3.1 Mixed methods research

A research method for projects where there are both qualitative and quantitative data collections is often referred to as *Mixed Methods Research*. Leech et al. [LO09] defines the method as:

> *In general, mixed methods research represents research that involves collecting, analyzing, and interpreting quantitative and qualitative data in a single study or in a series of studies that investigate the same underlying phenomenon.*

The design ensures triangulation, which means that data from different sources will support the findings [Zoh13]. While quantitative variables of interest are more defined at the beginning, a qualitative data collection makes it easier to explore unknown variables and can therefore contribute with new knowledge [Rob11]. It is

therefore important to be aware of the differences in qualitative and quantitative research and how to combine them. The strengths and challenges of using the *mixed methods approach* are outlined in Table 3.1 and discussed throughout this chapter. In addition, depending on what the research's purpose is, there are several types of multi-strategy designs that can be applied. In this case, the model Evaluation Research was used [Rob11].

Strengths	**Complexities**
Triangulation ensures enhanced validity through both qualitative and quantitative research.	Lack of *skills and training* in both qualitative and quantitative methods.
Completeness is accomplished through combined research approaches.	*Timing issues* cause by different time frames related to qualitative and quantitative research components.
Offsetting weaknesses in each single method approach and providing stronger inferences.	*Limitations* in cases where there are no obvious advantages of combining qualitative and quantitative findings.
Ability to deal with complex situations	
Explain findings to a greater extent as one can verify findings using a different approach	
Illustration of data by using qualitative data to better understand the qualitative data	
Instrument development and testing by using results of qualitative research to refine research questions in quantitative phase.	

Table 3.1: Strengths and complexities in mixed methods research [Rob11].

Evaluation research The purpose of an Evaluation Research is to measure the effect or effectiveness of some implementation, such as an invention or product [Rob11]. It is also suitable to use when issues with a program need to be highlighted

and the time frame is limited. In this project this is done through usability testing, in addition to qualitative interviews and a quantitative questionnaire.

There exist two types of evaluations; summative and formative. While the summative type investigates the effect of the implementation, the formative type is concerned about potential improvements that can be done in the implementation. This master thesis project makes use of both types.

According to Robson [Rob11] it is important to begin an Evaluation Research with a *need assessment*. A service or program is often considered to be implemented since the current situation asks for it or the current solution does not meet a certain perceived need [Rob11]. Thus, a need assessment is usually set up in order to know what to prioritize in the investigation. The organization in this study was considering implementing a DLP solution in order to secure the organization's data. The concern was that the employees would not use the solution due to potential perceived barriers. Thus, there was a need for a solution that minimizes the amount of barriers and at the same time provides the desired level of data protection.

Robson [Rob11] lists a set of criteria that an evaluation should meet. For the project at hand, a reasoning is given for each of the required criteria; utility, feasibility, propriety and technical adequacy.

- *Utility:* The project is useful for organizations considering implementing a DLP solution as the results can contribute to decision making related to what features to introduce and how to gain the best possible user experience.

- *Feasibility:* The project is conducted in practical and cost-effective terms as the interviews were conducted in the participants' office, using their personal computer and scheduled to fit their time schedule. Furthermore, the questionnaire was answered whenever the participants had time. The time frame was limited, and the only financial cost was the software tool *SurveyMonkey* [1] used to create and analyze the questionnaire.

- *Propriety:* The project was conducted in an ethical way since an inquiry was submitted to and approved by the Norwegian Center of Research Data regarding data gathering and research ethics. The participants were informed about what kind of data that was going to be collected and how it was going to be used and stored.

- *Technical adequacy:* Technical skill and sensitivity is considered when choosing what programs to use related to data analysis and storage.

[1] "SurveyMonkey", SurveyMonkey, accessed March 13, 2019, https://www.surveymonkey.com

As with other research methods, there are challenges with using Evaluation Research. One challenge is to engage participants [Rob11]. It is not unusual to experience lack of interest among the employees regarding information security. In addition, one might expect resistance to potential additional tasks in their daily work life. To reduce this resistance and motivate the participants to engage in the implementation, it was emphasized that their responses were valuable input to the organization's process of implementing the solution. In addition, their contribution could have a positive effect on how the solution would affect their daily work life.

How the background information was collected and how questionnaires, semi-structure interviews and usability testing were constructed and applied in this project is explained in more detail in the following sections.

3.2 Literature review

According to Okoli et al. [OS10] a literature review can be classified based on what the purpose is. In this research case, literature review is used as a theoretical foundation for primary research. Gaps in current research are identified and existing evidences are summarized in Chapter 2. Okoli defines systematic literature review as:

> *A systematic, explicit and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners.*

3.2.1 The information collection process

The information collection process that was used in this study consisted of three phases described in detail below.

Defining the research area: It was informed that the case organization was already in the process of piloting an implementation of a DLP solution. There were discussions about how this was going to be done successfully. It was also established that some of the members of the pilot group were skeptical towards the implementation mainly because they believed it would involve additional work. They expressed concern that the employees would simply choose to ignore the solution. Being aware of these concerns and also the organization's need, a discussion with the supervisors was conducted which lead to an agreement to focus on the user experience aspect of the implementation. The substantial question was to find out how to implement the solution in a way that would make the employees accept and actually use it.

Background research: After defining the research area, a literature review was started in order to search for previous studies related to information security and user experience. In addition, a number of existing DLP solutions and their features were studied. The search engine Google Scholar [2] was used to find relevant scientific papers. Social networking sites, where researchers and scientists share content and results, were used to search for relevant information. Among the most used in this study were ResearchGate, IEEE Xplore, Emerald Insight and ScienceDirect. Table 3.2 shows the search terms that were most actively used.

Information Security	
	Behaviour
	Classification
	Policies
	Organizations
	Barriers in technology
	User experience
	Technology Acceptance Model
	Technology adaptation
	Usability
Data Loss Prevention	
	Azure Information Protection
	User-driven security
	User experience

Table 3.2: Search terms actively used.

Analysis and evaluation: When doing the background research, the main focus was on reading abstracts and conclusions in the found literature. In this phase however, the most relevant literature was identified and selected for more in-depth study. As possible research questions and the direction of the research became clearer the following requirements were considered in order to help decide which studies were most relevant.

- *Setting:* The setting should be directly or indirectly related to the research area and preferably also in an organizational setting.

- *Participants:* The participants in relevant studies should be comparable to the case organization, that is, include employees with different background.

[2] "Google Scholar", Google, accessed January 20, 2019, https://scholar.google.no/

- *Sampling methodology:* The study should preferably be based on both qualitative and quantitative methods.

The result of the literature review is presented and categorized in each appropriate section in Chapter 2. In addition, since an existing software product is a central part of the research, detailed information and characteristics about the solution must be examined [Rob11]. The DLP solution used in the research is described in Chapter 4.

3.2.2 Validity and reliability in literature review

According to Dellinger [Del05], a literature review is not only about collecting evaluations and results of single studies. The review's validity and reliability are also affected by the researcher's interpretation of the meaning of the evaluations. Thus, there is a potential for subjective measurements of what is known and unknown in the field of study [Del05]. To what extent these assessments can be considered valid depends on the degree to which the researcher has focused on this in the process. It can therefore be challenging to reproduce the exact same review as researchers have different areas of interest. In addition, some papers might have been missed or new research may have been conducted in the area, resulting in inaccuracies. The literature review was conducted with these aspects in mind, but there is still no guarantee that all relevant research is included.

3.3 Usability testing

A key quality criterion for any product or service is usability. According to Diah et al. [DIAD10], usability testing brings benefits, such as low training cost, increased productivity and improved user satisfaction. It is therefore a suitable method to use when studying challenges and potential barriers related to the introduction of a DLP solution. For example, a usability test can reveal how intuitive a solution is. According to ISO 9241-11 [ISO18], usability is defined as follows:

> *Usability is the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*

The properties referred to in the ISO definition are described in Table 3.3.

3.3.1 The usability test approach

In this study, the focus of interest is the employees' perceptions and interactions with the DLP solution. Since the solution can be customized, the results of usability

Effectiveness	Accuracy and completeness with which users achieve specified goals
Efficiency	Resources used in relation to the results achieved
Satisfaction	Extent to which the user's physical, cognitive and emotional responses that result from the use of a system, product or service meet the user's needs and expectations

Table 3.3: ISO's definitions of usability properties [ISO18].

tests may affect which features that will be included in the implementation and also how they are configured. Before the testing started, the participant was interviewed in order to collect background information.

Inspired by the methodology of usability testing used in the research by Diah et al. [DIAD10], the usability testing in this study followed a similar approach. The methodology is presented in the flow chart in Figure 3.1.

Planning usability test: Planning is an important part of usability testing as this is where goals and potential problems are identified. In this case, the goal was to find potential barriers in a DLP solution and to ensure a good user experience with security in mind. Common scenarios were identified and put in context with the solution to be tested. It is important to emphasize that the participants are not being tested themselves. They contribute to reveal weaknesses and find the most desirable features of the solution.

Identify usability test: The usability test model consists of three parts that need to be identified; parameters, method and participants.

Identify parameters: It is clear that the *effectiveness* and *efficiency* parameters shown in Table 3.3 are related to objective characteristics, while the *satisfaction* property is related to subjective characteristics. In a usability evaluation of a system the focus is on one or more of these characteristics of usability [KFR10]. As this study focuses on the user experience, the *satisfaction* property was emphasized. However, effectiveness and efficiency may have an impact on the user's satisfaction. In order to gather subjective data and assess the *satisfaction* characteristic, semi-structured interviews were conducted both before and after the usability test.

Identify method: The user-experience research method *Desirability Study* was chosen [Roh14]. Rohrer defines this method as follows:

> *Participants are offered different visual-design alternatives and are expected to associate each alternative with a set of attributes selected from*

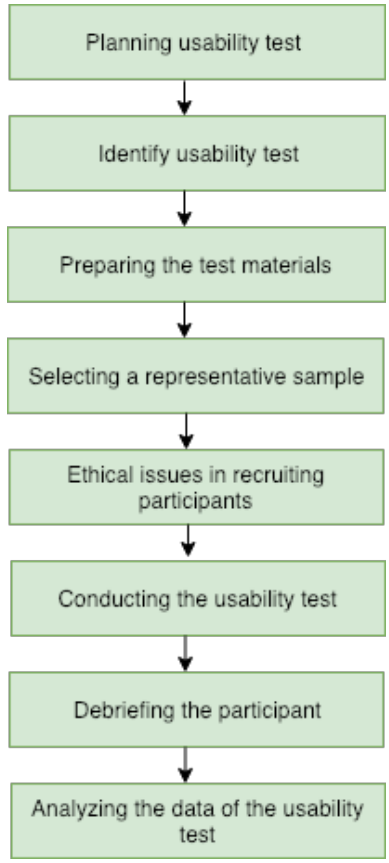


Figure 3.1: Methodology of usability testing [DIAD10].

a closed list; these studies can be both qualitative and quantitative.

In this case, the participants were offered alternative implementations of the AIP features that were considered. The attributes associated with the alternatives allowed the participant to assess which was the most satisfying. For some of the alternatives a checklist was used, but in some cases open-ended questions were more suitable. The complete scenario scheme can be found in Appendix B. Only one facilitator conducted the usability test sessions. The semi-structured interviews took about 20 minutes each and the usability test took approximately 30 minutes to complete. Each interview and test lasted for about one hour.

Identify participants: As one of the research goals was to find out if there was

a difference between managers and employees classification routines, the testing included representatives from both groups. The usability tests were conducted in the organization's office.

Preparing the test materials: In order to identify relevant scenarios, the DLP solution's capacity, properties and limitations were investigated. The most relevant features are described in Chapter 4. The scenarios were then developed in collaboration with the supervisors, who are experienced with the DLP solution used in this research and also knew what areas were the most important to focus on for the case organization. The scenarios were created to identify an acceptable balance between usability and security, e.g. to what extent pop-ups could be introduced without being ignored and still have a security enhancing effect. The different scenarios and their purpose are described in Table 3.4.

The scenarios were then tested on a fellow student. Since the tests were conducted by a single facilitator and also since changing program settings for the different scenarios would have been very time consuming, only parts of the usability test were conducted by running the DLP program. For the remaining test scenarios screenshots of the program displays were used. As most users are familiar with using Office programs it was not considered problematic using screenshots instead of a running version of the program. A checklist was created in order to ensure all scenarios were completed. This can be found in Appendix B. In addition, a step-by-step installation guide of the solution was sent out to the participant in advance so that the program was ready to be used. The required equipment for the tests were an audio recorder (the facilitator's smartphone), the participant's computer with the DLP solution installed and the test form.

S #	**Scenario Description**	**Purpose**
1	Explore the DLP functions in Microsoft Word and provide arguments for the desired level of control when applying a classification label	Test the balance between what the system should do and what the users want to control themselves.
2	Assess the justification requirement when performing an action that violates the organization's policy.	Check if the feature is perceived as useful and if a pop-up in such cases is desirable.
3	Assess to what extent templates should be classified in advance	Test the balance between what the system should do and what the users want to control themselves.

4	Suggest and explain how to best integrate customers' classification schemes in the solution	Explore what is the most suitable interface for integrating multiple schemes. Get an insight into the amount of features that are desirable before the solution is perceived as too clustered
5a	Explore the DLP features in Outlook and provide arguments for how the system should respond when managing confidential content in email	Finding the balance between user experience and security by exploring what amount of pop-ups and prevention of actions is acceptable
5b	Evaluate how the solution should control situations where email and attachments have different classification	Finding the balance between user experience and security by exploring what amount of pop-ups and prevention of actions is acceptable
5c	Evaluate how the solution should control subject and file names as these may reveal sensitive and confidential information	Finding the balance between user experience and security by exploring what amount of pop-ups and prevention of actions is acceptable

Table 3.4: Usability test broken down into scenarios.

Selecting a representative sample: According to Hinderer [Hin98], an effective representative sample is crucial to collect reliable data during usability tests. The participants experiences and opinions will only be reliable for identifying meaningful improvements if the participants both reflect the characteristics of the targeted users of the product or service and are likely to use it [Hin98]. In this master thesis project, participants both with and without a manager role was required. In order to recruit participants from both groups and with different background, one of the supervisors who work for and knows the organization assisted in the invitation process. According to Nielsen [Nie00] when using usability test, 5 users provides almost as good results as when using many more test participants. In addition, involving more participants will make the testing more expensive without providing significantly more precise results [Nie00]. Thus, in this study a total of 7 participants were tested from two different groups. Four of the participants had a manager role.

Ethical issues in recruiting participants: As the participants were going to be audio recorded, they had to be informed about this in advance. Before the usability test was conducted, the participant signed an agreement concerning this.

Conducting the usability test: The test was conducted in the organization's office. Only the participant and the facilitator were present in the room. An audio recorder captured the participants meanings and reflections. As *satisfaction* is the property in focus, the participant's thoughts regarding user experience were the most important. In this sense, the time it took to complete a scenario was not considered important as there were discussions between the facilitator and user during the test.

Debriefing the participant: When the test was completed, a semi-structured interview was conducted to get an insight into how the solution was perceived. The participant shared his or hers experience with the solution and additional thoughts and input were noted.

Analyzing the data of the usability test: To get a full overview of the participants' user experience, the interview and usability testing material was transcribed and analyzed. The details of the analysis phase are described in Section 3.7. The participants feature requests and perceived barriers were in focus, in addition to how the solution possibly would affect their classification routines.

3.3.2 Validity and reliability in usability testing

Important factors in order to assess the quality of the method chosen are discussed below. The paper written by Riege [Rie03] introduces four variables that can be assessed in order to improve the quality of the test design; construct validity, internal validity, external validity and reliability. The paper is used as the main reference in this subsection and it is referred to other papers where relevant.

Construct validity [Rie03] refers to the degree a specific test measures what it is supposed to. In cases where the researcher has a close connection with the research object; the organization or participant, it can lead to subjective judgments, such as selective memory, selective attention and selected encoding. Thus, in order to ensure construct validity, this must be avoided. As there were no personal or close connection between the researcher and participant in this case, the chances for introducing a bias were considered small. Multiple sources of evidence were used in the data collection phase in order to protect against researcher bias. In addition, data was gathered using multiple data collection methods. This is referred to as triangulation.

Internal validity [BH13] is concerned with factors of the selected human subjects that can affect the result. In this case, a possible threat to internal validity would be to only involve participants with an IT background. However, it can be difficult to assess their IT knowledge. In order to ensure a neutral introduction to the usability test, the following statement was used:

> *We will go through some scenarios using a software solution that is going to be implemented in the organization. The solution aims to protect documents and other information in the organization.*

No further instructions or training were given. It was also made sure that the participants knew what a usability test was and that the purpose was to test if the solution was useful and easy to use.

External validity [BH13] refers to the extent the results from the study can be generalized to other situations or groups of people. People, place and time can affect this variable. The fact that most of the participants were researchers may have affected the results. In addition, the participants were not introduced to all possible DLP features in the solution in order to limit the scope and focus on the ones that are assumed to be most important to the case organization. This may also have affected the external validity as the result might be different for other types of organizations.

Reliability [Rie03] refers to the extent the study can be replicated. It means that procedures and operations can be repeated by other researchers conducting the same study and that they will get similar results. As the participants tested are not static measurements, there will be differences. However, the differences revealed in later studies might be an interesting source of information in the research area. To ensure reliability, the usability test and interviews were audio recorded to get as accurate data as possible. The usability test material used is included in Appendix B.

3.4 Semi-structured interview

Semi-structured interviews are frequently used in mixed method research when the aim often is to explore hypotheses and gain new knowledge about a topic. Robson [Rob11] states that interviews conducted face-to-face make it possible to respond to relevant statements, ask follow-up questions and clarify misunderstandings, which is not possible in questionnaires. Both planned and unplanned questions are asked from a checklist of topics to be covered. In addition, it is both a practical and suitable approach in cases where the researcher is also the interviewer because of his or her in-depth knowledge about the research area [Rob11]. Both recording and taking notes during the interview is preferable. Taking notes ensures the information is available if something goes wrong. In addition, the notes can be used at the end of the interview to reflect on ideas, feelings and memory of the discussion [DN13].

The interviews were pilot tested by a fellow student. Questions were adjusted after a few iterations to make them as clear and understandable as possible. Two

interviews were conducted per participant; one before the usability test and one after. The first interview was sent out before the usability test in order for the participants to start reflecting about their security routines. The results from the interviews and usability testing are presented in Chapter 5.

3.4.1 Constructing the interviews

According to Robson [Rob11] a semi-structured interview often starts with an introduction part with warm-up questions, including relevant subtopics. For each subtopic, a number of key questions are asked. Then, the main part follows with in-depth questions related to the different topics. Finally, at the end of the interview, there are closing comments [Rob11]. Here the participants are asked if they have additional thoughts that they would to like include. Composing an interview in this way is also supported by Tjora [Tjo18] as shown in the illustration in Figure 3.2. Both the interviews before and after the usability test had the described composition.

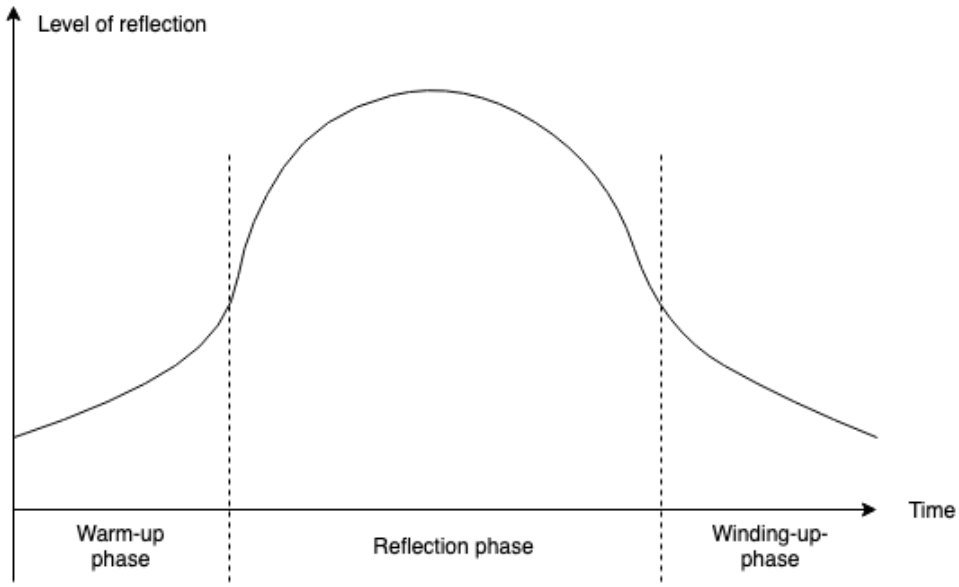


Figure 3.2: Tjora's suggested composition of an interview [Tjo18].

Most of the questions were formulated as open-ended to encourage the participants to answer with their own words, making it possible to reveal problems and explore aspects that spontaneously could arise [DN13]. The semi-structured inter-

view conducted before the usability test included questions grouped in the following categories; background, introduction, classification routines, classification awareness and expectations towards the DLP solution. The questions included in the interview conducted after the usability test where grouped in these categories; usability and perception of ease of use, predicted effect on their classification routines and usefulness of the solution in their daily work. The interview guide can be found in Appendix A.

3.4.2 Validity and reliability in semi-structured interviews

There exist several risk factors that may challenge the quality of a semi-structured interview and that one should be aware when conducting the interview. The most common pitfalls are described below.

Location of the interview [Tjo18]: Tjora states that conducting an interview in the interviewee's office during their working hours, as done in this research, introduces a high possibility for disturbances, such as phone calls and interruption from colleagues, as well as limited available time. On the other hand, Tjora also emphasizes the need for the participant to feel safe and being in a known environment. The fact that the interviews were conducted in the participants own office, may have contributed to a comfortable setting. The interview and usability tests were both conducted within working hours since the participants in general were busy. Also, their office was considered an appropriate location since this is where they will be using the DLP solution.

Duration of the interview [Tjo18]: According to Tjora, having a sufficient length on the interview is important to make the participant feel comfortable. However, Robson [Rob11] states that interviews that last longer than one hour would be considered time-consuming and thus affect the number of people who want to participate. A recommended length of an interview is somewhere between 20 minutes and one hour. Each of the interviews in this study lasted for about 20 min, in total about 40 minutes per participant.

The interviewer's experience [DN13]: One of the risk factors with an inexperienced interviewer is that relevant data may be ignored or missed. The reason is that knowing when to ask follow-up questions or probe responses requires experience with conducting interviews. In addition, an experienced interviewer knows how to ask open-ended questions and capture new concepts that may arise during the discussions. This may increase the validity of the study. As the interviewer in this research had limited experience with conducting interviews there is a possibility for this risk factor. However, the interviewer was aware of this and the other pitfalls presented and prepared with these in mind.

Recording and note taking [DN13]: Another concern is related to recording and taking notes. On one hand, there is a possibility that taking notes during the conversation can disturb the flow in the interview and cause misunderstandings through lack of concentration. On the other hand, writing notes only after the process may cause a loss of details. It was therefore decided that only key words would be written down during the interview in addition to audio recording. There is always a risk that some participants may feel uncomfortable knowing that they are being recorded and it may also affect how freely they speak. To reduce this risk the participants were ensured that the recording was only going to be used by the interviewer to make sure everything said was captured and that all recordings were to be deleted when the project was completed.

Ethical considerations [DN13]: During an interview, sensitive issues often arise, and the interview situation is subject to ethical considerations and risks. The participants were informed about what kind of information that was going to be gathered before the interview started. When participants are required to answer questions related to sensitive or personal aspects of their life it will often affect how they choose to answer. However, as the only sensitive information in the interview conducted in this study is the participant's job position and all data is kept anonymous, this was not considered an issue.

Validity and dependability [Tjo18]: In contrast to participants in a survey, participants in an interview are to a greater extent expected to reflect over their personal opinions and experiences as the researcher ask open-ended questions. However, when interviewed, participants often expect a formalized setting and are inclined to provide short answers. This is why a prepared set of follow-up questions are needed. In addition, there is a possibility that participants leave out details related to their experiences and perceptions since they might think they are not of interest to the interviewer or they want to keep the information to themselves. This is something that can be difficult to compensate for. In addition, as the interviews may vary there is a risk that the questions asked to the participants are different. To compensate this, the interview guide in Appendix A was used.

3.5 Questionnaire

Questionnaire is a common quantitative research method used in social research [Rob11]. It is important to have a clear definition of what kind of information to collect and also to ensure the formulations of the questions are well prepared to reduce risks of bias. As the time frame was limited and it was desirable to collect answers from a large number of employees having different job roles, a self-completion questionnaire was chosen [Rob11].

The questionnaire was constructed after the interviews and usability tests were completed. This approach was chosen in order to have a questionnaire that could be more focused on important aspects that were revealed during the interviews. In addition, questions related to the research questions and hypotheses that needed a larger amount of responses were included.

3.5.1 Respondent recruitment

One of the main challenges with questionnaires is to engage respondents [Rob11]. People are busy and completing a survey is often considered time consuming. It is therefore important to communicate how quickly it can be answered. It should also be informed that the participation is anonymous and that the data will be kept confidential. This was taken into account and also, instead of sending an email to the whole organization, the questionnaire was uploaded to the organization's internal platform and the employees were notified. This ensured that the employees could answer the questionnaire when they had the time, and avoided the problem with emails being ignored, deleted or ending up in the trash folder.

In total, 36 responses to the questionnaire remained for data analysis. Among the respondents, only three were managers. However, as it was voluntary to answer the questionnaire, it was not possible to control the total number of respondents and the distribution of respondents between departments in the case organization. These types of limitations to the research may cause biases for example if the majority of the respondents are employees interested in the topic.

3.5.2 Constructing the questionnaire

The TAM model described in Section 2.5 was used as guidance when constructing the questionnaire. Additional factors were included in the model to answer key research questions in this study. These are related to ease of use and usefulness of a software solution, its perceived barriers and how it may affect the employees' classification routines. In order to measure a possible effect on the employees' classification routines, there was a need to investigate to what extent they were familiar with the organization's classification policy. In addition, how aware they are of information security in general may be a dominant factor in their answers as it could indicate how they make and will make security related decisions [WS01]. The factors *Information security awareness* and *Organization policy awareness* were therefore added to the TAM model.

A number of questions that falls into the categories in the customized TAM model shown in Figure 3.3 were defined. The answers were expected to give an indication of whether or not the system would be used. In addition, they were expected to

reveal the employees view on the importance of implementing it and how willing they would be to change their habits and prioritize security in their daily work.

There is a disagreement about whether or not the *Attitude* variable is necessary to include in the TAM model. Some studies have found that there are no statistical differences in results when using a model including the variable compared to when not using it [LBLB17]. For this reason, and also since attitude can be closely related to intention to use, the variable is not included in the model used in this study.

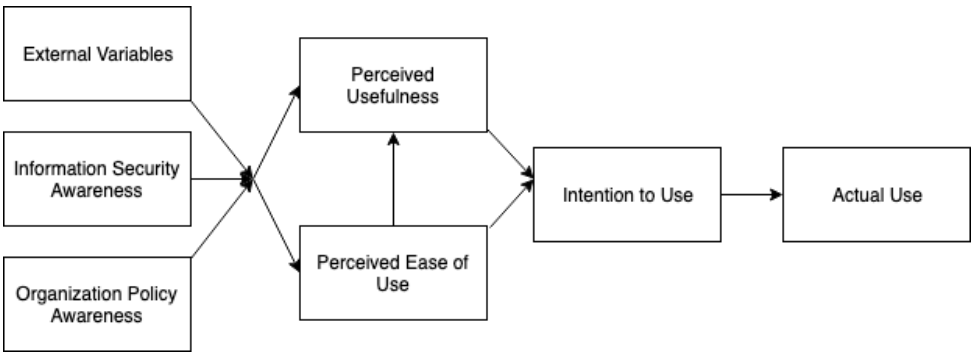


Figure 3.3: My defined version of TAM.

When constructing the questionnaire, a tip sheet written by Harrison was used as a guideline [Har07]. Important tips that were considered were to keep the questionnaire short, think about the order the different questions are asked and what types of questions to include; open-ended or closed-ended [Har07]. According to Zohrabi [Zoh13], a questionnaire consisting of both open-ended and closed-ended questions is preferable. Open-ended questions make it possible to discover new knowledge since the participants can answers in their own words. However, close-ended questions are more efficient to analyze and there is also a higher possibility that respondents would skip the open-ended questions. Therefore, some of the close-ended questions in the questionnaire also included a free text option to use if none of the alternatives were suitable [Har07]. When asking closed-ended questions it is important to include a suitable scale that is neither too restricted nor provides too many options. Harrison recommends a scale between 5-7 points. In the questionnaire used in this study, the closed-ended questions have a scale of 5 points.

It is important that the questions are formulated precisely to ensure they provide accurate answers to the research questions [Rob11]. The questionnaire was pre-tested by a fellow student to ensure that the questions were understandable and easy to read. According to Harrison one should avoid technical terms and jargon in a survey to ensure that the respondents understand the questions easily. As an example, to avoid

that users were confused or overwhelmed by unknown and complex terminologies, the words AIP and DLP were left out of the questionnaire and replaced with the word *security tool*.

The questionnaire consists of four parts. It starts with a warm up part, with relatively easy questions to answer, to get the participant started. These are related to current classification routines. The second part is concerned with knowledge and use of the organization's classification policy. The third part presents features in the solution to be implemented and includes questions concerned with the participants' perception of its usefulness and ease of use, potential barriers and how it will impact their daily work. The last part is the *End of survey* where the participants can add any comments. Some of the questions used were inspired by Shropshire et al. [SWS15], Safa et al. [SSS+15] and Bonilla et al. [LBLB17] which all make use of the TAM in their research. Table 3.5 shows the questions associated with the factors in the customized TAM model. The complete questionnaire can be found in Appendix C.

TAM Factor	**Question(s)/Statement(s)**
External variables	1. Do you have a personnel management role?
Information Security Awareness	1. I am aware of the consequences of classifying wrong 2. I believe classification of documents is important 3. I work with projects that may be exposed to information security risks, such as malicious attacks and industrial espionage. 4. Does or would working with projects exposed to information security risks affect your awareness regarding information security and the organization's security policy? 5. Getting work done fast has a higher priority than following the security policy

Organization policy awareness	1. I must meet additional classification requirements to the general requirements for the organization (e.g. customer classification schemes) 2. I am familiar with the organization's classification policy 3. I am often unsure about what classification level to apply 4. When I am unsure about which classification level to apply, I ... 5. Applying the organization's classification scheme policy is ...
Perceived Usefulness	1. The security tool will enable me to practice the organization's classification policy 2. The security tool will increase my job productivity 3. The security tool will be useful in my job 4. The security tool will make me more aware of the organization's classification policy 5. The security tool will improve my classification routines 6. There is a need for the security tool in the organization 7. What factors would prevent you from using of the tool?
Perceived Ease of Use	1. The security tool seems clear and understandable 2. Using the security tool will require low effort

Intention	1. I Intend to use the security tool.

Table 3.5: Variables from the TAM model used in the questionnaire.

3.5.3 Validity and reliability in questionnaire

In order to get reliable results, it is important to be aware of potential pitfalls. Boynton et al. [BG04] and Zohrabi [Zoh13] introduces important variables related to validity and reliability of a questionnaire. These are described below.

Construct validity [Zoh13]: Construct validity is concerned with measuring what is intended to be measured. In questionnaires, people tend to answer how they would like to behave in certain scenarios, not how they *actually* behave. The fact that the participants are informed that the questionnaire in this study is anonymous, is assumed to reduce this risk and contribute to more accurate answers.

Internal validity [Zoh13]: Internal validity is concerned with the risk that questions are misunderstood or appear unclear to the participants. The consequence of misunderstanding would be inaccurate responses. The questionnaire used was quality tested by a fellow student and the supervisors to ensure the questions were worded clearly.

External validity [BG04]: External validity is concerned with to what extent the results are generally valid. The context, type of organization and selection of respondents may influence this. It is therefore important to outline relevant limitations. In addition, Robson [Rob11] states that non-response bias is a risk as there should be a high response rate to represent the majority of the target group. This is addressed further in the discussion in Chapter 6.

Reliability [BG04]: Reliability is concerned with to what extent the questionnaire can produce consistent results when repeated by other researchers or at a different time. When results differ, it is most likely due to a different set of participants assuming that the questions and formats are identical.

3.6 Case context

The case organization is a knowledge organization consisting of 2000 employees. A knowledge organization is defined as an organization where the knowledge is regarded as the most important factor for the organization to be successful [Gru06]. In such

organizations, employees actively create and share knowledge [OO06]. The focus is more on the people working there, called *knowledge workers* [Gru06], rather than on following explicit procedures. This might result in a more open organization culture, where the employees are used to make own decisions, than compared to for example organizations in the process industry, where they are told what to do.

The organization had already chosen and started the process of implementing a DLP solution. The solution is described in Chapter 4. To limit the scope, this was the only solution tested in this study. The motivation for introducing a DLP solution was that the organization works with critical projects, also for external customers, and thus it is very important that their data is protected and that the employees handle documents in a secure manner. In addition, as there might also be a risk for industrial espionage, it is crucial that only certain authorized employees have access to specific project information. A DLP solution can contribute to this type of security provided that it is being used and accepted by the employees. The case organization has completed several security awareness campaigns [Str18]. This may already have influenced their awareness of security risks in general. However, there has not yet been conducted a campaign targeting protection of data specifically, such as classification of documents and emails.

A pilot group in the organization consisting of employees with an IT background has been involved since the early stages of this research study to identify potential issues with the DLP implementation. The organization's IT department is the stakeholder in this master thesis project. One of their main responsibilities is to perform actions to secure the organization's information and the implementation of the DLP solution is a part of this.

In the investigation of *H3*, "manager" refers to an employee with a personnel management role. Personnel managers were of interest as their relation to a project may be different compared to a project manager. A personnel manager is more likely to sign contracts involving a high cost and more responsibility, and may therefore be more concerned with information security. There are about 240 employees with a personnel manager role in the case organization.

3.7 Data analysis

In this phase, the data collected from the questionnaire, semi-structured interviews and usability testing was analyzed and mixed. Data from qualitative and quantitative data sets can either be linked or transformed into one data set. The transformation can be to convert qualitative data into quantitative data or convert quantitative data into qualitative data [San00]. As the interviews could provide additional new knowledge, the two data sets were analyzed separately and linked where they had a

common reference point in the research questions.

3.7.1 Qualitative analysis of interviews and usability tests

After each interview and usability test, notes were taken and any thoughts that immediately came to mind were summarized. Finally, transcripts of the interviews were made to include all of the participants' statements and ease the analysis process.

To analyze the qualitative semi-structured interviews and statements provided during the usability test, a thematic analysis method was found suitable to apply to the transcripts. Boyatzis [Boy98] states that:

> *Thematic analysis enables scholars, observers, or practitioners to use a wide variety of types of information in a systematic manner that increases their accuracy or sensitivity in understanding and interpreting observations about people, events, situations and organization.*

In thematic analysis, the qualitative data is encoded by a chosen "code". According to Boyatzis [Boy98] there are several ways to encode qualitative data. It can for example be done by following a list of themes and indicators, either generated based on the gathered raw information or theory and prior research [Boy98]. Boyatzis also emphasizes that only focusing on the underlying phenomenon may exclude valuable information in the raw material. However, including all observations made in the interviews and usability tests, may be outside the context of the research questions. When analyzing the transcriptions in this master thesis project, the focus was mainly on information that answered the research questions directly. Tjora [Tjo18] introduces a way of analyzing data, called *Stepwise-Deductive Inductive Method*. The stepwise model is supposed to reduce complexity and avoid jumping to conclusions. The model makes use of *in vivo* coding where the idea is to base the codes on the participants statements or concrete situations in the observations rather than on predefined themes. In this way the raw material is preserved. A similar approach is recommended by Malterud [Mal12]. Her method is suitable for cross-case analysis of different types of qualitative data, such as interview studies and observational studies. The strategy *Systematic Text Condensation* by Malterud was therefore chosen and applied in the analysis [Mal12]. However, Tjora's arguments were still kept in mind throughout the analysis. *Systematic Text Condensation* is designed with intersubjectivity in mind, in the sense that others can follow the same procedure and understand the conclusions. The procedure consists of four steps that were performed as described below:

Total impression – from chaos to themes: First, the transcripts were scanned and read from a bird's eyes view separately. Here, themes were identified in the

material and a sort method was decided on. The statements were coded using descriptive words, and data where the research questions and hypotheses were answered directly was categorized using colors specific to each question and hypothesis. The highlight tool in Google Docs[3] was used for this purpose.

Identifying and sorting meaning units – from themes to codes: In this step, the parts of the interviews that could answer the research questions and hypotheses were in focus. The statements were investigated in order to further analyze the meaning and identified by codes. Examples of codes were the desired degree of control level and types of barriers in the solution.

Condensation – from code to meaning: In this step the data was reduced from a large number of raw transcript pages to one single page for each of the participants in order to reveal aspects of interest. The answers from each participant were first summarized in separate documents and then further condensed into one scheme.

Synthesizing – from condensation to descriptions and concepts: In this step, the concepts were further developed and transformed into consistent statements regarding the participants experience from the demonstration of the solution. In addition, an important part of data analysis is to perform an assessment of the findings. The findings were then compared with prior research and relevant literature. The discussion is found in Chapter 6.

3.7.2 Qualitative analysis of questionnaire

The electronic questionnaire was created and analyzed using the online survey software tool *SurveyMonkey* [4]. The cloud-based software has several features for collecting, viewing and analyzing the responses in a survey.

View survey responses Responses can be grouped into question summaries, individual responses and open-ended responses. For visual analysis, dynamic charts were generated automatically, making the process easier. In addition, statistics such as time used to respond to a question or response completeness can be provided.

Using rules to analyze data When a view of data is chosen it is possible to further apply rules to answer the research questions more specifically. Types of rules are filter rules, compare rules and show rules. Certain criteria can be applied to filter out responses based on groups of respondents. In this research, a compare rule on

[3] "Google Docs", Google, accessed February 20, 2019, https://www.google.com/docs/about/

[4] "How to analyze results", SurveyMonkey, accessed March 13, 2019, https://help.surveymonkey.com/articles/en_US/kb/How-to-analyze-results

management role was used, to investigate any correspondence between role and the questionnaire statements. A screenshot of the feature is shown in Figure 3.4.

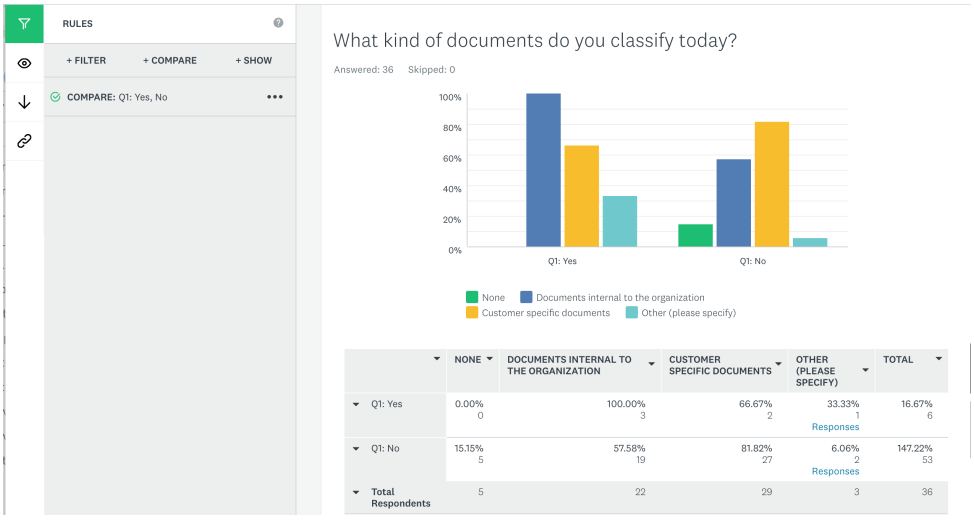


Figure 3.4: A screenshot of the filtering by Compare Rule function in SurveyMonkey.

3.8 Ethics / privacy concerns

Since the data collected in the interviews includes personal data such as the participant's name, e-mail and job position, an inquiry was sent to and approved by the Norwegian Center for Research Data (NSD). Before the interview and usability test was conducted, the users were informed about the purpose of the study, the types of data that would be collected, anonymization and storage limitation. In the interviews and usability test, the facilitator's password protected smartphone was used as an audio recorder. To ensure anonymity and confidentiality, all data recorded will be deleted at the end of the project. This was communicated to the participants. The questionnaire was made available online and there was no need to provide name or email address to answer the questionnaire. In addition, the client computer IP-addresses were not traceable, making it impossible to trace the answers back the users. When the research was completed, all the data gathered was deleted. The received confirmation from NSD can be found in Appendix D.

This chapter presented the methodology Evaluation Research that consists of both of quantitative and qualitative data collection methods. The reason why the method was suitable for this master thesis project was explained and justified. In

the next chapter, the main features of the DLP solution used in the research will be described.

Chapter 4
Azure Information Protection

In the market today, there are several vendors offering commercial Data Loss Prevention solutions. The solution selected by the case organization in this research is Azure Information Protection (AIP) developed by Microsoft[1]. AIP is a cloud-based solution that is compatible across all Office365 products. It aims to protect and secure documents, email and sensitive data that is shared inside and outside the organization.

Only features relevant for this specific research are presented here. AIP includes a number of additional features that may be applied and used by an organization, but in order to narrow the scope, these are not described here.

The description in this chapter is based on information from Microsoft's online documentation. All software screen captures used in this section are from Microsoft's webpages.

4.1 Features

The first step when setting up AIP is to specify where the data is stored and configure the *scanner* so that it is able to discover and classify sensitive information in all documents used by the organization, both on-premises and in the cloud. The scanner will do a one-time initial scanning of all existing data objects in the data store. The next step is to configure *labels* and *policy settings*. Policy settings are explained further in Section 4.2.

Labels is a central feature of AIP which is used to classify data. They represent the organization's classification levels and can be customized according to the organization's requirements. When applied to documents and emails, labels add restrictions to further actions on these objects.

[1] "What is Information Protection", Microsoft Corporation, updated May 20, 2019, https://docs.microsoft.com/en-us/azure/information-protection/what-is-information-protection

AIP supports a number of ways that labels can be applied to documents. These are controlled by the system settings. Labels can be specified both by using application commands and in File Explorer through context sensitive menus. Applying a label to a document can be a mandatory action by the user when creating a document as shown in Figure 4.1. Another alternative is that the system itself selects a default value for new documents as shown in the example in Figure 4.2. A third alternative is that the system recommends a classification label by applying rules defined in the configuration to content in the document. It is then up to the user to decide if it is a suitable level. Figure 4.3 shows a custom tooltip bar with a recommendation.

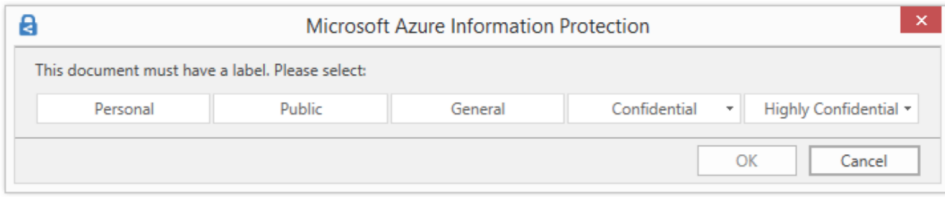


Figure 4.1: Mandatory to set value when creating a document in Word.

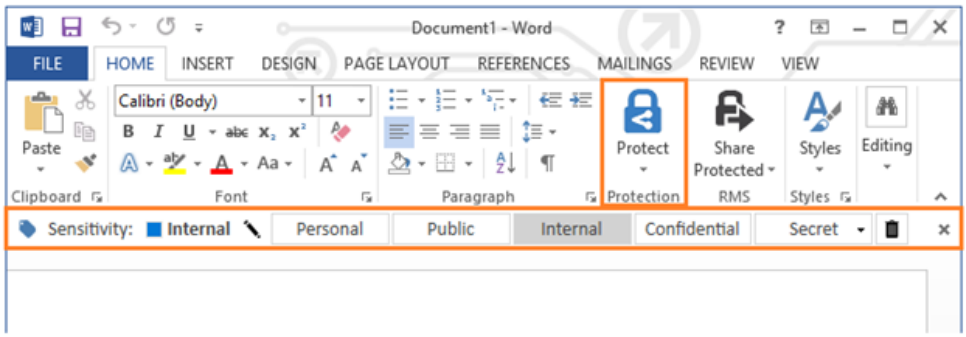


Figure 4.2: Default values set from creation of document in Word.

As mentioned above there are different ways that the user-driven classification options can be presented in the user interface. This is decided by the organization. When a label is applied to a document, the data is protected, and it is possible to track the document. Analyzing data flows makes it possible to detect undesirable events and prevent data leakage [2].

[2] "What is Information Protection", Microsoft Corporation, updated May 20, 2019, https://docs.microsoft.com/en-us/azure/information-protection/what-is-information-protection

In addition, there are configuration settings that control security functions that runs without user involvement or feedback to the users. Examples of these are fixed access rights, content blocking, and logging of email traffic, all based on metadata or information patterns in the data. As the users are not involved in these system decisions, it is not within the scope of this research.

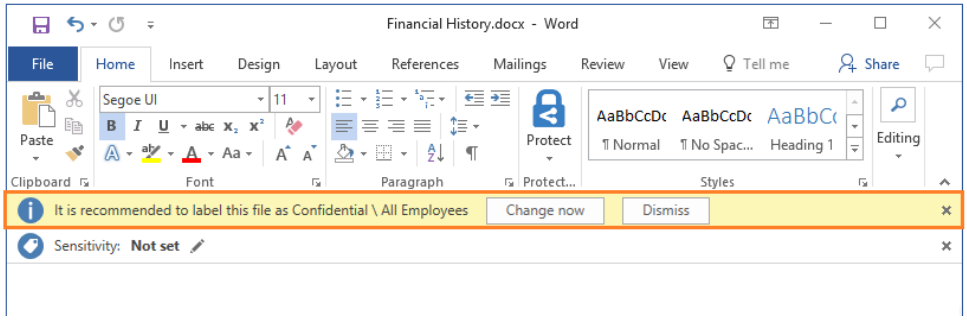


Figure 4.3: AIP recommends classification label in Word.

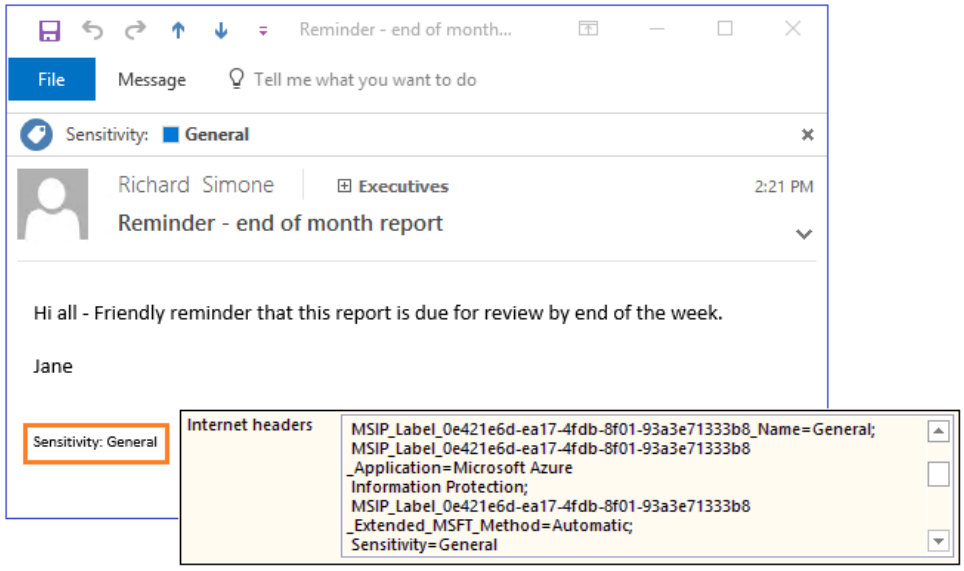


Figure 4.4: AIP in Outlook.

Headers, footers and watermark are examples of visual indicators that can be applied to a document or email. An example of this type of indicator in an email is shown in Figure 4.4. Adding metadata to files and emails in clear text makes it

possible for other services, such as DLP solutions, to identify what classification level is applied and what the appropriate action is. Regardless of who the documents are shared with or where the documents are stored, the classification is identifiable.

AIP provides additional protection functions for email. Recipients who cannot open a protected email in their personal email client e.g. Gmail, can log into a browser and view the content by using a one-time passcode. Figure 4.5 shows the different protection options a user has when sending a confidential email. In this case, the recipient is the only person who can read the email. Furthermore, emails can be configured to always provide the same protection for its attachments, but the attachments do not inherit the email labelling [3].

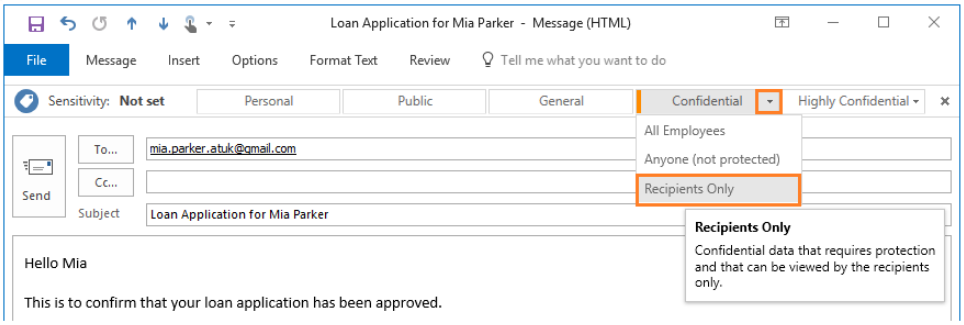


Figure 4.5: AIP in Outlook.

Figure 4.6 shows the customized toolbar embedded in Excel and with a default value set to *General*.

4.2 DLP policies

AIP allows you to configure DLP policies that comply with business standards, industry regulations and your organization's policies. With a DLP policy it is possible to [4]:

- Identify sensitive information across many locations, such as Exchange Online, SharePoint Online, and OneDrive for Business.

[3] "Frequently asked questions about classification and labeling in Azure Information Protection", Microsoft Corporation, updated April 17, 2019, https://docs.microsoft.com/en-us/azure/information-protection/faqs-infoprotect#when-an-email-is-labeled-do-any-attachments-automatically-get-the-same-labelling

[4] "Overview of data loss prevention", Microsoft Corporation, updated March 5, 2019, https://docs.microsoft.com/en-us/office365/securitycompliance/data-loss-prevention-policies

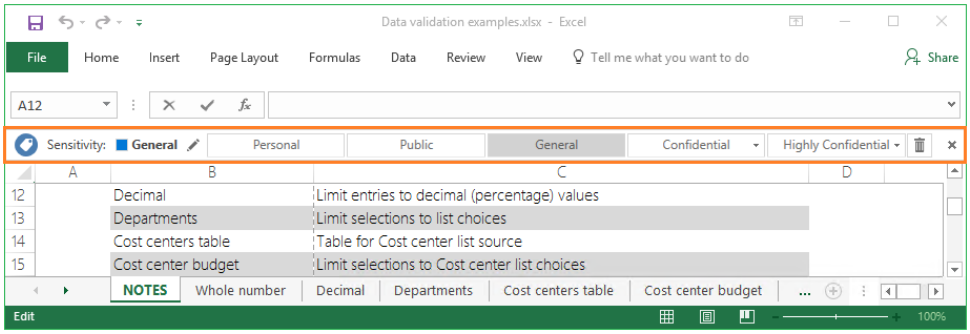


Figure 4.6: AIP in Excel.

- Prevent the accidental sharing of sensitive information.
- Monitor and protect sensitive information in the desktop versions of Excel 2016, PowerPoint 2016, and Word 2016.
- Help users learn how to stay compliant without interrupting their workflow.
- View DLP reports showing content that matches your organization's DLP policies.

A DLP policy is configured by creating a set of rules. As shown in Figure 4.7, a rule consists of a set of *conditions* that must be fulfilled for *actions* to take place. A condition is concerned with content and context. Content means the type of information that is included in the document or email, such as sensitive information and *labels*, while context is concerned with whom the information is shared.

Actions restrict who can access the content. For example, the document owner could be the only person with access. To change the restrictions, the owner may have to remove sensitive information or perform other remedial actions. In the case of email, an action could prevent it from being sent.

User notifications and *User overrides* can be used to educate the users about the policies without preventing them from doing their work. For instance, if a user wants to send sensitive information to a person outside the organization, a notification will inform the user that the action is not doable and that a user override, that is a justification for the action, must be conducted. An example of the justification feature is illustrated in 4.8. Finally, an *incident report* is sent to the administrator with details about the performed events.

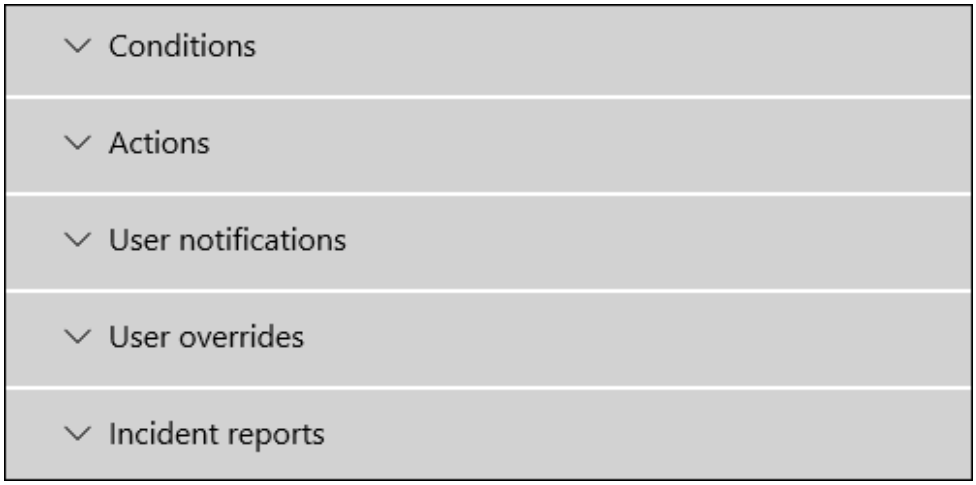


Figure 4.7: Rule specifications in AIP.

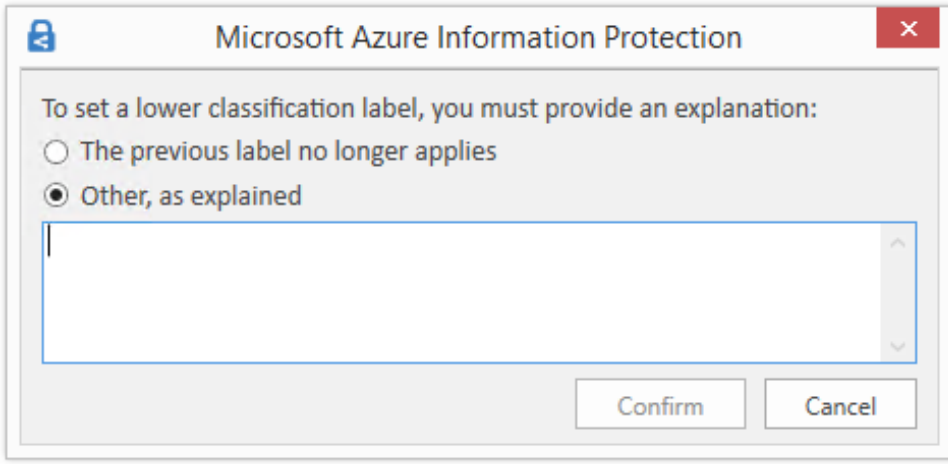


Figure 4.8: The justification feature in AIP.

When the AIP policy downloads to computers that have installed the AIP client, the system is configured with settings and the organization's labels from the configured global policy. If there is a need to supplement these for specific users, e.g. for users working with projects requiring special settings or customer labels, it is possible to

create a custom *scoped* policy for these users [5]. However, for users working with projects with different policy settings at the same time, this adds complexity to the system that may affect the user experience. The custom *scope* policy feature must therefore be used with care.

4.3 How information is protected

AIP makes use of Azure Rights Management (ARM) which is a protection technology integrated with Office365 and Azure Active Directory. Independent of the location of the document, protection in terms of encryption, authorization policies and identification is preserved. In addition, the technology can be used with other DLP solutions that are cloud-based or on-premises, and that does not make use of labelling. To limit the scope, this research only considers the parts of the solution that provide classification of documents and emails.

This chapter has described the main features in the DLP solution AIP. With this information in mind, the scenarios and research results can be better understood. The next chapter presents the findings in the interviews, usability tests and questionnaire.

[5] "How to configure the Azure Information Protection policy for specific users by using scoped policies", Microsoft Corporation, updated May 14, 2019, https://docs.microsoft.com/en-us/azure/information-protection/configure-policy-scope

Chapter 5
Results

This chapter presents the results gathered from the semi-structured interviews, usability tests and the questionnaire. As the interviews and usability tests were used as a basis for what should be explored in the questionnaire, the results from the interviews and usability tests are presented separately from the results from the questionnaire. The analysis of the results from the interviews and usability tests provides a good basis for answering *RQ1* since it included a detailed presentation of the relevant features of the solution. The results from both the interviews and questionnaire were used as a basis for answering *RQ2* and its related hypotheses; *H1*, *H2*, *H3*. Both managers and non-managers from the case organization participated in the interviews and usability tests. This grouping into roles aimed to shed light on hypothesis *H3* which was included to reveal any differences in classification routines.

5.1 Results from interviews and usability tests

As the interviews were held in Norwegian, the citations included are translated into English. An interview guide was used when conducting the interviews. This can be found in Appendix A. A detailed description of the scenarios can be found in Appendix B and the purpose behind them is described in Table 3.4.

5.1.1 DLP features

Five scenarios were described and followed by a discussion with the participants in order to investigate the balance between security and usability in the solution. Multiple DLP features were introduced and discussed. The results from each scenario are presented separately.

Scenario 1 In the first scenario, three different user interface alternatives for controlling classification in Microsoft Word were introduced. By choosing an alternative, the users expressed their desired degree of manual control of the classification.

As shown in Figure 5.1, in total 6 out of the 7 participants preferred option 2; having a default classification level set. The participants justified their choice with different reasons. One was that the toolbar is more visible and accessible with this option compared to the others. In addition, having *Internal* as the default classification level prevents the undesirable effect of accidentally having documents classified as *Open* or not classified. One participant said that the majority of internal documents are written in Word and option 2 is therefore a good choice as it provides a safe basis in most cases. Another point made was that it provides the best possible protection with low effort and minimal risk of error. Two employees also pointed out that they frequently use templates when opening new documents. Thus, having a default security level applied to templates would be desirable. This is further explored in scenario 3.

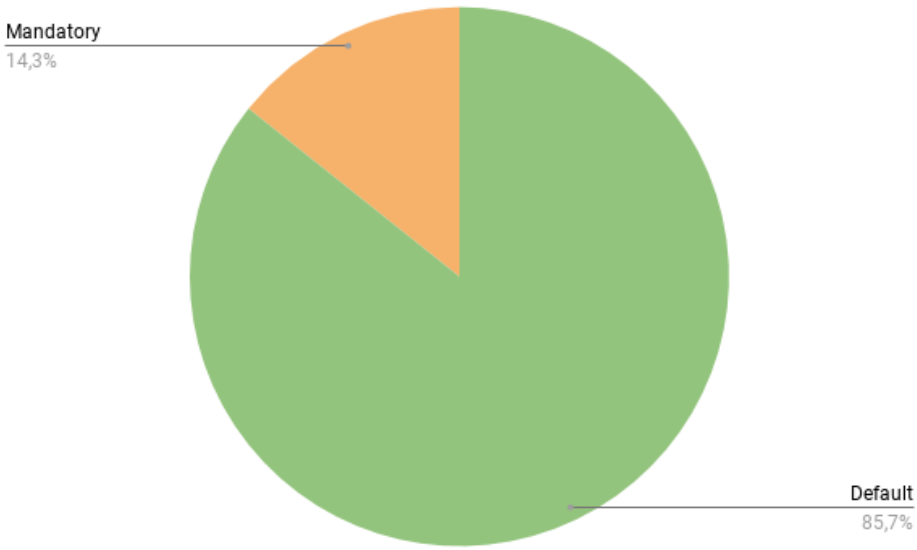


Figure 5.1: The distribution of answers to alternatives for controlling the classification level (Scenario 1).

However, some disadvantages with option 2 were also seen by the participants. Having a default label can potentially prevent the user from making an active decision and forget to change the label when required at a later time. Another disadvantage that was pointed out, is that choosing not to classify is an option and thereby totally avoiding the classification action. One participant pointed out that documents are often sent out to customers that use different classification systems. Having the option of not classifying a document would potentially lead to a large number of

unclassified documents as this will make it easy for the customers to manage the document. Finally, it was noted that option 2 also limits the available screen space.

Only one participant would prefer option 3; having a mandatory classification action when creating a document. The main argument was that this would ensure that the classification was not ignored. It was argued that a mandatory pop-up would not be an annoyance since creating new documents is something that happens infrequently in most users' daily work. Another advantage is that users have to decide on the classification level before adding content to the document. On the negative side, some participants believed that option 3 could be perceived as bureaucratic and irritate people.

Another participant suggested an alternative which is a combination of option 2 and 3. Option 2 would be the best choice, but with having the pop-up from option 3 appearing when the user is prompted to save the document for the first time. The reason was that it could be more clear to the user what classification level to choose after writing some content.

None of the 7 participants preferred option 1; letting the solution recommend a classification level. They did not trust the software to predict the correct classification level. It would also have to be a dynamic scanning process, which could complicate the feature even more. Furthermore, customers have different restrictions and classification policies and an automatic classification feature would not fit well in these situations. Also, conditions change over time and what is confidential today may be open tomorrow. One participant argued that ideally this would be the best option, providing that the solution predicted the correct classification level in all cases. However, this is not very likely, and thus, it would be more of a confusing or annoying feature.

Scenario 2 A possible feature that can be included when implementing AIP is to have a justification pop-up when the user performs a classification level action that is violating the organization's rule set. The purpose of this scenario was to identify if this function was seen as useful and also how it should be implemented.

All but one of the participants responded that they would like the justification function to be included, as illustrated in Figure 5.2. The one participant against it argued that the competence level of people in the organization is high and that one should trust that they make the correct decisions.

It was argued that even though it may be perceived as a barrier it is an important function in the case organization, mainly because of the IT security challenges we face today. Including the function means adding an extra click, but this is manageable. One thing that was suggested to consider was to only apply the function when

changing the classification level from a strict level to a less strict level. Another participant pointed out that having the function was acceptable as long as it does not occur too often. The main concern made was that restricted documents are often sent out of the organization and in these cases the classification level has to be set lower in order for the customer to open it. This means that the process will happen frequently and thus can be perceived as annoying.

In general, the participants expressed the importance of including traceability mechanisms as part of this type of action. However, they had some concerns related to having to provide an explanation for their action.

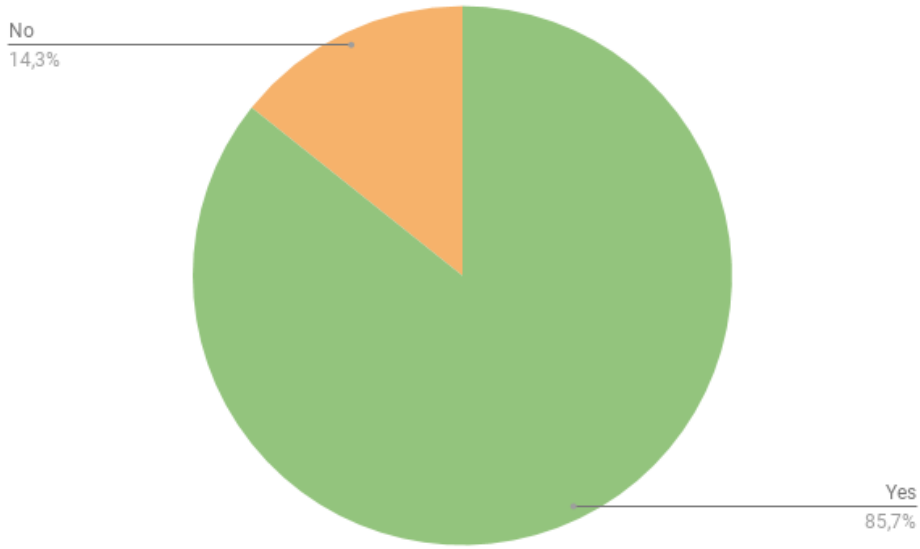


Figure 5.2: The distribution of answers to whether or not the justification feature should be included (Scenario 2).

One participant argued that the main reason for including a mandatory explanation was that it would force the user to reflect on the justification of the action. There should be well thought-out reasons for changing a document's classification level, especially when changing to a lower level, since it may lead to the information being exposed. Furthermore, traceability is an important factor since the organization produces a large number of documents. This applies both to who was accountable and being able to recall justifications made.

Among the skeptical notes made were a fear that the explanation requirement would feel like an intervention in the workflow. Such functions may lead to users

trying to find ways around it, especially if they occur frequently. To avoid this, it is important that the process of changing a classification level takes minimal time.

One participant argued that this functionality would be useless unless it was checked by someone. In order to make speed up the process, it might be tempting for the user to just type *Ok* in the explanation field. It was suggested that the feature could be restricted to templates only.

Based on experience from a previous workplace, one participant said that people often chose to write the same justification each time, for example *Not relevant anymore*. However, this is better than not writing anything and justification text should be short in any case. The user still has actively made a decision and signed it.

Scenario 3 The third scenario was concerned with pre-classified templates. The purpose of this scenario was to identify to what extent the participant would find this useful and also how it should be implemented, for example with regards to document types.

In total 5 out of the 7 participants responded that pre-classified templates was a desirable feature for all types of documents. This is visualized in Figure 5.3. The remaining two participants argued that both notes and reports can have all different levels of classification and that it would be difficult to set a suitable default level.

One of the participants argued that the main reason for having pre-classified document templates is that it ensures there is a consciously chosen level for all new documents and thus avoiding unprotected documents. Another point made was that this makes it possible to apply different default classification levels for different types of documents. However, it is important that a good evaluation has been made for the pre-classification level. One example that was given was the fact that PowerPoint documents are often used to present results and, as such, in many cases are more exposed than Word documents. It was also emphasized that having a pre-set level should not make it difficult to change by the user.

Regarding the choice of default level, it was suggested to have it set to *Internal* for all the templates. Today, many templates are classified as *Open* and some customers have expressed concern about this, believing that it means that the documents can be shared with everyone, including people outside the organization. Having the default level set to *Internal* will increase trust.

One participant, who works with customer projects only, expressed a concern that default template level would mean it had to be changed to a lower level every time it was to be sent outside of the organization. This is similar to what was commented in Scenario 2 and an important issue to be solved. Another concern expressed was

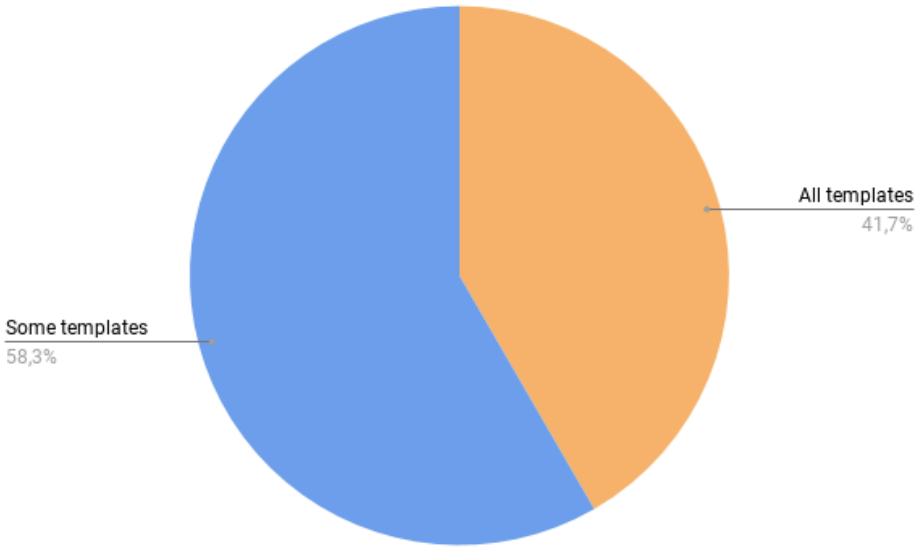


Figure 5.3: The distribution of answers to what extent templates should be pre-classified (Scenario 3).

that pre-classification might make people less aware of classification levels in general and neglect to assess the level in special cases.

Two of the participants suggested a feature where documents are created within a context of a project and inherits the project's classification level. In cases where there is a mismatch between the classification of the document and the project, there should be a pop-up warning addressing this.

Scenario 4 As most of the participants are working with customer projects, a scenario related to how to handle this situation was presented. As shown in Figure 5.4, 3 out of 7 participants did not want to add customer's classification schemes in the solution. They do not regard this as practical and considered that assessments and responsibilities should be left to the user. It would be unrealistic to trust that technology supports all aspects of these cases. In addition, two participants stated that adding more toolbars or command buttons could reduce screen space even further.

Four suggestions on how to integrate the customers' classification schemes in the AIP solution came up during the interview. One was that one should show the customer schemes in the same bar but split with the organization's classification

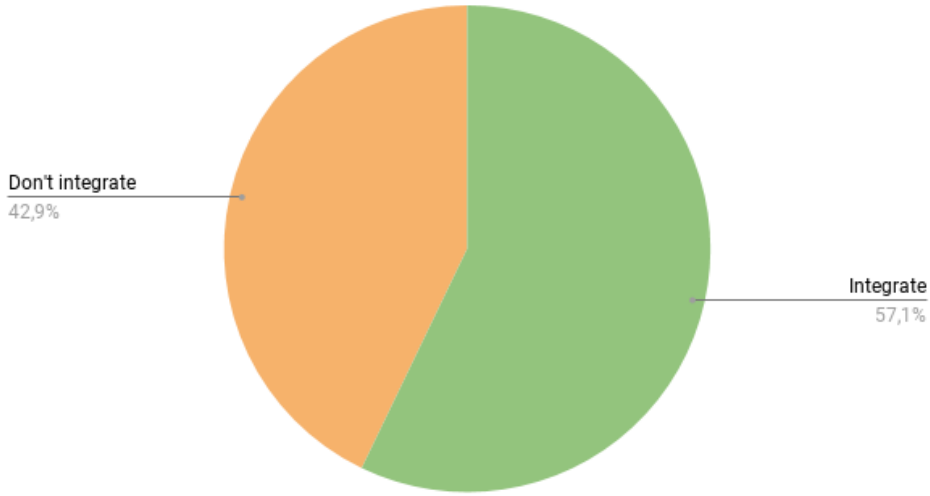


Figure 5.4: The distribution of answers to whether customers' classification scheme should be integrated into the AIP solution (Scenario 4).

scheme on one side and the customers on the other.

The second suggestion was to add a separate toolbar with the customer's classification scheme. This would make it easy to map the organization's scheme with the customer's scheme and identify the differences, e.g. the organization's highly confidential label could correspond with the customer's confidential label. In some cases, a negotiation regarding classification took place at the beginning of the project. This could result in the customer accepting the organization's classification scheme or modifying their definitions. Having both schemes available in the solution will improve the internal communication regarding classification.

A third suggestion was to add an additional label named *Customer Internal* in the organization's classification toolbar. This label should correspond with a highly confidential level. The reason for this is that employees may not have enough knowledge to assess what the customers consider confidential, for example trade secrets.

The fourth suggestion was to include a label with the customer's name to the toolbar, indicating that other classification schemes are relevant. Today it is often hard to identify documents that have additional classification schemes and a customer

specific label would be helpful in this situation.

One of the reasons mentioned for not integrating the customers classification schemes, was that it would be too complex as one has to keep the system in sync with changes in the customer's scheme. In addition, one would have to classify every document twice; both the customer's template and the AIP toolbar, which is a cumbersome process. Another participant stated that: *The solution is easy for people to learn. I think one should use the system as a basis and in projects say that additional conditions are required. It is much simpler to implement a regime like that than adding additional functions to the solution.*

It was also stated that in some cases it is sufficient to discuss with the customers whether or not the organization's labels are acceptable and hopefully come to an agreement. In other cases, customer specific documents are uploaded remotely on to a shared server without any classification scheme being applied. How customer documents are handled to a large extent depend on the project being worked on.

Scenario 5 This scenario was concerned with different aspects related to how AIP protection should be applied to emails. It consisted of three email related functions that the participants were asked to evaluate.

5a: Managing confidential content in email Given the fact that the organization does not allow emails with content classified as *Highly Confidential* to be sent, the participants were asked if this label should still be visible. All of the participants wanted to be visible, mainly because removing it would be confusing. It was stated that showing it would also work as a reminder and prevent the user from classifying incorrectly. More than half of the participants commented that they think there might be a risk that the users would try to circumvent the restriction and classify lower in order to be able to send the email. However, one of the participants added that this may not be very likely, and something not to be too concerned with, given that all employees have agreed to comply with the organization's rules.

Regarding what action the system should take in cases when the user tries to send an email labelled as *Highly Confidential*, all participants except one said it should be prevented from being sent. The reason why one participant did not think the system should stop an email labelled as *Highly confidential*, was that attachments were regarded as the most critical, while daily communication and clarifications is in the email text. The participant suggested that one should be able to send emails classified with any level, assuming that the content is encrypted. However, if there are attachments classified as *Highly Confidential*, the email should be prevented from being sent.

As presented in Figure 5.5, 4 out of 7 participants preferred that a pop-up warning

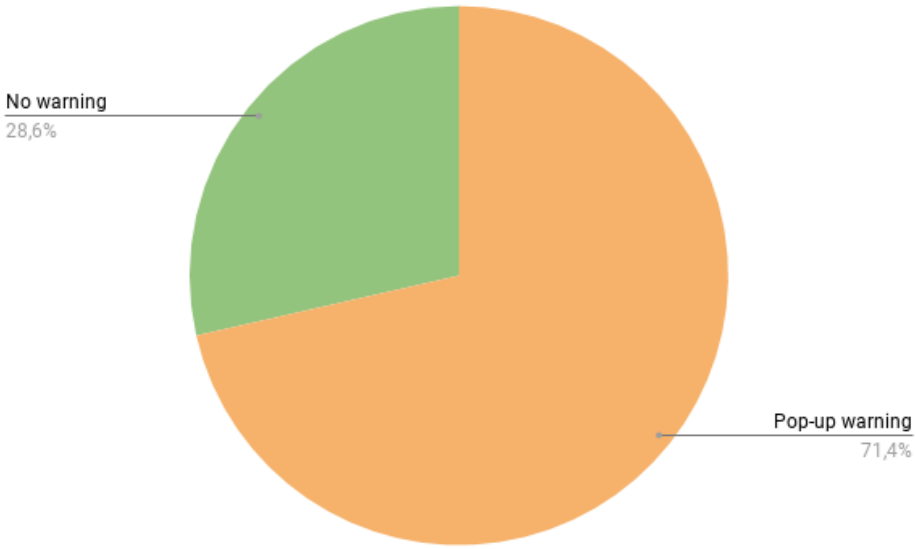


Figure 5.5: The distribution of answers to whether a pop-up warning should be displayed in cases where the email was classified as *Highly Confidential* (Scenario 5a).

was displayed in cases where the email was classified as *Highly Confidential*. One of the participants suggested that the pop-up message should contain information about what procedure to follow, and maybe who to contact for clarification, to be able to send *Highly Confidential* information by email. This would ensure that there is less risk of violating the organization's classification policy. Another suggestion that was made was to make the send button inactive when the *Highly Confidential* label was selected.

The *Highly Confidential* label also shows a small tooltip icon that the user must move the mouse over to see additional text about the restrictions that apply. One of the participants commented that this function was somewhat subtle and that it would be better if a short text was always showing.

5b: Managing Different Classification on Email and Attachments This scenario was used to explore what would be the most desirable response from the solution in cases where the email and the attached document classification level differs. As visualized in Figure 5.6, 5 out of 7 participants preferred that a pop-up warning was shown only when the attachments had a higher classification level than the email. The main reason for this was to make the users aware of the situation and

force an action. Still, one participant did not see a need for the system to prevent the email from being sent, as long as it was a conscious action by the user. It was also expressed concern that too many warnings will make users ignore them.

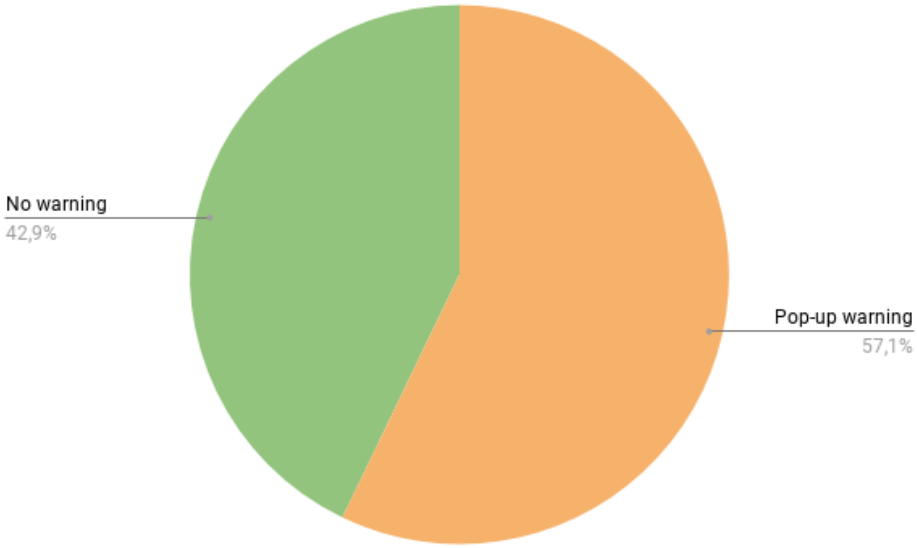


Figure 5.6: The distribution of answers to whether a pop-up warning should be shown only when the attachments have a higher classification level than the email. (Scenario 5b).

Two participants suggested that the classification level for the email was automatically changed to match the highest level of the attachments, possibly also including a system warning about this. One participant would prefer that sending an email with attachments classified higher should be prevented automatically. Another participant suggested that this should only happen in cases where the attachments are classified as *Confidential* or higher.

It was also stated that, in general, emails should not contain sensitive information. If possible, alternatives, such as cloud services, should be used. It was added that the most important factor to avoid data leakage is a good organization security culture.

5c: Managing subject and file names This scenario was used to examine whether and how the solution should manage the email subject field content. In total 6 out of 7 participants did not want the solution to check the subject field. This is illustrated in Figure 5.7. One participant would like to have this feature, but it should be a yellow message bar reminding the user to check the content in the

subject field and the names of attached documents when working with emails and attachments labelled as *Confidential* or higher. One participant suggested that the solution could rename attached documents based on an organization convention.

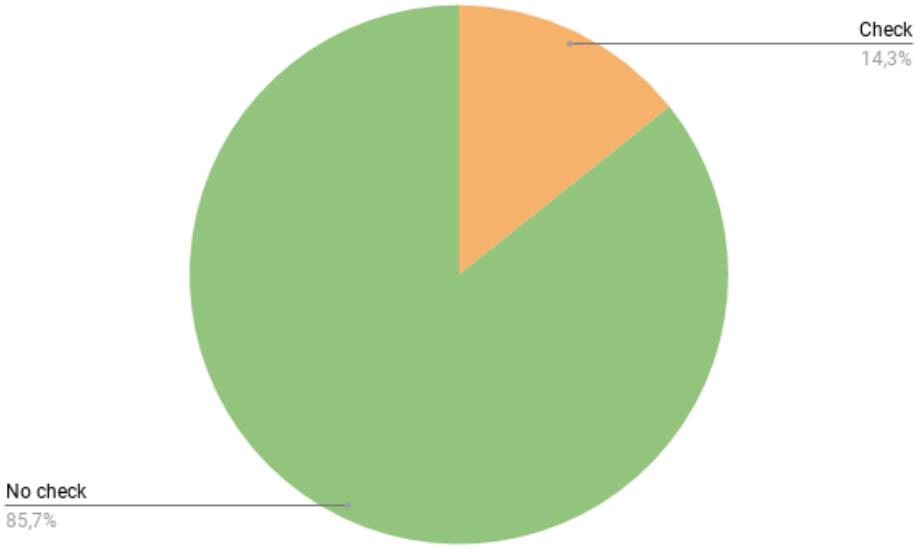


Figure 5.7: The distribution of answers to whether the solution should check if the subject field includes sensitive information (Scenario 5c).

Arguments given for not performing a check and showing a warning:

- The feature could be annoying as it may appear often
- The scenario is not perceived a problem today
- When relevant, it must be left for the user to handle
- Organization policies should be sufficient assuming employees are aware of them. A security culture must be established in parallel
- Generic attachment names are preferable. It is the receiver's responsibility to rename if desirable.

Two of the participants stated that a warning or pop-up would be suitable only in cases where the email is labelled as *Highly Confidential* or *Confidential*. One of them emphasized that one could compare this to a risk matrix; assess the likelihood and

consequences of events, where you put most attention to cases where the consequences for the company are highest.

Table 5.1 summarizes the findings. It shows the distribution of choices for the questions asked following the presentation of each scenario. Overall, the participants agree on most scenarios, except Scenario 4 and Scenario 5a.

Scenario 1		**Scenario 2**		**Scenario 3**		**Scenario 4**	
Recommendation	0	Optional	0	All templates	5	Integrate	4
Default	6	Mandatory	7	Some templates	2	Don't integrate	3
Mandatory	1						

Scenario 5a		**Scenario 5b**		**Scenario 5c**	
No pop-up	0	No pop-up	0	No pop-up	6
Pop-up	4	Pop-up	5	Pop-up	1
Prevention	7	Prevention	3	Prevention	0

Table 5.1: Summary of results from the usability test.

5.1.2 Potential barriers in the solution

The following results from the post-test interview is concerned with potential barriers in the solution that were identified by the participants. As shown in Figure 5.8, 5 out of 7 participants believe there is a need for the solution and are positive about using it. As one of the managers expressed: *If this solution is implemented correctly, one will get a more visual picture of the risk associated with the document. This is especially useful if the solution can integrate the customer's security labels. As such, it will mark a new chapter in the organization regarding security in our daily work. I believe it will be an eye-opener for a lot of people.*

Those who expressed skepticism stated that they would only use it if it was enforced by establishing a common policy. One of them stated: *The list of stupid policies in this organization is long, but I have to comply with them. I will use the solution if it is implemented, but I hope it won't be. People think there are too many rules already and this can be a contributing factor for them to leave. The solution is bureaucratic.* However, after using the solution for a week the participant cited above sent an email expressing a change of mind: *I was relatively critical to the solution in the interview, but I installed it. A colleague disagreed with my opinion and said the solution was a good idea. Now that I have used it for a couple of weeks, I think it may work. It is fairly easy to use, and it makes me think more about who I am writing to and what I am writing. However, I have defined for myself that Internal also in practice means Project Internal. I must be able to share documents with people*

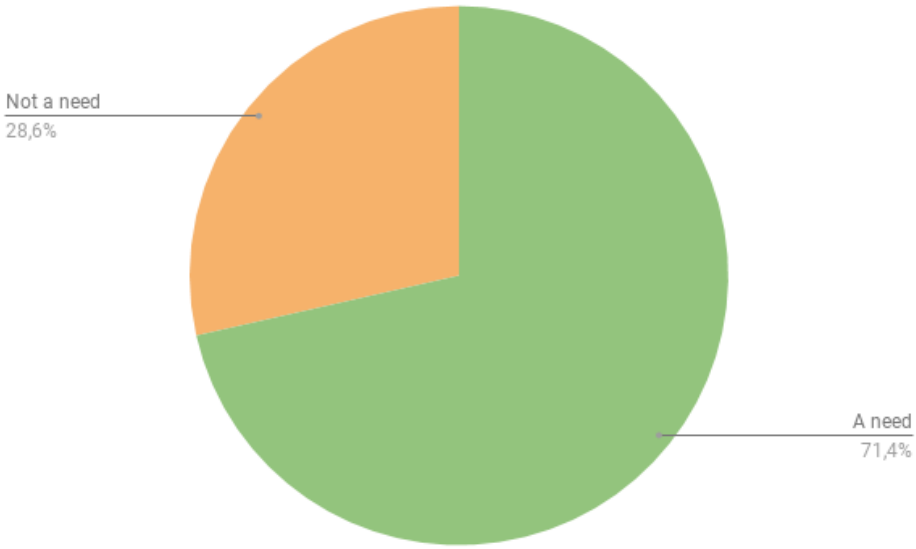


Figure 5.8: The distribution of answers to how the need for the solution is perceived among the participants.

that we have a formal project collaboration with without defining it as Open. I must be able to share documents classified as Internal with customers. Or else we will need another category (which has some disadvantages). The participant's concern regarding how to involve customer classification is discussed in Section 6.1.3.

Even though the participants expressed that they were positive to use the solution, they identified a number of potential barriers. Below are two interview quotes from participants expressing their concerns with the solution: *My only concern is that it becomes too rigid and that you often have to close boxes to make them disappear. If the focus is to remove the boxes instead of thinking about the information you are managing, the purpose behind it is gone. However, so far, I am satisfied with the solution*, and *The solution looks feasible at first sight, but my fear is of course that it creates more work. It all depends on how strict it is implemented.*

The barriers most frequently mentioned in the interviews are listed below:

- Enforced features can be perceived as bureaucratic
- Frequent pop-ups can be frustrating

- Lack of verification and trust in the solution may cause false safety
- Little or no consequence of performing actions causing indifference to the use value
- Too many restrictions make users consider it rigid
- It does not work in OneNote. Office programs are inconvenient
- Reduces the available screen space
- Too much technology relative to the size of the problem
- The current security policy is sufficient
- Reduced system performance
- No perceived added value in the daily work

5.1.3 Classification routines and awareness

The possible effects the DLP solution may have on the employees' classification routines and awareness were investigated. The results provide a good basis for answering *RQ2* and its associated hypotheses.

In total, 4 out of 7 participants responded that they consult someone when they are unsure about what to classify a document. This is visualized in Figure 5.9. One participant said that it is natural to consult the project manager if the document is project related and people working in Human Research if the document concerns personal information. It was also said that the client or the owner of the information was the right person to ask if the confidentiality level was unclear. By experience this situation happens frequently.

The remaining participants stated they would select to classify a document at a high level in cases where they were unsure. One participant reasoned that, although being aware that this would add restrictions to the document, it was still the easiest option. For example, by applying a higher label than needed, a document can be more difficult to share at a later stage. One participant argued that cost of classifying a document too high is smaller than the cost of classifying it too low.

The effect of the DLP solution on routines and awareness was explored. As shown in Figure 5.10, all participants except one believed that the solution would improve

Figure 5.9: The distribution of answers to what the participants do when unsure about what classification level to apply (H1).

their own and other employees' policy awareness and classification routines. One participant said that people are aware of the different classification levels, but they might struggle describing them. In that sense, this solution will work as a reminder. Another participant stated that the solution will result in more documents being classified. It was also stated that with the solution you must make an active decision and be more explicit when classifying.

It was also emphasized that emails are often sent uncritical and that this solution will contribute to the employees' awareness when sending emails. In addition, you do not have to check policy documents when unsure, since the solution provides a short description of the classification levels when hovering over the toolbar labels. It was also argued that the solution will help save time when classifying documents. One of the participants stated: *I think the biggest advantage with this solution is that it brings the classification policy closer to the user.* Another statement made was: *If implemented correctly it will give a good visual indication of the risk related to the document.*

The one participant who was skeptical regarding increased awareness stated: *Educating people by implementing a software solution could work, but this is not sufficient enough. It solves problems that barely exist.*

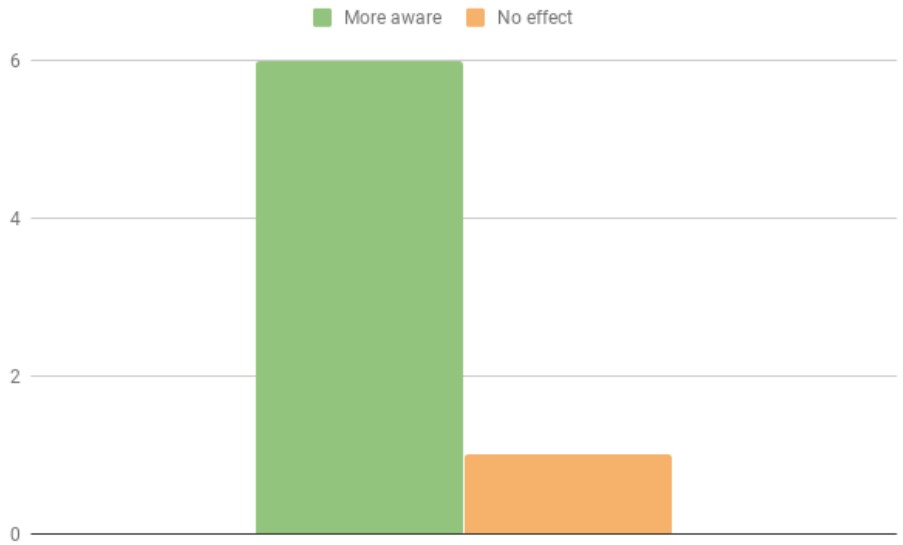


Figure 5.10: The distribution of answers to whether the DLP solution will increase the employees' awareness of the organization's classification policies (H2).

One of the participants argued that policy awareness is more related to organizational culture. It was argued that the case organization has a culture that is hard to change. This might be a challenge. This is supported by the following interview quote: *I would say there are two schools. In my department, which has worked a lot with the industry, it is a good discipline to comply with classification procedures. Other departments are more idealistic and academic, believing that everything should be open and published and one should do what is best for the society.*

The difference between managers and non-managers classification regarding classification routines and awareness was also investigated. All the non-managers and three quarters of the managers state that they know the organization's classification scheme. One of the managers, who was hired recently, claimed to not yet know the scheme well. This is shown in Figure 5.11. However, the person was familiar with classification schemes from a previous job assignment, especially when related to personal employee data.

One employee admits that the familiarity with the scheme should have been better. It was read 8 years ago at the beginning of the employment. Today, he believes he knows what is correct but has to check it when relevant. He always does an assessment of the document when it is created and when it is finished, so ensure

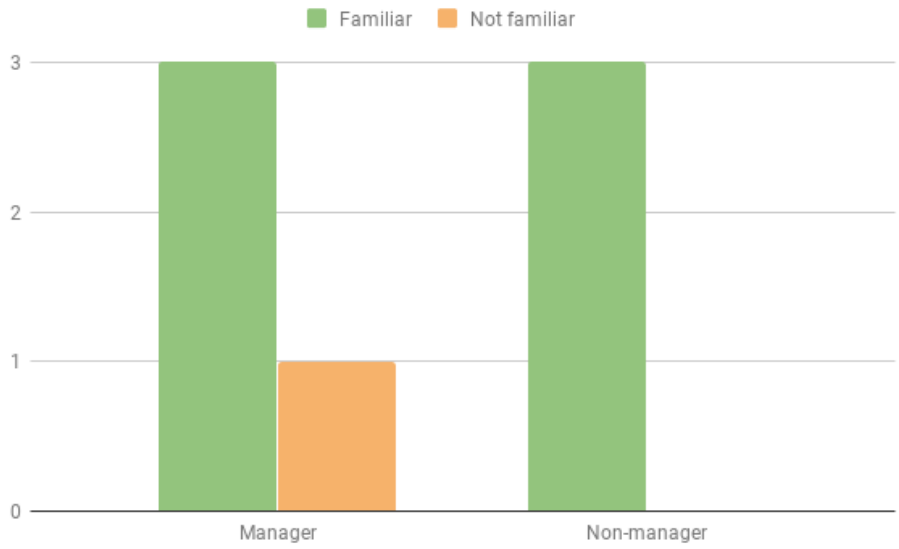


Figure 5.11: The distribution of answers regarding familiarity with the organization's classification policy for managers and non-managers (H3).

no mistakes are made. Another employee claims to know The Security Act[1] better than the organization's policy, but still feels that he knows the organization's policy fairly well. A point made by one of the managers: *People classify documents based on experience and this is what really decides whether or not a document becomes Confidental or Open in different projects.*

Two of the managers state that they classify documents often, while one manager does not, since he is not working with projects. Regarding what types of documents they classify, one of the managers mainly classifies templates, while one of the others mainly classifies notes and reports and most often when they are to be distributed.

All the managers said that they trust the employees to handle and classify documents correctly. None of them perform audit control. Some of the managers expressed that they were more concerned with protecting data that includes personal information about their employees. One manager said that in cases where information could be confidential it is discussed with the employees.

None of the employees feel that their manager imposes requirement regarding

[1] "Lov om nasjonal sikkerhet (sikkerhetsloven)", LOVDATA, accessed April 10, 2019, https://lovdata.no/dokument/NL/lov/2018-06-01-24

classification on them. However, the managers said they would act if they discover violations of the policy.

5.1.4 Actions the organization can take to ensure a successful implementation

At the end of the interviews, the participants were asked if they had suggestions or ideas for how the organization should implement the solution to make it successful and what could be done to strengthen classification routines in general.

One of the participants believed that implementing a DLP solution would be the simplest way to improve classification routines. Just distributing an updated security rules and procedures manual would not be effective in the long run. However, the solution must be thoroughly tested first as there are often bugs in these kinds of solutions.

Another participant expressed that implementing the DLP solution in itself would not have a large effect on the employees' classification routines: *The only thing that will improve classification routines is the managers engagement and how the importance of security is explained and conveyed to the employees. People will see the importance of it if they understand that it has value.*

Several participants suggested that a security campaign should be used to introduce the DLP solution to the employees. Especially, since it would increase the security awareness among the employees and make it easier to understand the motivation for implementing the solution. One participant stated: *It can be reasonable to remind people on a regular basis about the classification routines and what the different levels mean. There have been information security campaigns in the organization before and I believe we should continue with that. One can never be fully trained. Information is perishable, people forget, and the rules change. The solution should be introduced through a campaign and not just by adding an extra button in Outlook. I think sending out a video or electronic information could be sufficient.* Another participant stated that: *There should be a security campaign as part of full a communication strategy. One should consider if the solution replaces something existing that is more unwieldy. When you introduce something new you should remove something else. This is a balance that a communication plan should focus on.*

5.2 Results from the questionnaire

The questionnaire can be found in Appendix C. The results are not presented in chronological order, but instead according to their relevance with respect to themes related to the research questions. These are; applying a security tool, classification routines and awareness, and actions the organization can take to ensure a successful implementation. The questionnaire was active for two weeks and in total 36 responses were received. The number of responses for each question was 30 or higher. As shown in Figure 5.12, only three respondents have a personnel management role.

Some of the statements have answer alternatives *Strongly agree*, *Agree*, *Neither agree nor disagree*, *Disagree* and *Strongly disagree*. To better relate the results to the research questions and hypotheses, in some cases the answer alternatives *Strongly Agree* and *Agree* will be combined. This also applies to the alternatives *Disagree* and *Strongly Disagree*.

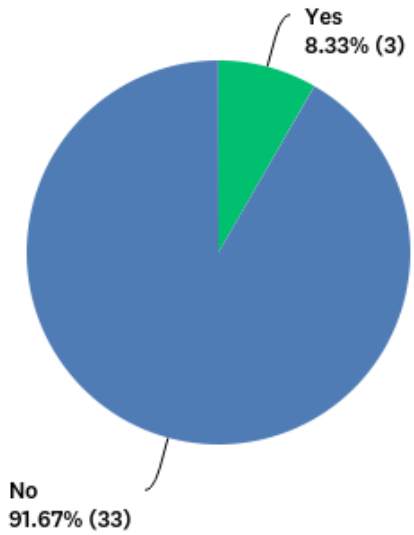


Figure 5.12: The distribution of answers to Question 1; *Do you have a personnel manager role?*

5.2.1 Applying a security tool

The section shows the answers to questions and statements related to possible implementations of the DLP solution, its usefulness and ease of use. The results provide a good basis for answering *RQ1*.

Question 15 *How do you perceive the toolbar in Word?* As seen in Figure 5.13, the majority of the respondents found the location and size of the toolbar appropriate. In addition, near about half of the respondents said that it was easy to understand. 20% of the respondents expressed other concerns, such as that the toolbar occupied too much vertical space.

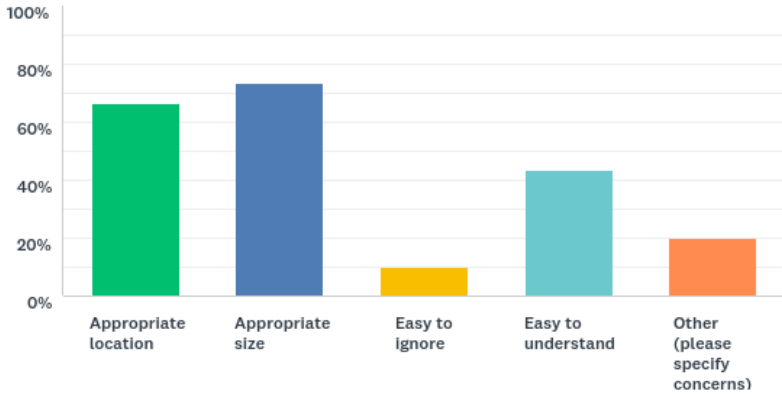


Figure 5.13: The distribution of answers to Question 15 which was concerned with the perception of the toolbar.

Question 16 *How should the solution handle the situation where a user wants to send Highly Confidential content by email.*? As there is a policy in the organization stating that *Highly Confidential* content should not be sent by email, it was desirable to examine how the respondents think the solution should handle this situation. More than half of the respondents answered that the solution should show a warning with options and prevent the email from being sent. Figure 5.14 shows the results.

Question 17 *Is the justification feature usable?* The question was concerned with the usability of a justification box when the user changes the classification to a lower level. As shown in Figure 5.15, about two-thirds of the respondents find this feature usable.

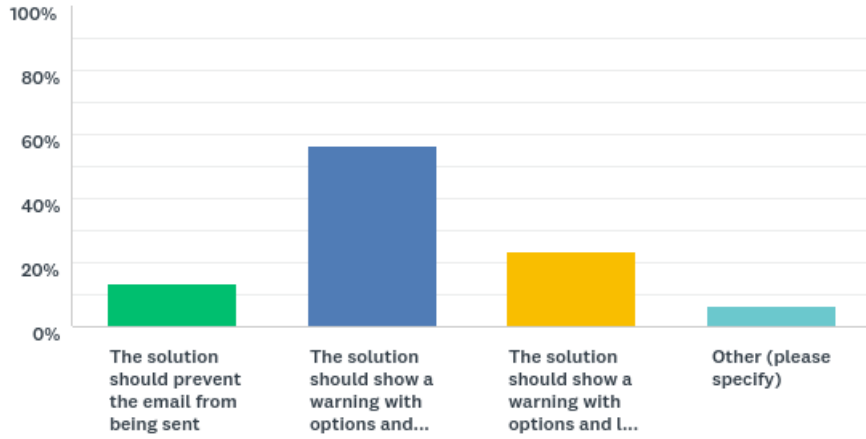


Figure 5.14: The distribution of answers to Question 16 which was concerned with how the solution should handle the situation where a user wants to send *Highly Confidential* content by email.

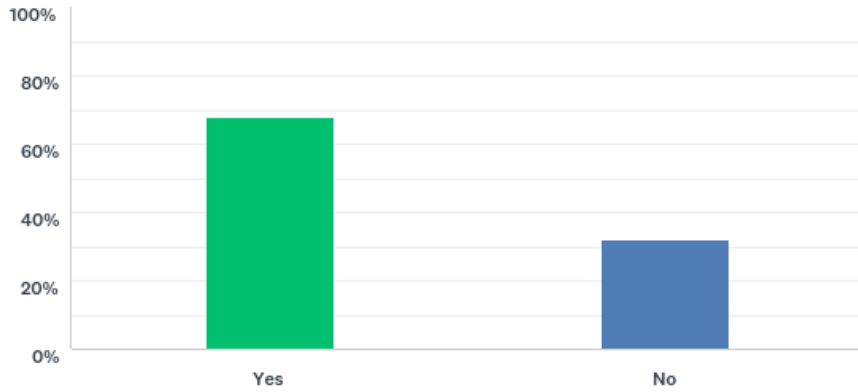


Figure 5.15: The distribution of answers to Question 17 which was concerned with the usability of the justification box.

Question 18 *Given that the justification feature above is implemented in the solution, in what cases should it be used?* The two alternatives with the highest, and almost equal score, were *Only when changing the label from the two most confidential levels to a lower classification level* and *Only when changing to a lower level.* This means that over 60% of the respondents believe the justification feature is most

relevant when changing to a lower level. The result is presented in Figure 5.16.

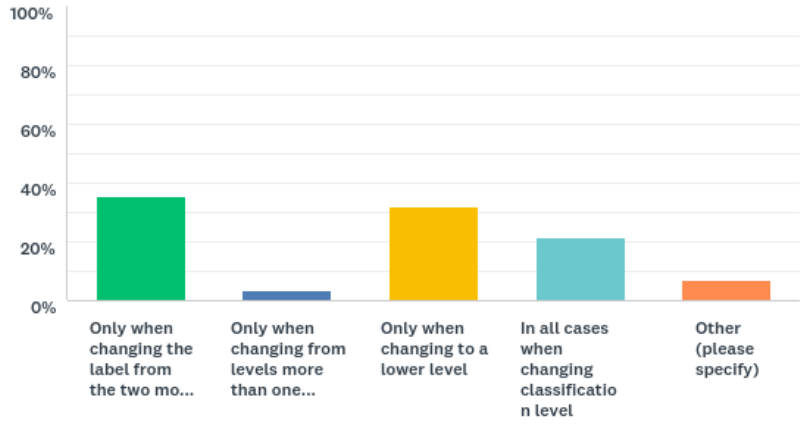


Figure 5.16: The distribution of answers to Question 18; *Given that the justification feature above is implemented in the solution, in what cases should it be used?*

Statement 19 *I believe that the security tool should be applied to ....* Figure 5.17 shows the result from the statement regarding what document types should be classified. It is clear that the majority believe that both emails and documents should be classified.

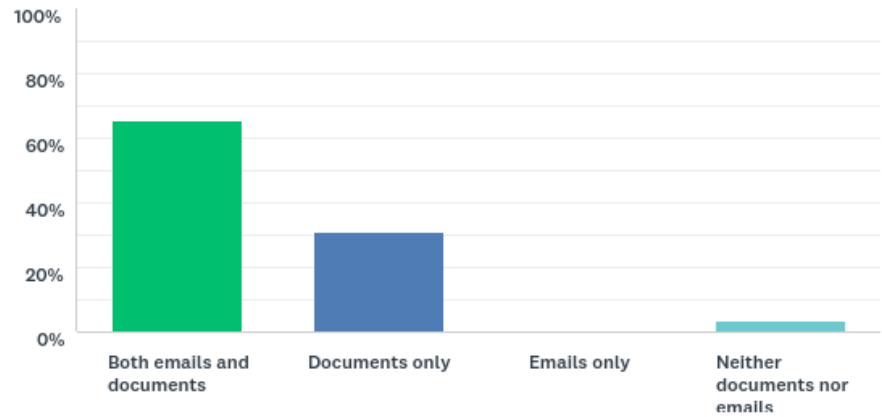


Figure 5.17: The distribution of answers to Statement 19; *I believe that the security tool should be applied to ... .*

Statement 20 *The security tool seems clear and understandable.* As shown in Figure 5.18, approximately 70% of the participants agree with this statement.

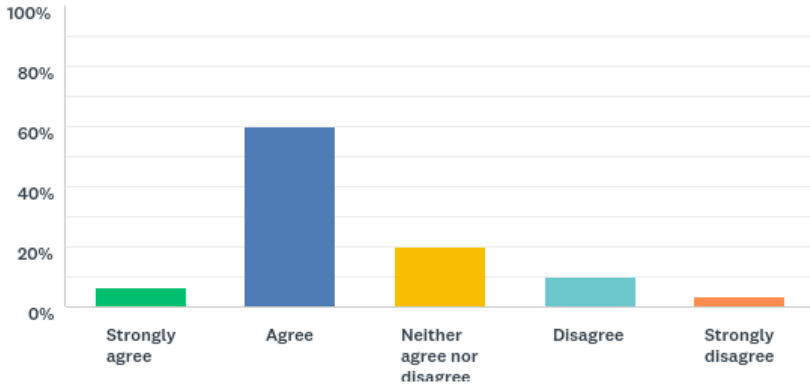


Figure 5.18: The distribution of answers to Statement 20; *The security tool seems clear and understandable.*

Statement 21 *Using the security tool will require low effort.* As shown in Figure 5.19, more than 60% of the participants agree with this statement.

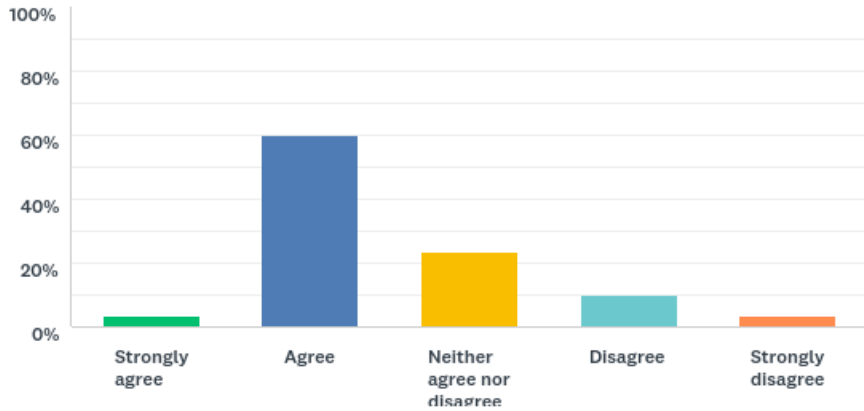


Figure 5.19: The distribution of answers to Statement 21; *Using the security tool will require low effort.*

The answers to the following statements and question shed light on the respondents' perception of the usefulness of the solution and potential barriers.

Statement 24 *The security tool will decrease my job productivity.* As visualized in Figure 5.20, the two alternatives with the highest, and almost equal score, were *Neither agree nor disagree* and *Disagree*. In total, these alternatives were chosen by more than 80% of the respondents.

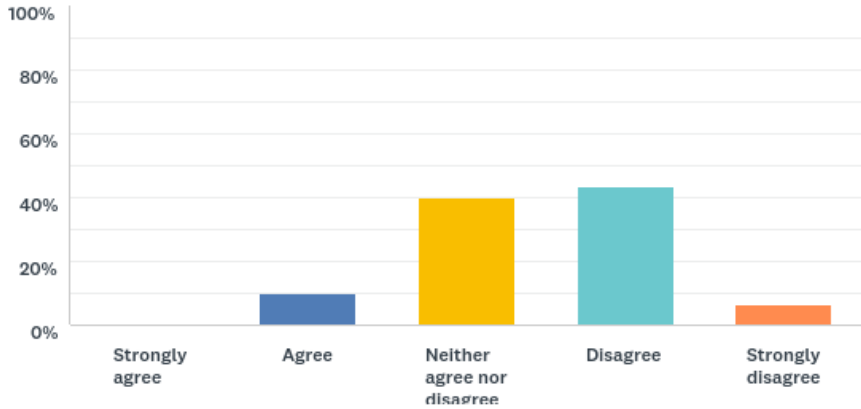


Figure 5.20: The distribution of answers to Statement 24; *The security tool will decrease my job productivity.*

Statement 25 *This tool will be useful in my job.* More than half of the respondents agree to the statement. Nearly all the other respondents chose the alternative *Neither agree nor disagree*. The results are illustrated in Figure 5.21.

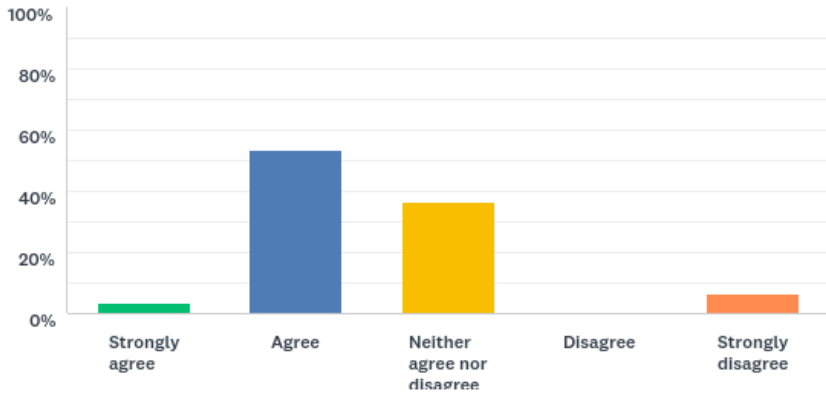


Figure 5.21: The distribution of answers to Statement 25; *This tool will be useful in my job.*

Question 28 *What factors would prevent you from using of the tool?* The respondents were allowed to choose multiple alternatives. In descending order, the five most chosen options were *Lack of information from the organization*, *Lack of enforcement from the organization*, *Annoying pop-ups*, *Unnecessary features* and *Lack of experience*. The result is shown in Figure 5.22.

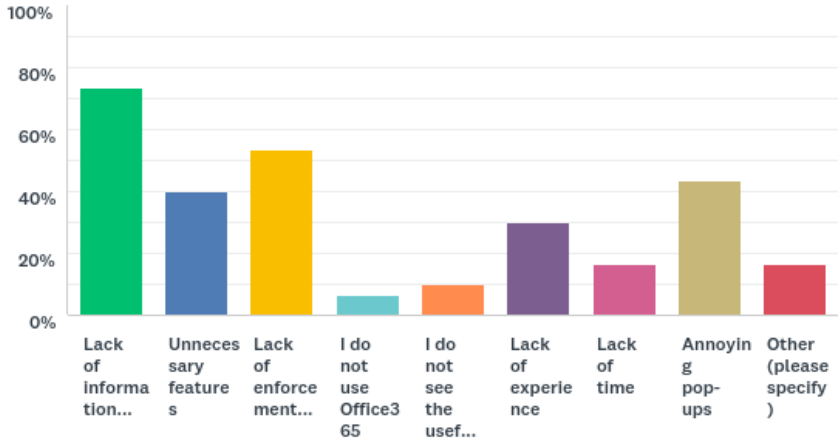


Figure 5.22: The distribution of answers to Statement 28; *What factors would prevent you from using of the tool?*

.

5.2.2 Classification routines and awareness

This section contains answers to questions and statements that made the employees reflect on their classification routines today (statements 2 through 9), and also about how they believe their routines and awareness will be affected by implementing the DLP solution. The results provide insights that can be used to answer *RQ2*.

Question 2 *What kind of documents do you classify today?* The respondents were allowed to choose multiple alternatives. The two most chosen options were *Customer specific documents* (81 %) and *Documents internal to the organization* (61 %). Only 14 % answered they did not classify any documents. The result is shown Figure 5.23.

Statement 3 *I must meet additional classification requirements to the general requirements for the organization.* As presented in Figure 5.24, nearly half of the respondents agree with tis statement. However, a significant and almost equal number of respondents chose the two alternatives *Neither agree nor disagree* and *Disagree*.

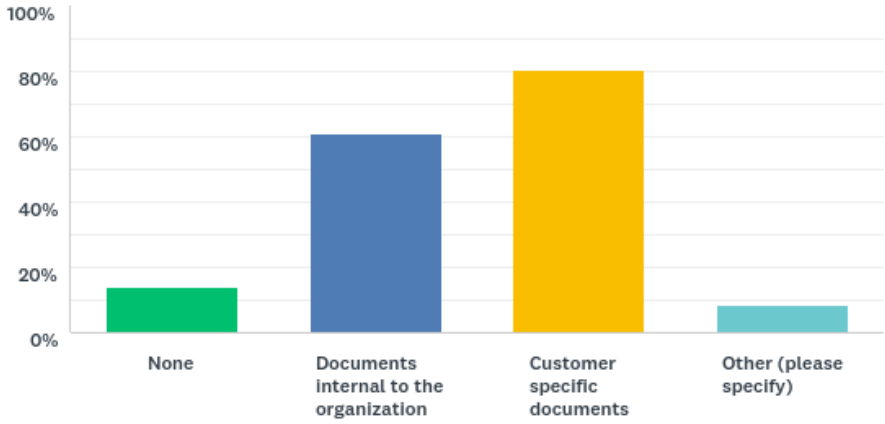


Figure 5.23: The distribution of answers to Question 2; *What kind of documents do you classify today?*

.

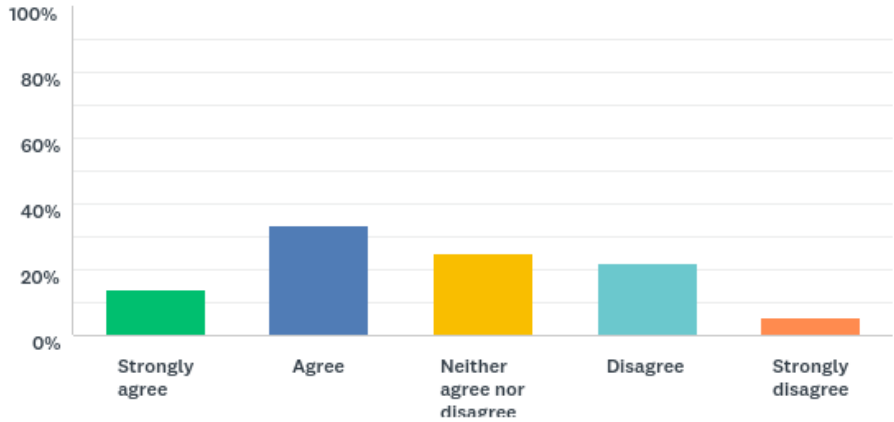


Figure 5.24: The distribution of answers to Statement 3; *I must meet additional classification requirements to the general requirements for the organization*

.

Statement 4 *I work with projects that may be exposed to information security risks, such as malicious attacks and industrial espionage.* Almost two-thirds of the respondents agree to this statement. A significantly amount chose the alternative

Neither agree nor disagree. The result is shown in Figure 5.25.

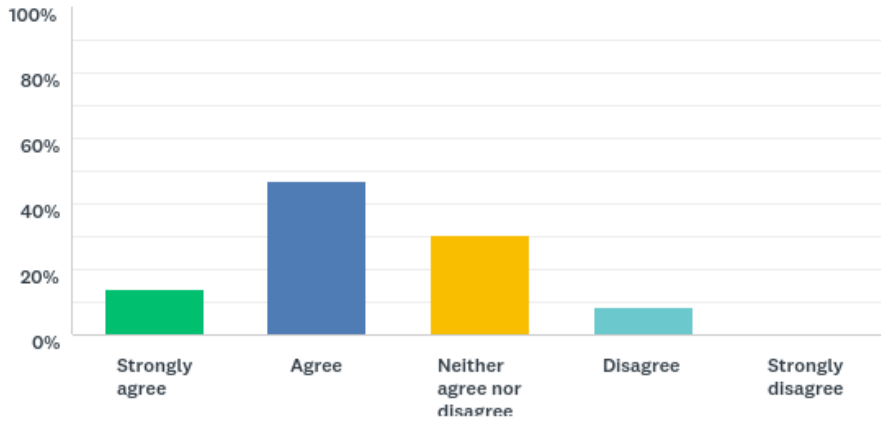


Figure 5.25: The distribution of answers to Statement 4; *I work with projects that may be exposed to information security risks, such as malicious attacks and industrial espionage.*

.

Question 5 *Does or would working with projects exposed to information security risks affect your awareness regarding information security and the organization's security policy?* As shown in Figure 5.26, almost all respondents agree with this statement.

Question 6 *How often do you classify documents?* The result show that most of the participants classify documents *Less than once a month* or *A few times a month.* The result is illustrated in Figure 5.27.

Statement 7 *I believe classification of documents is important.* As shown in Figure 5.28, nearly 100 % of the respondents agree to this statement.

Statement 8 *Getting work done fast has a higher priority than following the security policy.* As shown in Figure 5.29, a significant majority of answers disagreed with this statement.

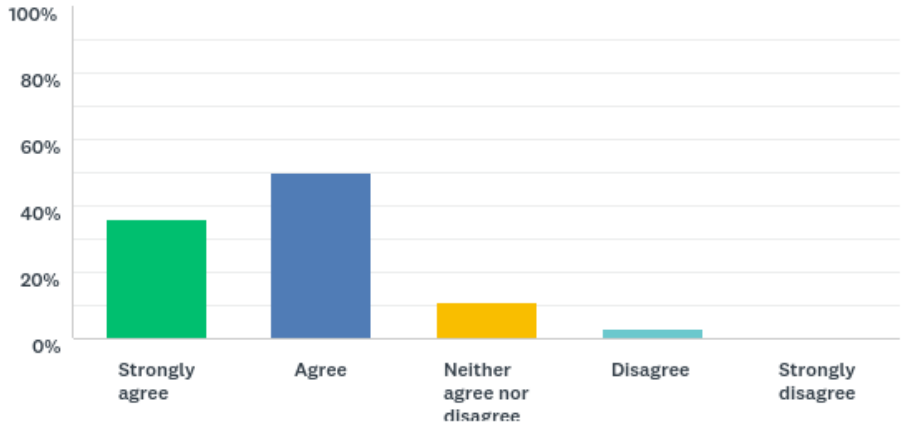


Figure 5.26: Statement 5; *Does or would working with projects exposed to information security risks affect your awareness regarding information security and the organization's security policy?*

.

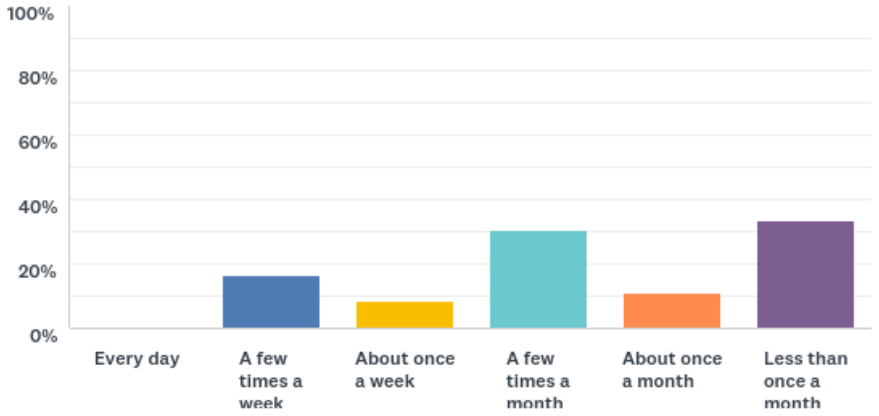


Figure 5.27: The distribution of answers to Statement 6; *How often do you classify documents?*

.

Question 9 *How often do you create documents based on the organization's templates?* As visualized in Figure 5.30, more than half of the respondents answered that they *Usually* create documents based on the organization's templates and about 20 % responded that they *Always* do.

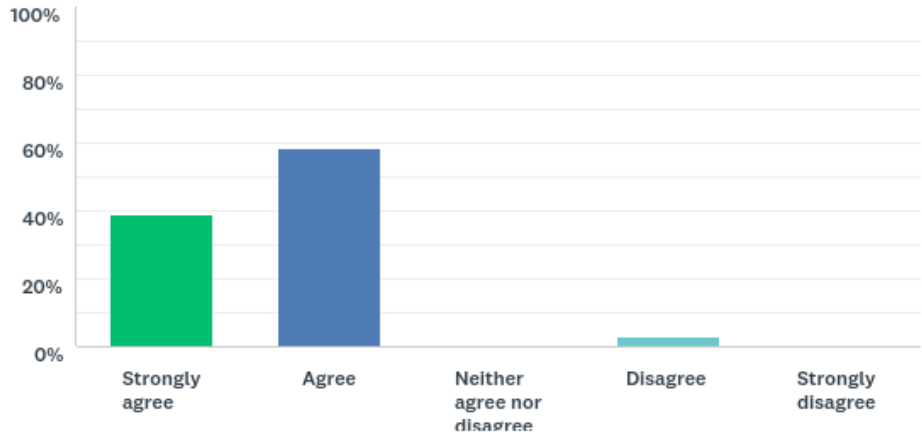


Figure 5.28: The distribution of answers to Statement 7; *I believe classification of documents is important*

.

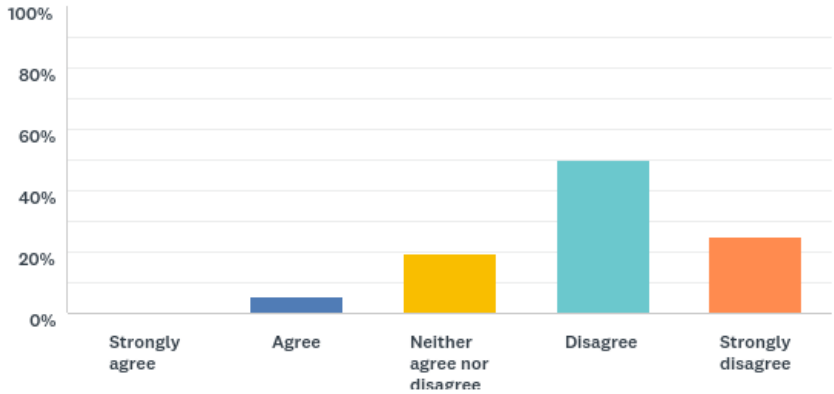


Figure 5.29: The distribution of answers to Statement 8; *Getting work done fast has a higher priority than following the security policy.*

Statement 10 *I am familiar with the organization's classification policy.* As shown in Figure 5.31, more than 80 % of the respondents agree with the statement.

Statement 11 *I am aware of the consequences of classifying wrong.* More than two-thirds of the respondents agree to this statement. Only 10 % responded that they disagree. The result is show in Figure 5.32.

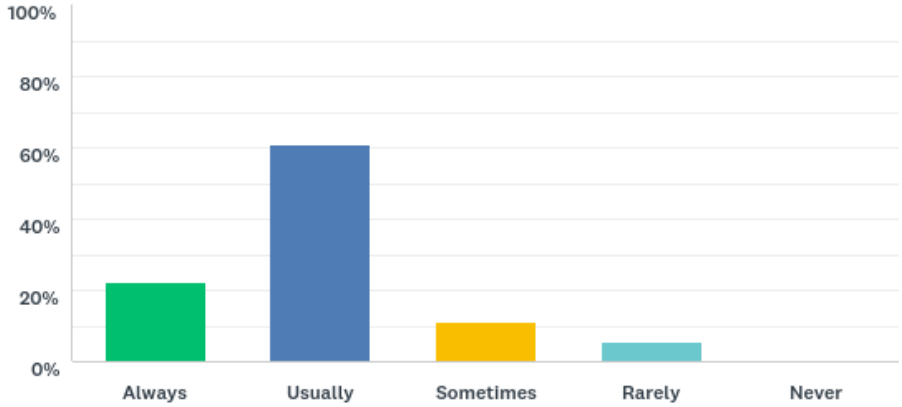


Figure 5.30: The distribution of answers to Statement 9; *How often do you create documents based on the organization's templates?*

.

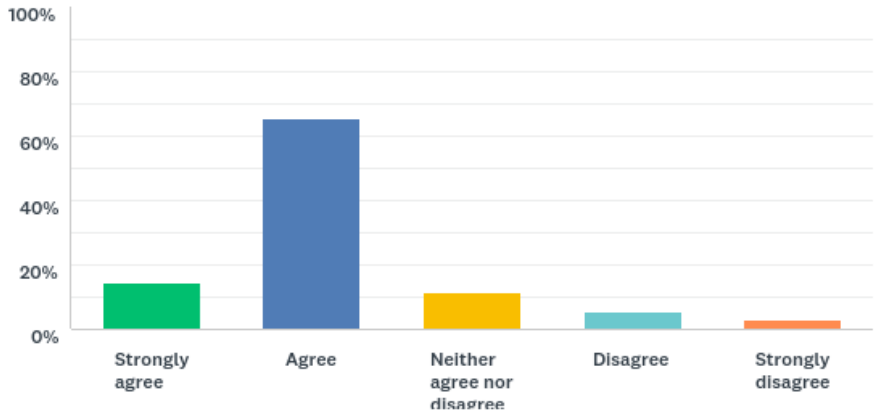


Figure 5.31: The distribution of answers to Statement 10; *I am familiar with the organization's classification policy*

.

Statement 12 *I am often unsure about which classification level to apply.* Almost half of the respondents disagree with this statement. However, about 30 % of the respondents are often unsure about which classification level to apply. Figure 5.33 shows the results.

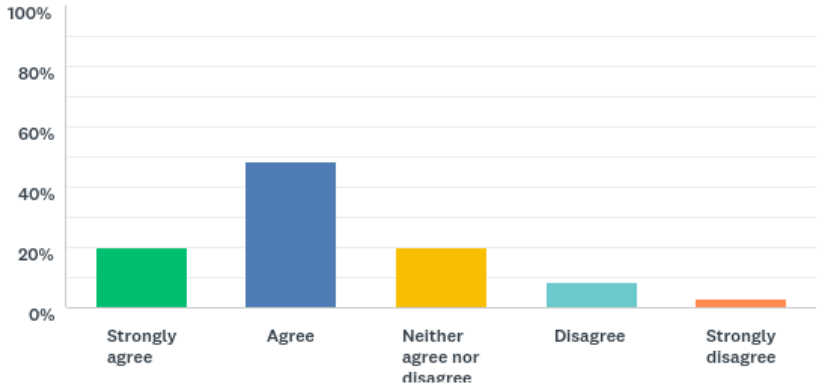


Figure 5.32: The distribution of answers to Statement 11; *I am aware of the consequences of classifying wrong*

.

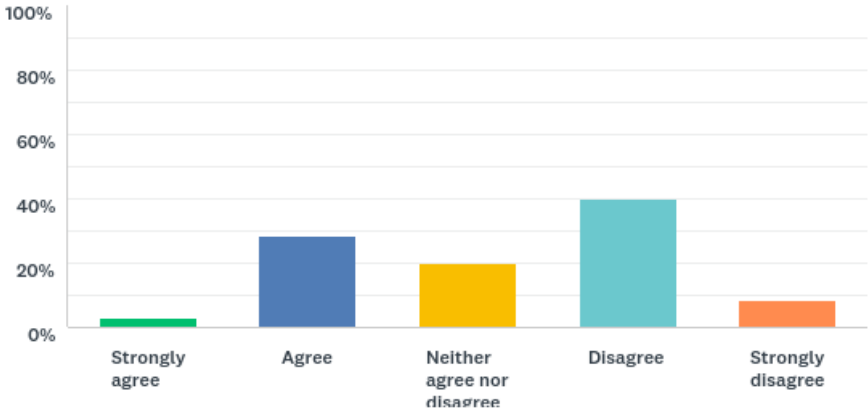


Figure 5.33: The distribution of answers to Statement 12; *I am often unsure about which classification level to apply*

Statement 13 *When I am unsure about which classification level to apply, I....* In cases when they are unsure, the majority of the respondents choose to *ask the project manager*. The option with the second largest amount of responses was *classify at a higher level*. The results are shown in Figure 5.34.

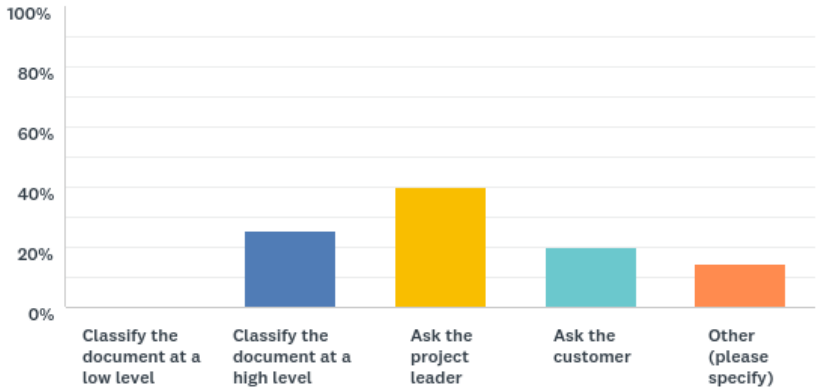


Figure 5.34: The distribution of answers to Statement 13; *When I am unsure about which classification level to apply, I...*

Statement 14 *I believe applying the organization's classification scheme is...* As shown in Figure 5.35, almost all respondents believe that it is a necessity. About 20 % of the respondents find it easy and another 20 % find it difficult to apply the policy.

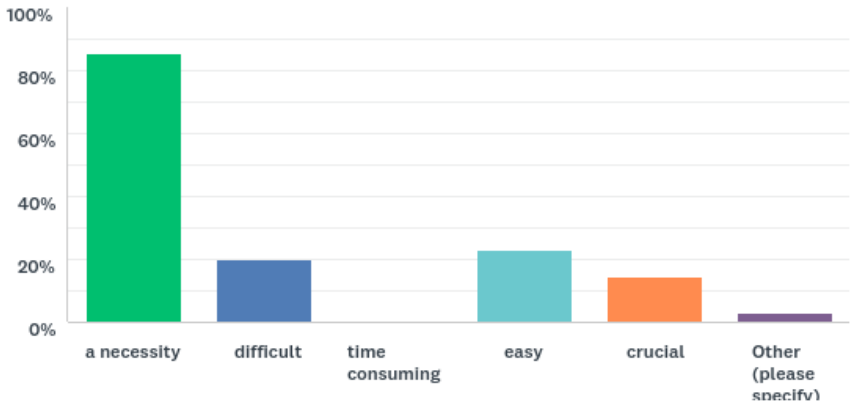


Figure 5.35: The distribution of answers to Statement 14; *I believe applying the organization's classification scheme is...*

.

The following statements are concerned with how the participants believe their routines will be affected by the DLP solution.

Statement 22 *The security tool will enable me to more easily practice the organization's classification policy.* A clear majority of the participants agree with this statement, as illustrated in Figure 5.36.

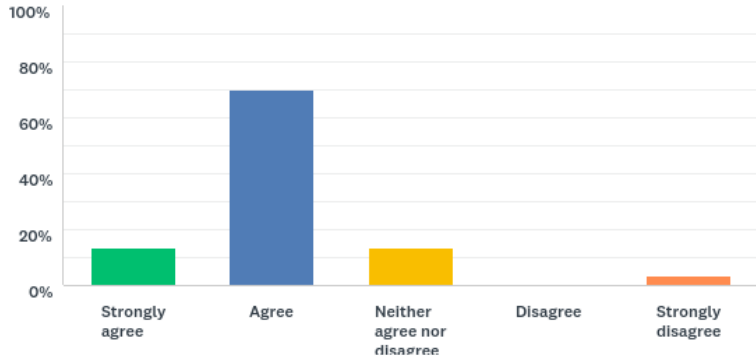


Figure 5.36: The distribution of answers to Statement 22; *The security tool will enable me to more easily practice the organization's classification policy.*

Statement 23 *The security tool will help me better protect customer data.* As presented in Figure 5.37, a clear majority of the respondents agree with the statement.

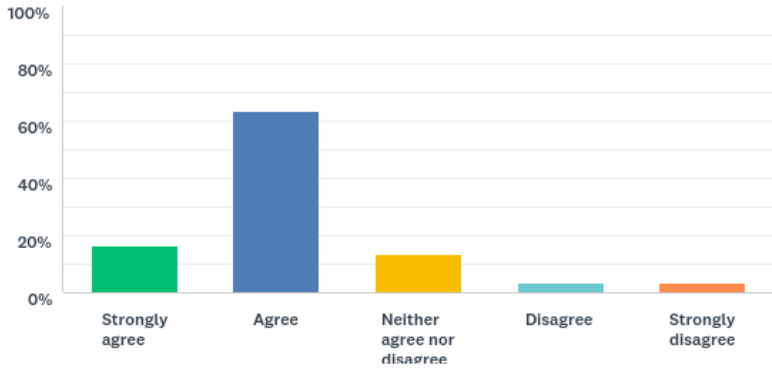


Figure 5.37: The distribution of answers to Statement 23; *The security tool will help me better protect customer data.*

Statement 26 *The security tool will make me more aware of the organization's classification policy.* As shown in Figure 5.38, more than two-thirds of the respondents agree to this statement.

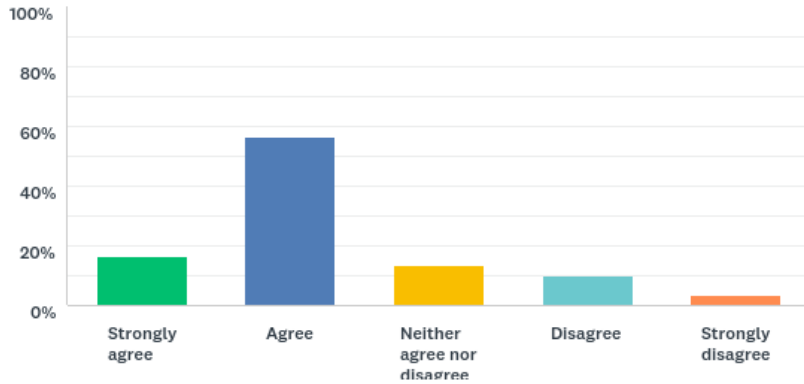


Figure 5.38: The distribution of answers to Statement 26; *The security tool will make me more aware of the organization's classification policy.*

Statement 27 *The security tool will improve my classification routines.* As illustrated in Figure 5.39, a vast majority of the respondents agree with this statement.

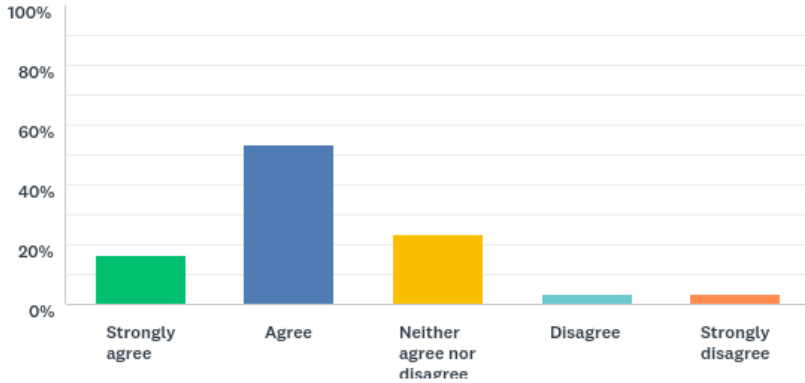


Figure 5.39: The distribution of answers to Statement 27; *The security tool will improve my classification routines.*

Statement 29 *There is a need for this security tool in the organization.* Slightly more than half of the respondents agree with this statement. However, most of the remaining respondents are unsure. The result is presented in Figure 5.40.

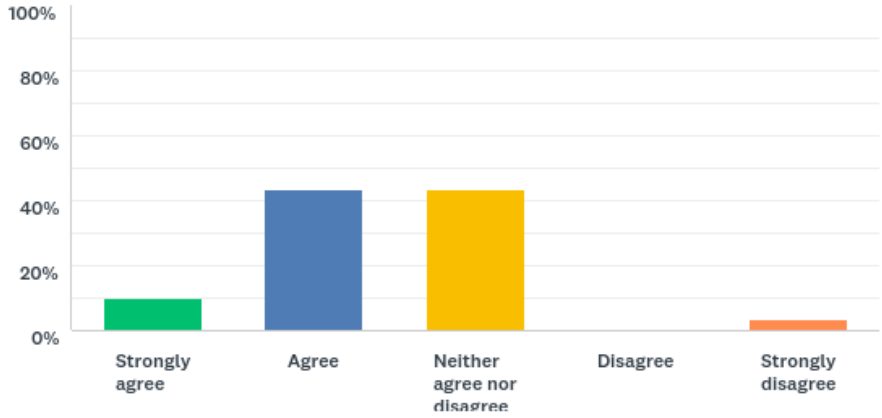


Figure 5.40: The distribution of answers to Statement 29; *There is a need for this security tool in the organization.*

Statement 30 *I intend to use the security tool* As presented in Figure 5.41, 90 % of the respondents find it likely that they will use this tool.

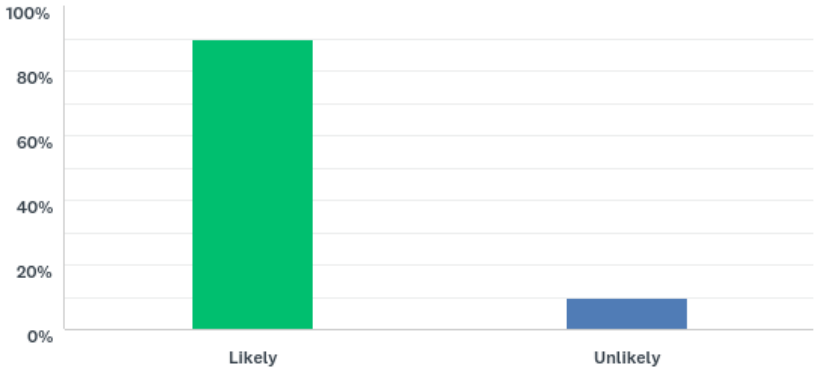


Figure 5.41: The distribution of answers to Statement 30; *I intend to use the security tool.*

The differences in classification routines between managers and non-managers was investigated. The Compare Rule function in SurveyMonkey was used to group the results based on the answers to Question 1, *Do you have a personnel management role?* The rule was applied to statements 2,5,6,7,8,10,11,12,13,14 as these were considered relevant for the investigation of *H3*.

Statement 2 *What kind of documents do you classify today?* The result show that all managers classify documents, while some of the non-managers do not classify any documents. The result is illustrated in in Figure 5.42.

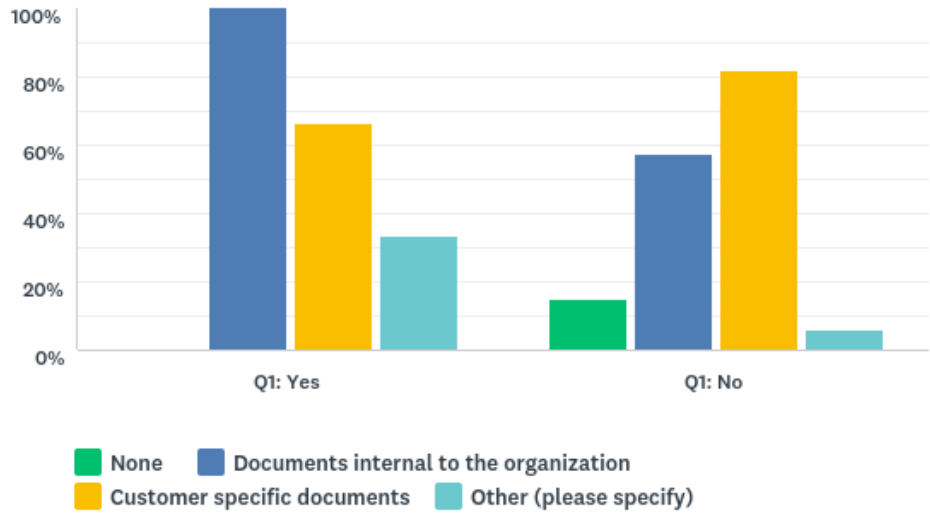


Figure 5.42: The distribution of answers with the Q1 Compare Rule applied to Statement 2; *What kind of documents do you classify today?*

Question 5 *Does or would working with projects exposed to information security risks affect your awareness regarding information security and the organization's security policy?* As shown in Figure 5.43, there is no significant difference between managers and non-managers regarding this question.

Question 6 *How often do you classify documents?* The result shows that, overall, the managers classify documents more often than non-managers. This is visualized in Figure 5.44.

Statement 7 *I believe classification of documents is important.* in Figure 5.45. As visualized in Figure 5.43, there is no significant difference between managers and non-managers regarding this statement.

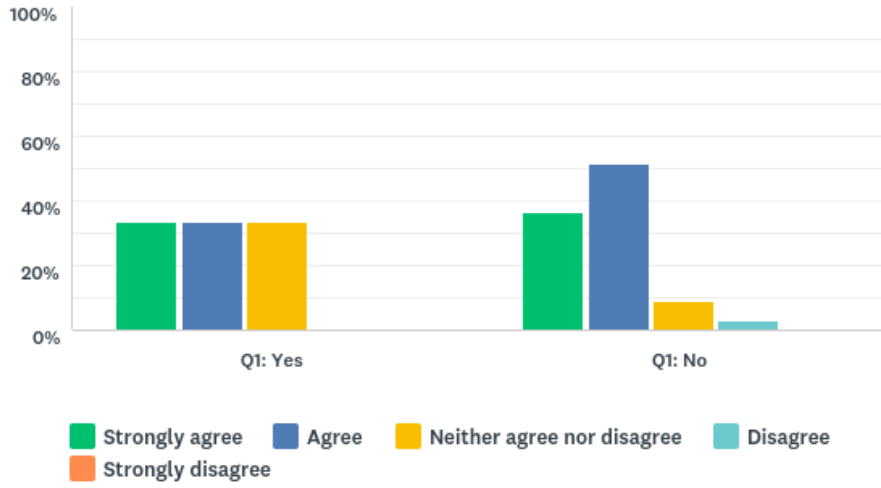


Figure 5.43: The distribution of answers with the Q1 Compare Rule applied to Question 5; *Does or would working with projects exposed to information security risks affect your awareness regarding information security and the organization's security policy?*

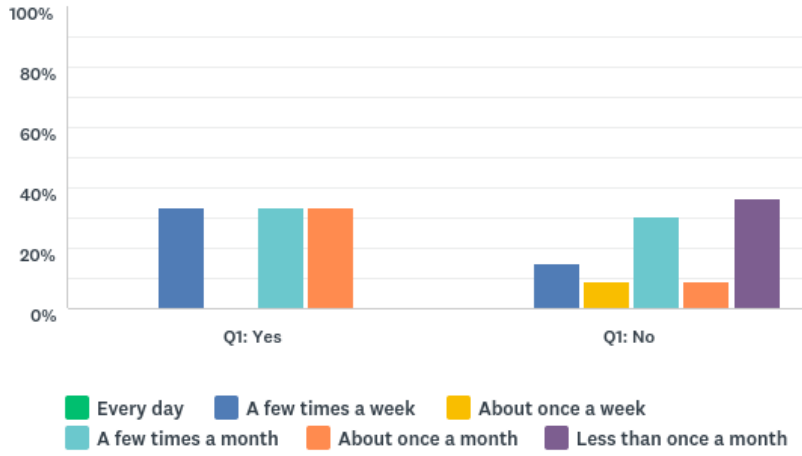


Figure 5.44: The distribution of answers with the Q1 Compare Rule applied to Question 6; *How often do you classify documents?*

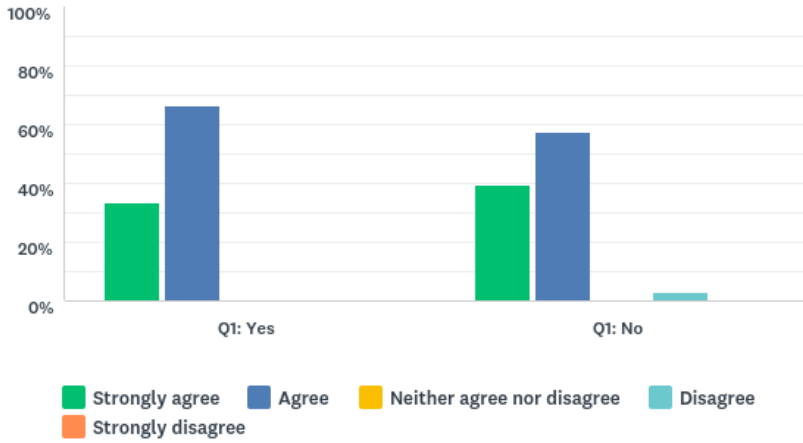


Figure 5.45: The distribution of answers with the Q1 Compare Rule applied to Statement 7; *I believe classification of documents is important*

Statement 8 *Getting work done fast has a higher priority than following the security policy.* As presented in Figure 5.46, the managers disagree slightly more than non-managers with this statement.

Statement 10 *I am familiar with the organization's classification policy.* The result show that the managers agree slightly more with this statement than non-managers. This is illustrated in Figure 5.47.

Statement 11 *I am aware of the consequences of classifying wrong.* As shown in Figure 5.48, all the managers agree with this statement, while among the non-managers about 70 % agree.

Statement 12 *I am often unsure about which classification level to apply.* As visualized in Figure 5.49, all the managers disagree with this statement, while among the non-manager there is a fairly equal spread over the alternatives *Agree*, *Neither agree nor disagree* and *Disagree*.

Statement 13 *When I am unsure about which classification level to apply, I....* The result presented in Figure 5.50 show that there are no significant differences between managers and non-managers.

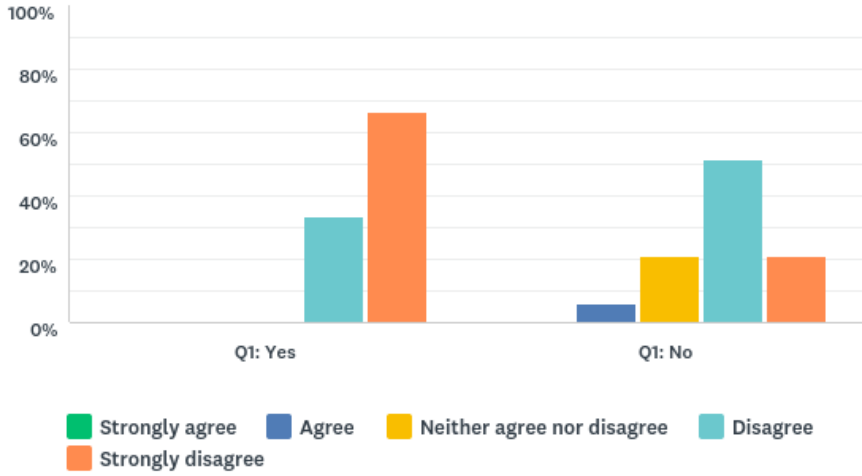


Figure 5.46: The distribution of answers with the Q1 Compare Rule applied to Statement 8; *Getting work done fast has a higher priority than following the security policy*

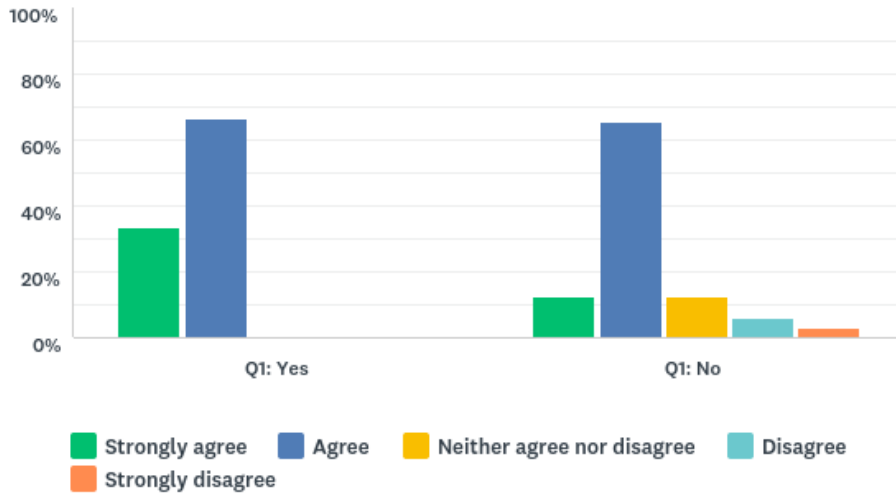


Figure 5.47: The distribution of answers with the Q1 Compare Rule applied to Statement 10; *I am familiar with the organization's classification policy*

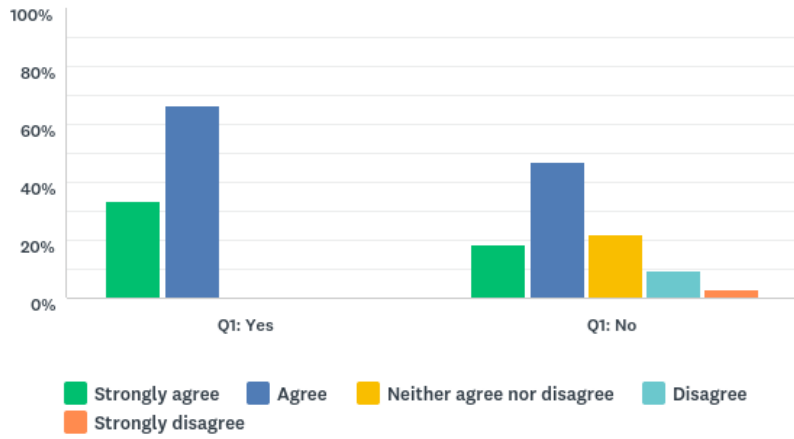


Figure 5.48: The distribution of answers with the Q1 Compare Rule applied to Statement 11; *I am aware of the consequences of classifying wrong.*

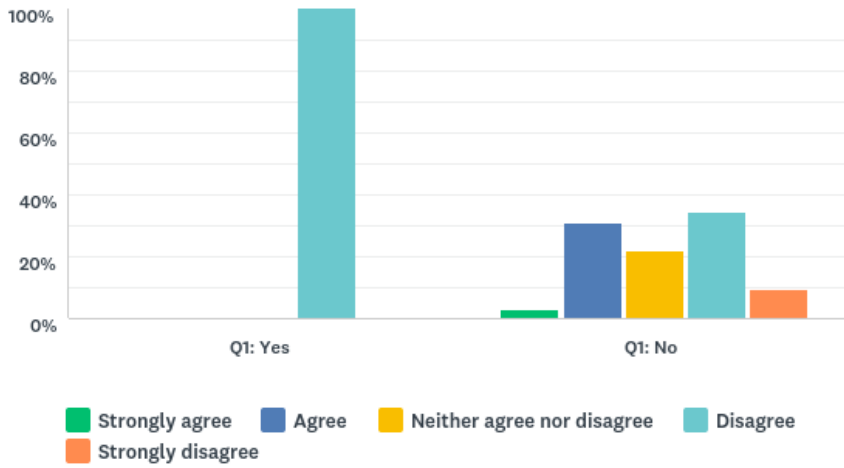


Figure 5.49: The distribution of answers with the Q1 Compare Rule applied to Statement 12; *I am often unsure about which classification level to apply.*

Statement 14 *I believe applying the organization's classification scheme is ....* All of the managers believe it is *a necessity*, and this applies to almost all the non-managers as well. None of the managers responded that they found it difficult

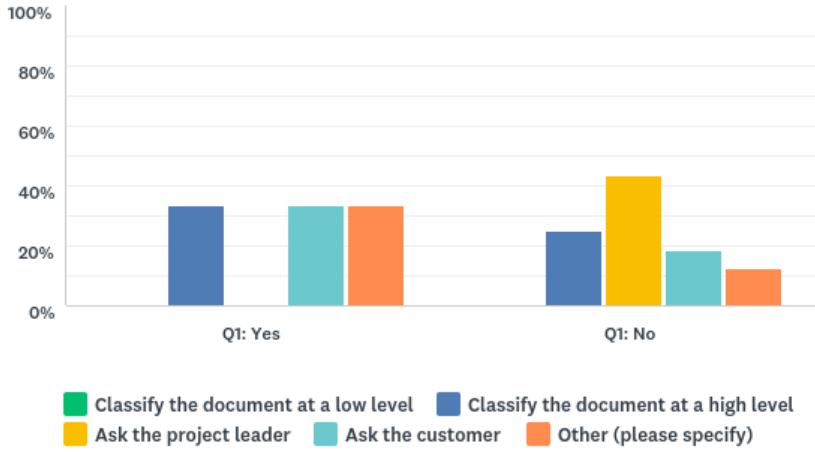


Figure 5.50: The distribution of answers with the Q1 Compare Rule applied to Statement 13; *When I am unsure about which classification level to apply, I...*

or easy, but about 20 % non-managers responded that they found it easy or difficult. In addition, the managers find it slightly more crucial than the non-managers. The result is illustrated in Figure 5.51.

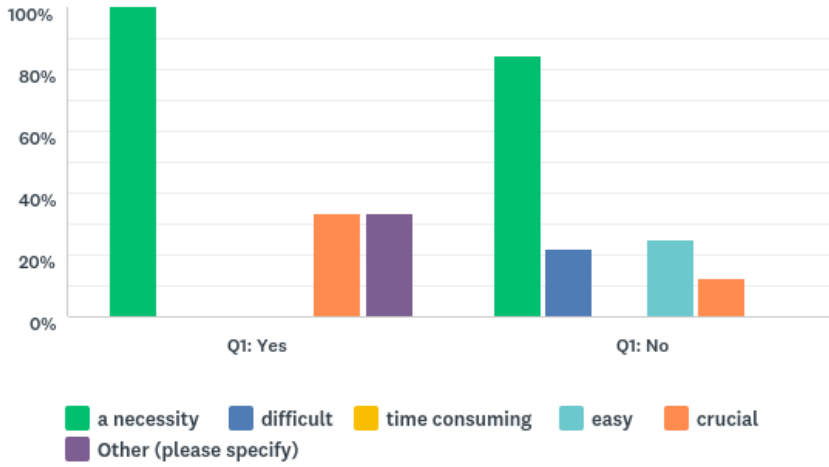


Figure 5.51: The distribution of answers with the Q1 Compare Rule applied to Statement 14; *I believe applying the organization's classification scheme is ...*

5.2.3 Actions the organization can take to ensure a successful implementation

The question below was included to get a better insight into what actions the organization should focus on before implementing the solution.

Question 31 *What do you expect from the organization before the solution is implemented?* In descending order, the three most chosen options were *Information (meeting, email)*, *Enforcement by the organization* and *Online user guide*. The result is shown in Figure 5.52.

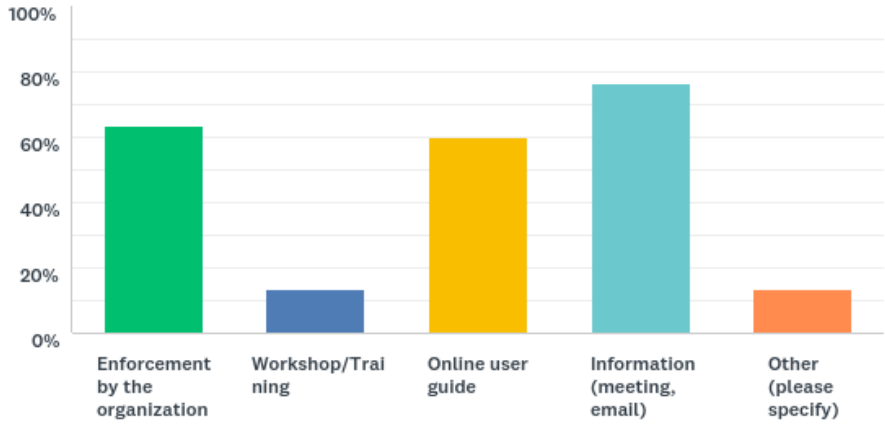


Figure 5.52: The distribution of answers to Question 31; *What do you expect from the organization before the solution is implemented?*

This chapter presented the results from the usability tests, interviews and questionnaire. In general, the interviews, usability tests and answers to the questionnaire revealed that the participants were positive towards the solution. However, regarding the actual implementation, they had some suggestions and concerns. The result also expressed to what extent the solution would affect their daily work. In the light of previous empirical studies in the field and other aspects, the results are further discussed in the next chapter.

Chapter 6
Discussion

In this chapter the results from the interviews, usability tests and questionnaire are discussed and compared with findings from existing literature. The chapter is structured as follows; first, the results are discussed in relation to the research questions and existing knowledge, then the limitations of the study are identified and presented and finally, some directions for future work are suggested.

6.1 RQ1: To what extent can DLP features be introduced before they are perceived as barriers and reduce the user experience?

As there is a lack of previous research on user experience of security tools, the main reference in *RQ1* is the study of the security tool *Polaris* by Dewitt et al. [JDK06] and the research by dePaula et al. [dPDD+05]. The discussion of the results relevant for *RQ1* is divided into three parts, which also correspond with the purpose outlined for each scenario in Table 3.4 (Chapter 3):

- User vs. system control
- Workflow efficiency and potential barriers
- Integration of customers' classification scheme

6.1.1 User vs. system control

A previous study by dePaula et al. [dPDD+05] argues that it can be problematic to remove security control from the user, as only they can decide appropriately when and how to share information. dePaula et al. adds that care should be taken when considering removing power from the user if one believes that the system can do a better job. Moving security features into the background also removes the user's

understanding of the implications of these actions and potential security problems [dPDD+05]. Thus, it is about finding a balance between using automatic tools in decisions and decision-making by employees. The results from Scenario 1 and 3 show that the employees prefer that documents and templates are pre-classified with a default label that they can change in later time, rather than having the solution recommend a classification level dynamically based on the scan of the content. It was expressed concern whether the solution's recommendations could be trusted and that it would be difficult for the solution to recommend properly for all types of documents. It was also emphasized that it should be easy for the user to change the pre-set level when required. However, the choice of default classification levels for templates should be well thought through, especially since document types such as notes and reports can have all different levels of classification and finding a suitable default value can be difficult. Thus, it can be concluded that the employees are positive to some level of system control regarding classification of documents, but that they want to be able to make the final decision themselves. These findings are supported by the arguments cited from dePaula et al. [dPDD+05].

6.1.2 Workflow efficiency and potential barriers

Regarding the effect the security solution would have on their workflow, a clear majority of the respondents believe that using the tool will require low effort and not decrease their job productivity. In addition, more than half of them agree that the tool will be useful in their job, while the rest were unsure. The majority of the respondents believe that both emails and documents should be classified. In general, the results show that following the security policy has a higher priority than getting work done fast. These findings indicate that the general attitude towards the usability aspects of the implementation is positive.

Dewitt et al. [JDK06] state that there is a strong possibility that alerts in security tools will annoy users and that they as a consequence will be ignored and just clicked away without being read. However, Dewitt et al. also says that the informative effect of alerts should not be underrated. Again, it is about finding a balance. Zurko et al. [ZKSB02] found that either the security culture or security-related user interfaces must be changed if warning users by making them click boxes to proceed with their work should have an effect. They also state that:

> *The more frequently security warnings appear in everyday use, the more users will learn to click "OK" without thinking or even remembering that they have done so.*

It is also emphasized that one should treat false alarms as serious security vulnerabilities and not regard them as acceptable irritants. The responses revealed that most

of the participants expressed concern that having a pop-up requiring classification when opening a new document would be an annoyance and perceived as bureaucratic. The results from both the usability tests and the questionnaire show that most of the participants support including the justification function. However, there was some concern that the explanation text might be written without care if the box occurred too frequently. The results also show that a justification was only considered relevant when changing the label from a confidential classification level to a lower level. The most important advantage of including the function was considered to be traceability. Regarding AIP in Outlook, most of the participants were comfortable with pop-ups when trying to send emails labeled with the most confidential classification level. This also applied to cases where attachments have a higher classification label than the email. Regarding text in subject fields and attachment names, almost all participants did not want a system check and pop-up warning. These findings show that there is a clear concern that frequent pop-ups may have a negative effect on workflow efficiency. However, in cases where the value of the function is recognized, it is acceptable. This agrees with the conclusion in the study by Dewitt et al. [JDK06].

An important finding in the research on the usability of the security tool Polaris by Dewitt et al. [JDK06] was the additional difficulties introduced when security features are added to an existing system rather than being integrated from the start. Dewitt et al. found it unlikely that the security tool *Polaris* would be successful as it was a post-hoc consideration [JDK06]. In the case of AIP, which is also an add-on, this could be a concern. However, as most employees are familiar with Office programs, one might conclude that the implementation would not be perceived as comprehensive compared to introducing a completely new program. Another point made by Dewitt et al. [JDK06] was:

> *Polaris should be more tightly integrated with the operating system so that context sensitive menus can be used.*

One of the features in AIP is that context sensitive menus is an integrated part of the solution, for example a part of the file explorer. As such, this can be regarded beneficial for a successful integration of the solution. A study on website interface design in general, Guay et al. [GRR19] found that it was problematic to have a cluttered interface and that important content likely would be obscured [GRR19]. This is supported by Fidas et al. [FVA10] who state that a bad user interface experience due to security features will have negative effects on the security [FVA10]. The additional toolbar that AIP adds to Office programs was in general considered acceptable, both with regards to space and location. In addition, the ease of use of the security tool in general appeared clear and understandable. With the above findings in mind, it is clear that the fact that the AIP solution uses both context

sensitive menu and has a tight integration with an already well-known user interface, is a clear advantage.

Dewitt et al. [JDK06] also suggest that identified barriers in programs should be worked through before they are imposed on the users, for instance as a corporate security policy. The results revealed several potential barriers in the AIP solution, and these are listed in Section 5.1.2. Of these, the five most chosen were *Lack of information from the organization*, *Lack of enforcement from the organization*, *Annoying pop-ups*, *Unnecessary features* and *Lack of experience*. One of the barriers mentioned in the interviews was *Little or no consequence of performing actions causing indifference to the use value*. It was pointed out, regarding the toolbar functionality, that applying classification did not have the expected consequence of a visible change with the document, for example on the document front page. As a consequence, the classification labeling would involve two actions, both using the toolbar and editing the document, and thus reducing the usefulness. This kind of mismatch between expected and real functionality is supported by findings in Dewitt et al. [JDK06]:

> *One participant assumed that Polaris was automatically protecting their files at all times, when in fact some of the applications they were using were not under the protection of Polaris. This user had a high expectation of the security software in that they didn't expect to have to take any explicit action in order to be protected.*

In conclusion, as barriers would affect the efficiency of the employees' workflow and thus decrease the effectiveness of the protection offered by AIP, actions must be taken to reduce these before implementing.

6.1.3 Integration of customers' classification scheme

Nearly half of the respondents often work with customer projects where they meet additional classification requirements that differ from the organization's classification policy. This indicates that integration with external classification schemes is a relevant issue. Most of the participants were positive about integrating the customer's classification scheme, mainly because it makes it easier to relate it to the organization's scheme. In addition, having a customer specific button or label by implementing a *scoped* policy, as described in Chapter 4, would work as a reminder. However, there was a concern that it would add undesired complexity, both to the implementation and to the user interface.

As mentioned in Section 5.1.1, one participant, who works with customer projects only, expressed a concern that the classification level had to be changed to a lower

level every time a document or email was to be sent outside of the organization. One option to cope with this situation is to include a label named "Customer External", which is customized to remove access rights. In this way, the customer needs no specific access privileges, which makes it easier to deal with the document or email. The classification is still visible, but serves no purpose in terms of access restrictions, only as a visual indicator of what classification label that is applied. However, as this adds more responsibility to the employee in terms of security handling, there might be a need for additional policy instructions.

According to Microsoft, there is no limit to the number of labels that can be added to the toolbar. However, a large number of buttons may have a negative impact on the user experience. As mentioned earlier and supported by Guay et al. [GRR19], it is important to avoid that the interface becomes too cluttered.

In conclusion, there is an acceptance that integrating customer classification schemes will affect the user experience to some extent. There are potential issues, such as how the solution will handle a situation where there is a significant number of customer specific labels. Consequently, there is a need for further research into this matter.

6.2 RQ2: How does a DLP solution affect the employees' classification routines?

As a basis for the discussion of *RQ2*, the current situation regarding classification routines in the organization was identified. The result shows that the employees classify documents monthly on average. In addition, the majority believe classification of documents is important and regard it as a necessity. They also express that they are familiar with the organization's classification policy. In general, these results indicate there is an existing security culture within the organization and a general understanding of the importance of data security. This finding is supported by Veiga et al. [Vei16] who concluded that reading the policy had a positive effect on security culture within the organization.

The respondents were also asked about their expectations of the usefulness of the solution and how they believed their routines would be affected by it. A clear majority believe that the solution will both enable them to more easily practice the organization's classification policy and improve their classification routines. In addition, almost all the respondents expressed that it will help them better protect customer data, and more than half of them feel there is a need for the security tool in the organization. The findings agree with the fact that more 90% of the respondents answered that they find it likely that they will use the tool.

All of the findings above show that the employees are in general positive towards document classification and also the use of supportive tools. This is in line with the conclusion in the study by Safa et al. [SSS+15] that the level of experience with information security and data protection in an organization has a positive effect on security consciousness. However, as found by Safa et al., a possible explanation for the positive results may also be that most of the employees who chose to respond to the questionnaire, had an interest in information technology and information security. This kind of bias could have been revealed if the specific job role and experience with information security had been included in the questionnaire. Safa et al. also found that users' perception of control often does not correspond with how they actually behave. With this in mind, it may be that the results in this project are affected by some respondents' overestimated self-efficacy.

6.2.1 H1: Employees tend to classify documents as confidential by default, which sometimes might result in information being stricter classified than required

Most of the respondents state that they do not prioritize getting work done fast above compromising security. In cases where they are unsure about what classification level to apply, most of them say they ask someone, either the project manager or the customer. The questionnaire results show that only 25% choose to classify the document at a high level. Thus, it is not a crucial issue to predict how an introduction of the DLP solution will affect the number of false positives. However, as the result shows, most of the respondents believe that the DLP solution will improve their classification routines. Furthermore, this may indicate that an introduction of the solution will reduce the number of false positives. Overall, these findings show that the employees believe classification is important. This is in contrast with the result found in the study by Dewitt et al. [JDK06] where the participants preferred speed to security. Possible explanations for the difference could be type of organization, previous experience, limited number of participants and a non-representative sample. Based on the results of this project, the hypothesis *H1* was proven to be false.

6.2.2 H2: A DLP solution will make employees more aware of an organization's classification policies

A study by Weirich et al. [WS01] shows that users will not make good security decisions unless they believe they are at risk. The results from the questionnaire shows that two-thirds of the respondents work with projects that may be exposed to information security risks. Almost all respondents believe that working with such projects affect their awareness regarding information security and the organization's security policy. Furthermore, two-thirds state they are aware of the consequences of classifying wrong. However, neither the usability test nor the questionnaire included

questions aiming to test the users' ability to classify correctly. As mentioned earlier, Safa et al. found that users' perception of control often does not correspond with how they actually behave. By making participants and respondents classify a set of organization specific documents of different types and classification levels, one could potentially reveal the awareness of the classification policies in more detail. Overall, the result shows that the general awareness among the employees regarding security risks is high, but introducing a DLP solution may still have a significant effect on the awareness, as discussed in the next paragraph.

According to the study by Dewitt et al, [JDK06], a visual indication in software solutions that data is under risk is more effective that education in order to increase awareness of information security and motivation for protecting it. This is supported by findings in this project, where more than two-thirds of the respondents said that the security tool will make them more aware of the organization's classification policy. The most evident visual indicator in AIP is the toolbar and especially the color indication associated with each classification level label, which contributes to how the employees perceive the risk of the document or email. It was pointed out that although people are aware of the different classification levels, they might struggle describing them. The toolbar provides additional information by hovering the labels, so in that sense the solution will work as a reminder. Still, the visible indicators do not rule out the need for education. Reading the policy has a positive effect as shown by Veiga et al. [Vei16]. Based on the results of this project, the hypothesis *H2* was proven to be true.

6.2.3 H3: Employees with management roles are more aware of classification than other employees

In general, the results show that there is no or little difference between managers and non-managers both regarding how often they work with projects exposed to information security risks, their security awareness and their perception of the necessity of having a security policy. However, it was revealed that managers classify documents slightly more often and are moderately more aware of and familiar with the organization's classification policy. Based on this, it appears that the differences are small and that the managers do not feel that there is a big gap between the organization policy and practice of it among the employees. To some degree, these findings are supported by Strand [Str18] who in a research study conducted with the same organization as in this master thesis, found that the employees with a higher level of responsibility were more concerned with information security. Thus, the difference found in this project is not as clear as observed in previous research. However, Dewitt et al. [JDK06] found that users base their decisions on previous experience. Thus, in addition to their level of responsibility, the employees' classification habits can also be affected by their former workplaces and projects they have been involved in.

Unfortunately, the number of respondents with a management role in this project was only three. Thus, only a very limited data basis was available to assess this hypothesis.

Several earlier studies have indicated that the organization's security culture is affected by social factors and managers influence. As mentioned in Chapter 2, Knapp et al. [KMRF06] found that top management support had a positive impact on the security culture and enforcing the security policy in an organization. However, results from a study by Hagen et al. [HAH08] shows that only 60 % of the respondents felt that the top managers in their companies actually were engaged in the information security work. A reason for this could be that top management lack the technical skills and knowledge of information security [HAH08]. The interviews in this study revealed that the relationship between managers and employees with regards to document classification, was built on trust and that there were no audit reviews of the classification of documents. The managers trust their employees and do not impose requirements on them today. The employees do not feel that their managers care more about classification than they do. Some of the managers expressed that they were more concerned with protecting data that includes personal information about their employees. When comparing the findings with Knapp et al. [KMRF06] it should be kept in mind that they refer to top managers, whereas this project is concerned with employees with a personnel management role.

As referred to in Section 1.2, Stanton et al. [SMSJ04] found that employees' security behavior is affected by their manager's security behavior. It therefore made sense to assess whether initiatives targeting the classification routines of managers could be the most efficient. Since it was not found any significant difference in the behavior between managers and non-managers, there appears to be no obvious need for initiatives targeting classification routines of managers in particular. In conclusion, the hypothesis *H3* was proven to be mostly false.

6.2.4 Actions the organization can take to ensure a successful implementation

The study by Dewitt et al. [JDK06] revealed that if users are not using the security tool correctly or not knowing its features, the level of security would not increase. In fact, the users thought they were protected by the solution in cases where they were not [JDK06]. This shows that without proper education there will still be security risks. In light of this, the respondents in this project were asked what they expected from the organization before implementing the DLP solution. The most chosen answers were information through meetings and emails, enforcement by the organization or providing an online user guide. The study by Hagen et al. [HAH08] points out that worker participation in information security will improve the

individual awareness, ownership, motivation and acceptance of information security [HAH08]. Taking the feedback from the respondents into account, may contribute to a successful introduction of the solution.

According to Weirich et al. [WS01], in a socio-technical system, the effectiveness strongly relies on the user's willingness to perform that extra effort that a secure behavior requires. A clear majority of the respondents in this master thesis project state that following the security policy has a higher priority than getting work done fast. This indicates that there is a good basis for the implementation of the solution.

Weirich et al. [WS01] also states that security policies alone are not sufficient to ensure correct behavior. Furthermore, they conclude that in most organizations, force cannot be used to make employees comply with policies. Instead, they should be persuaded through tutorials, training and discourse [WS01]. This agrees with the feedback from the respondents in this project and, as such, appears to be a good approach.

To sum up, the following factors are important to consider for an organization in order to achieve a successful implementation:

- Employee involvement
- Identification and removal of barriers
- Enforcement
- Training and campaigns
- Online user guide

6.3 Limitations

The methods chosen and how the research was conducted may have affected the results. This section discusses potential limitations in the study.

The sample size in the quantitative parts was not optimal. Compared to the number of employees in the case organization, which is 2000, the number of respondents in the study is low. In total, there were 36 respondents to the questionnaire. If the results need only to be reasonable evident, it is acceptable to choose an Exploratory

Confidence level of 80 % [1]. According to a SurveyMonkey online calculator [2], given a population size of 2000, the sample size should have been 152, to achieve an 80 % confidence level and margin of error of 5 %. Using 36 respondents and a confidence level of 80 %, gives a margin of error of 10.6 %. The margin of error is high, and it is thus clear that the small sample size has limited potential for generalizing. One of the reasons for the low number of respondents could be that answering the questionnaire was voluntary or that only employees with an interest in the research field chose to respond. Making the questionnaire mandatory or using incentives could potentially have increased the number of participants. However, there is a risk that this could also have reduced the quality of the results as some participants might not put much effort into responding properly in order to complete the questionnaire fast. Additionally, as the questionnaire included predetermined alternatives, it might be that some of the answers are biased.

The distribution of respondents in the managers and non-managers groups was not optimal. In this master thesis project, there were only 3 managers responding to the questionnaire. The case organization consists of 2000 employees where 240 of these have a personnel management role. It is therefore clear that the number of respondents with personnel management role does not provide a sufficient data base.

In total 7 employees participated in the usability test and interviews. Four of these had a management role. The limited number of participants was due to time restrictions. With a higher number of participants would also have made it possible to get an equal ratio between managers and non-managers. However, previous research has shown that including a minimum of 5 participants in a usability test provides accurate results.

The testing time for each participant was limited and may have influenced their perception of the DLP solution features presented. As a result, they might have answered differently if knowing all the features in detail. For instance, one of the participants changed his mind about the usefulness of the solution after using it for one week after the usability test. As suggested by Nielsen [Nie00], using multiple iterations may improve the quality of the test. Again, due to time limitations, this was not possible.

The facilitator had only limited experience with carrying out interviews and usability tests. Being accompanied by an experienced interviewer could have been an advantage. However, both the interviews and the usability test were based on a prepared guide, reducing the risk of ignoring important topics.

[1] "How confident do you need to be in your research?", MeasuringU, accessed April 25, 2019, https://measuringu.com/confidence-levels/

[2] "Sample size calculator", SurveyMonkey, accessed April 25, 2019, https://www.surveymonkey.com/mp/sample-size-calculator/

Also, since the usability test was based on a demonstration of the solution and the questionnaire only included screenshots of it, the effect the solution has on the employees' classification routines was not explicitly tested before and after an actual implementation. This means that one cannot conclude the exact effect an introduction of the solution would have on the classification routines, only the participants and respondents predicted effect.

Another limitation for the validity is the type of organization. As the case organization is a knowledge organization, their projects are often intended to be shared with the society. As such, the security culture might be different from for example a financial institution or a consultant company. It is therefore important to emphasize that the results found in this study regarding the implementation of features in the DLP solution are most relevant for similar types of organizations. In addition, it may be that chosen functions and problems and desired features identified by the participants in AIP is not applicable to other DLP solutions in the market.

6.4 Future work

The main focus in this project has been to investigate the balance between user experience and security when introducing a security software tool in an organizational context. The results were based on a demonstration of the security tool functions. For future research it would be interesting to measure the explicit effect on the classification routines of an actual implementation. In addition, different types of organizations and DLP solutions could be included to get more generalized results.

Furthermore, it could have been interesting to find out whether a security campaign related to DLP and AIP would have had an effect on the success of the implementation. This would require setting up a control group within the organization that does not participate in the campaign.

As pointed out in this research, supporting integration of multiple classification schemes from different parties is an important feature of any DLP solution. Given the additional complexity introduced, such as how to synchronize with customer schemes and updates efficiently without reducing the user experience, there is an obvious need for further investigations.

Chapter 7
Conclusion

A DLP solution is a viable tool for organizations in their effort to meet challenges related to security and protection of critical data today. However, introducing a new solution in an organization is always a risk as it in most cases to some extent will affect the employees daily work life and may be perceived as a barrier. This project has explored a balance between security and user experience when introducing a DLP solution in a knowledge organization.

RQ1 The first research question explored to what extent DLP features can be introduced before they are perceived as barriers and reduce the user experience. The results show that in general the employees were positive towards implementing the solution. However, it was found that it is essential that the employees perceive the features as useful and retain control over the classification functions and understand how they work. In addition, it is an advantage that the solution integrates well with existing software and an already well-known user interface. Furthermore, it should not introduce barriers, such as frequent pop-ups, that interrupt their workflow. However, in cases where the value of the function is clearly recognized, it is acceptable. To avoid barriers that may decrease the effectiveness of the protection offered, actions should be taken to reduce these before implementing the solution. As many organizations work with customer related projects, it is important for producers of DLP solutions to support integration of multiple classification schemes.

RQ2 The second research question investigated how a DLP solution could affect the employees' classification routines. It was found that employees do not tend to classify documents as confidential by default. Instead, when in doubt, they consulted either the project leader or the customer. Even though the results did not make it possible to predict how an introduction of the DLP solution would affect the number of false positives, it was revealed that most of the respondents believe it will improve their classification routines. The results also showed that employees with management roles are only slightly more aware of classification than other employees. Consequently, there is no obvious need for initiatives targeting classification routines

of managers in particular. Despite an existing security culture, it was revealed that the solution will both enable employees to more easily practice the organization's classification policy, improve classification routines and help better protect customer data. In addition, the DLP solution will make them more aware of the organization's classification policies. However, a successful implementation of a DLP solution demands for actions by the organization, such as providing information, enforcement and tutorials.

References

[ASM16] Sultan Alneyadi, Elankayer Sithirasenan, and Vallipuram Muthukkumarasamy. A survey on data leakage prevention systems. *Journal of Network and Computer Applications*, 62, 2016.

[Beg00] Thomas A. Beggs. Influences and barriers to the adoption of instructional technology. *ERIC - Institute of Education Sciences*, pages 1–14, 2000.

[BG04] Petra M. Boynton and Trisha Greenhalgh. Selecting, designing, and developing your questionnaire. *BMJ*, 328(7451):1312–1315, 2004.

[BH13] Ajay Bandi and Phil Heeler. Usability testing: A software engineering perspective. *2013 International Conference on Human Computer Interactions (ICHCI)*, pages 1–8, Aug 2013.

[BJH06] Andrew Burton-Jones and Geoffrey S. Hubona. The mediation of external variables in the technology acceptance model. *Information & Management*, 43(6):706 – 717, 2006.

[Boy98] Richard E. Boyatzis. *Transforming Qualitative Information: Thematic Analysis and Code Development*. SAGE, 1998.

[Cha11] Stephane Charbonneau. The role of user-driven security in data loss prevention. *Computer Fraud & Security*, 2011(11):5 – 8, 2011.

[CWK05] Mark Chan, Irene Woon, and Atreyi Kankanhalli. Perceptions of information security in the workplace: Linking information security climate to compliant behavior. *Journal of Information Privacy and Security*, 1:18–41, 07 2005.

[Del05] Amy Dellinger. Validity and the review of literature. *Research in the Schools*, 12:41–54, 01 2005.

[DGDdlFJ04] Paul Dourish, Rebecca E. Grinter, Jessica Delgado de la Flor, and Melissa Joseph. Security in the wild: User strategies for managing security as an everyday, practical problem. *Personal Ubiquitous Comput.*, 8(6):391–401, November 2004.

[DIAD10] Norizan M. Diah, Marina Ismail, Suzana Ahmad, and Mohd K. M. Dahari. Usability testing for educational computer game using observation method. *2010 International Conference on Information Retrieval Knowledge Management (CAMP)*, pages 157–161, March 2010.

[DN13] Owen Doody and Maria Noonan. Preparing and conducting interviews to collect data. *Nurse Researcher*, 20(5):28–32, 2013.

[DOSC16] Gurpreet Dhillon, Tiago Oliveira, Santa Susarapu, and Mario Caldeira. Deciding between information security and usability: Developing value based objectives. *Computers in Human Behavior*, 61:656–666, 08 2016.

[dPDD+05] Rogério de Paula, Xianghua Ding, Paul Dourish, Kari Nies, Ben Pillet, David Redmiles, Jie Ren, Jennifer Rode, and Roberto Silva Filho. Two experiences designing for effective security. *Proceedings of the 2005 Symposium on Usable Privacy and Security*, pages 25–34, 2005.

[FD86] Jr. Fred D.Davis. A technology acceptance model for empirically testing new end-user information systems: Theory and results. *MASSACHUSETTS INSTITUTE OF TECHNOLOGY*, page 291, 02 1986.

[FGD02] S.M. Furnell, M. Gennatou, and P.S. Dowland. A prototype tool for information security awareness and training. *Logistics Information Management*, 15(5/6):352–357, 2002.

[FVA10] Christos Fidas, Artemios G. Voyiatzis, and Nikolaos M. Avouris. When security meets usability: A user-centric approach on a crossroads priority problem. *2010 14th Panhellenic Conference on Informatics*, pages 112–117, Sep. 2010.

[GRR19] Sarah Guay, Lola Rudin, and Sue Reynolds. Testing, testing: a usability case study at university of toronto scarborough library. *Library Management*, 40(1/2):88–97, 2019.

[Gru06] Jan Grund. Kunnskapsorganisasjoner - hva er ledelses- og styringsutfordringene? *MAGMA*, 2006. Accessed: 2019-05-09.

[HAH08] Janne Merete Hagen, Eirik Albrechtsen, and Jan Hovden. Implementation and effectiveness of organizational information security measures. *Information Management & Computer Security*, 16(4):377–397, 2008.

[Har07] Chase Harrison. Tip sheet on question wording. *Harvard University Program of Survey Research*, page 4, 11 2007.

[Hin98] D. Hinderer. Challenges in participant recruiting for usability testing. *IPCC 98. Contemporary Renaissance: Changing the Way we Communicate. Proceedings 1998 IEEE International Professional Communication Conference (Cat. No.98CH36332)*, 2:417–426 vol.2, Sep. 1998.

[ISO18] ISO. Ergonomics of human-system interaction — part 11: Usability: Definitions and concepts. *ISO 9241-11*, (2), 03 2018.

[JDK06] Alexander J. DeWitt and Jasna Kuljis. Aligning usability and security: A usability study of polaris. *ACM International Conference Proceeding Series*, 149:1–7, 01 2006.

[KFR10] Ronald Kainda, Ivan Fléchais, and A. W. Roscoe. Security and usability: Analysis and evaluation. *2010 International Conference on Availability, Reliability and Security*, pages 275–282, Feb 2010.

[KMRF06] Kenneth J. Knapp, Thomas E. Marshall, R. Kelly Rainer, and F. Nelson Ford. Information security: management's effect on culture and policy. *Information Management & Computer Security*, 14(1):24–36, 2006.

[Kun03] Mike Kuniavsky. *Observing the User Experience: A Practitioner's Guide to User Research*. Interactive Technologies. Elsevier Science, 2003.

[LBLB17] Luis Miguel Lopez-Bonilla and Jesus Manuel Lopez-Bonilla. Explaining the discrepancy in the mediating role of attitude in the tam. *British Journal of Educational Technology*, 48(4):940–949, 2017.

[Lec15] Aurélie Leclercq. Managing byod: how do organizations incorporate user-driven it innovations? *Information Technology & People*, Vol. 28 Issue: 1, pp.2-33, 2015.

[LHM01] Carl E. Landwehr, Constance L. Heitmeyer, and John D. McLean. A security model for military message systems: retrospective. *Seventeenth Annual Computer Security Applications Conference*, pages 174–190, Dec 2001.

[LK10] Lyong S.L. Liu and Richard Kuhn. Data loss prevention. *IT Professional*, 12(2):10–13, March 2010.

[LO09] Nancy L. Leech and Anthony J. Onwuegbuzie. A typology of mixed methods research designs. *Quality & Quantity*, 43(2):265–275, Mar 2009.

[Mac91] Wendy E. Mackay. Triggers and barriers to customizing software. *ACM*, pages 153–160, 1991.

[Mal12] Kirsti Malterud. Systematic text condensation: A strategy for qualitative analysis. *Scandinavian Journal of Public Health*, 40(8):795–805, 2012. PMID: 23221918.

[Marnd] Luther Martin. Understanding DLP. *Auerbach Publications*, (n.d). Accessed: 2018-10-25.

[Nie00] Jakob Nielsen. Why you only need to test with 5 users. *Alertbox [Online]*, 03 2000.

[NKX09] Boon-Yuen Ng, Atreyi Kankanhalli, and Yunjie (Calvin) Xu. Studying users' computer security behavior: A health belief perspective. *Decision Support Systems*, 46(4):815 – 825, 2009. IT Decisions in Organizations.

[NM18] Van Marwan Omar Nguyen and Derek Mohammed. A security framework for enhancing user experience. *International Journal of Hyperconnectivity and the Internet of Things (IJHIoT)*, 1(1), 19-28, 3 Oct. 2018.

[NS05] Johan Van Niekerk and Rossouw Von Solms. A holistic framework for the fostering of an information security sub-culture in organizations. *Proceedings of the ISSA 2005 New Knowledge Today Conference*, pages 1–13, 01 2005.

[OO06] Ezra Ondari-Okemwa. Knowledge management in a research organisation: International livestock research institute (ilri). *Libri*, 56:63–72, 01 2006.

[OS10] Chitu Okoli and Kira Schabram. A guide to conducting a systematic literature review of information systems research. *Sprouts: Working Papers on Information Systems*, 10(26):1–51, 2010.

[PS17] C. Mercy Praba and Dr.G. Satyavathy. A technical review on data leakage detection and prevention approaches. *Journal of Network Communications and Emerging Technologies (JNCET)*, 7:6, 09 2017.

[Rie03] Andreas M. Riege. Validity and reliability tests in case study research: a literature review with "hands-on" applications for each research phase. *Qualitative Market Research: An International Journal*, 6(2):75–86, 2003.

[RK17] Brian Reed and Deborah Kish. Magic quadrant for enterprise data loss prevention. *Gartner*, 2017.

[RM10] Mogull RM. Understanding and selecting a data loss prevention solution. *DLP Whitepaper*, 10 2010.

[Rob11] Colin Robson. *Real World Research, 3rd Edition.* John Wiley & Sons Ltd, 09 2011.

[Roh14] Christian Rohrer. When to use which user-experience research methods. *Nielsen Norman Group - World Leaders in Research-Based User Experience*, pages 1–7, 10 2014.

[San00] Margarete Sandelowski. Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in nursing & health (Online)*, 23:246–255, 2000.

[SMSJ04] Jeffrey Stanton, Paul Mastrangelo, Kathryn Stam, and Jeffrey Jolton. Behavioral information security: Two end user survey studies of motivation and security practices. *Proceedings of the 10th Americas Conference on Information Systems*, page 175, 01 2004.

[SS15] Rizwana Shaikha and Dr. M. Sasikumar. Data classification for achieving security in cloud computing. *Procedia Computer Science*, 45:493 – 498, 2015. International Conference on Advanced Computing Technologies and Applications (ICACTA).

[SSS+15] Nader Sohrabi Safa, Mehdi Sookhak, Rossouw Von Solms, Steven Furnell, Norjihan Abdul Ghani, and Tutut Herawan. Information security conscious care behaviour formation in organizations. *Computers & Security*, 53:65 – 78, 2015.

[Str18] Kristine Larsen Strand. Influencing factors and effectiveness of a security awareness campaign. *Norwegian University of Science and Technology Department of Telematics*, June 2018.

[SWS15] Jordan Shropshire, Merrill Warkentin, and Shwadhin Sharma. Personality, attitudes, and intentions: Predicting initial adoption of information security behavior. *Computers & Security*, 49:177 – 191, 2015.

[Tjo18] Aksel Tjora. *Qualitative Research as Stepwise-Deductive Induction*. Routledge, 01 2018.

[TS14] Radwan Tahboub and Yousef Saleh. Data leakage/loss prevention systems (dlp). *2014 World Congress on Computer Applications and Information Systems (WCCAIS)*, pages 1–6, Jan 2014.

[Vei16] Adéle Da Veiga. Comparing the information security culture of employees who had read the information security policy and those who had not: Illustrated through an empirical study. *Information and Computer Security*, 24(2):139–151, 2016.

[WH03] Mark Wilson and Joan Hash. Information technology security awareness, training, education, and certification. *iTL Bulletin*, 10 2003.

[WHP11] Michael Wolf, Dwight Haworth, and Leah Pietron. Measuring an information security awareness program. *RBIS*, 15(3):9–22, 07 2011.

[WS01] Dirk Weirich and Martina Angela Sasse. Pretty good persuasion: A first step towards effective password security in the real world. *Proceedings of the 2001 Workshop on New Security Paradigms*, pages 137–143, 2001.

[ZKSB02] Mary E. Zurko, Charlie Kaufman, K. Spanbauer, and C. Bassett. Did you ever have to make up your mind? what notes users do when faced with a security decision. *18th Annual Computer Security Applications Conference, 2002. Proceedings.*, pages 371–381, Dec 2002.

[Zoh13] Mohammad Zohrabi. Mixed method research: Instruments, validity, reliability and reporting findings. *Theory and Practice in Language Studies*, 3(2):254–262, 02 2013.

Appendix A
Interview Guide

The interview guides used during the interviews conducted before and after the usability test are attached as PDF's below.

Interview Guide

Name: _____
Email: _____
Job position: _____

PRE-TEST INTERVIEW [20 min]

[Thank you for participating in this interview and usability test! As I wrote in the information sheet, we will go through some scenarios for a software solution that is going to be implemented in the organization. The solution will be used to protect documents and other information in the organization. First, I will ask some questions about you and your relationship with data classification. Then we will conduct a usability test to find potential weaknesses and strengths in the system. Finally, I will ask some questions related to your experience from testing the program.]

[Warm-up questions]

1. **Background**
   - 1.1. Have you participated in a usability test before?
     [If not: The purpose of this usability test is not to test you but to test the solution. It is important to emphasize that it is impossible to make mistakes.]
   - 1.2. Are you familiar with the solution that is going to be tested in this usability test?

2. **Introduction**
   - *2.1.* What do you associate with information security*?*
   - 2.2. To what extent do you think about data protection in your daily work?
     - 2.2.1. When sharing documents?
     - 2.2.2. When sharing emails?
     - 2.2.3. Is there any difference between how you handle emails and documents?
   - 2.3. What actions do you take to secure your data?
     - 2.3.1. Why do you take these actions?

[Reflection questions]

3. **Classification routines**
   - 3.1. To what extent are you familiar with the organization's classification policy?
   - 3.2. What is your relationship to document classification in the organization?
     - 3.2.1. How do you apply the policy?
     - 3.2.2. How often do you apply it?
     - 3.2.3. What kind of documents do you classify?

3.3. What do you do when you are unsure about what classification level to apply to a document? (E.g. ask your boss, do not classify)
    3.3.1. **Case:** Imagine that you are unsure about what classification level to apply to a document. Which of the following classification levels would you apply? Personal, public, internal, confidential or highly confidential? Why?
3.4. How do you find that your closest leader sets requirements for document classification?
    3.4.1. How has this affected your classification routines?
    3.4.2. How would a different set of requirements affect your routines? (E.g. audit control)

4. Classification Awareness

4.1. Do you think there is a risk related to classifying wrong or not classify at all?
    4.1.1. Is this something that concerns you?
    4.1.2. How can breaking classification rules be prevented?

5. Expectations for the DLP solution

5.1. What are your initial thoughts about such a solution? Do you see any challenges?

[Winding-up questions]

6. Extra

6.1. Is there anything else you would like to add?

POST-TEST INTERVIEW [20 min]

[Warm-up questions]

1. What are your initial thoughts about the solution?

[Reflection questions]

2. **Usability**
   - 2.1. How intuitive did you find the solution?
   - 2.2. Which parts of the solution were not intuitive?
   - 2.3. Do you find that the solution lacks something?

3. **The Effect on Classification Routines**
   - 3.1. How do you think that the solution will affect your familiarity with the organization's classification policy?
   - 3.2. Can it contribute to making you more confident about what classification level to apply? Why?
   - 3.3. How do you think it can affect your classification routines? (E.g. increase frequency, awareness).
     - 3.3.1. How can it change the way you secure data today?

4. **The Solution's Usefulness**
   - 4.1. How will the solution affect your daily work?
   - 4.2. Do you consider there is a need for this solution in the organization? Why/why not?
   - 4.3. What are potential reasons that could prevent you from using the solution?
   - 4.4. Do you expect to be using the solution? What should be done to motivate you to use the solution?
   - 4.5. Do you have any suggestions for ways to improve the classification routines?

[[Winding-up questions]

5. **Extra**
   - 5.1. Is there anything else you would like to add?

Appendix B
Usability Test

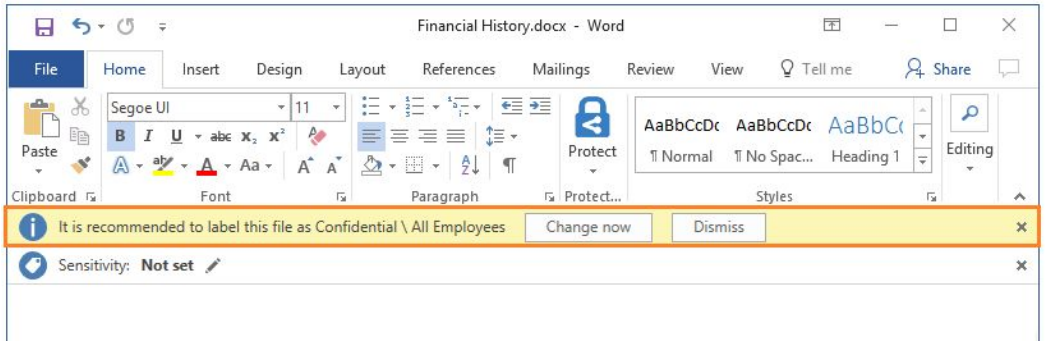
The usability test guide used during the usability test is attached as PDF below.

USABILITY TEST AZURE INFORMATION PROTECTION

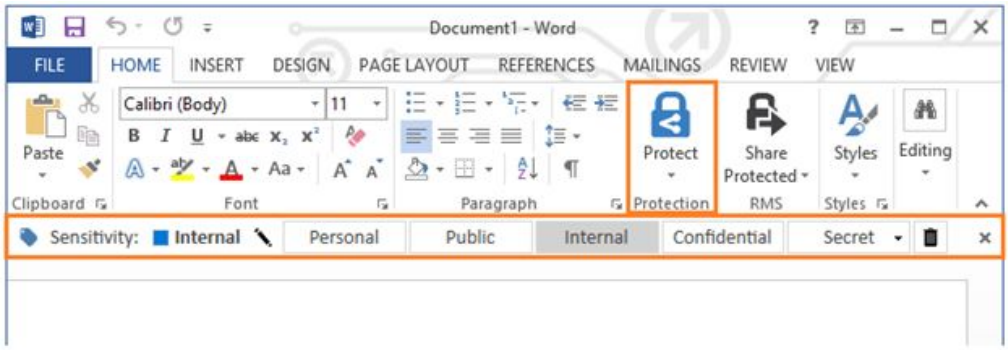
Name: _____

Scenario 1 - Overview *[6min]*
● Open an empty Word document. Use some time to explore the classification functionality. *[1min]* ● There are different levels of control related to what extent the user should apply a classification. These will now be presented and explained for you. *[5 min]* *<u>Questions</u>* - What level of control do you prefer? Why? ❏ 1 - Recommendation ❏ 2 - Default ❏ 3 - Mandatory - Is the degree of information satisfying?
Comment:

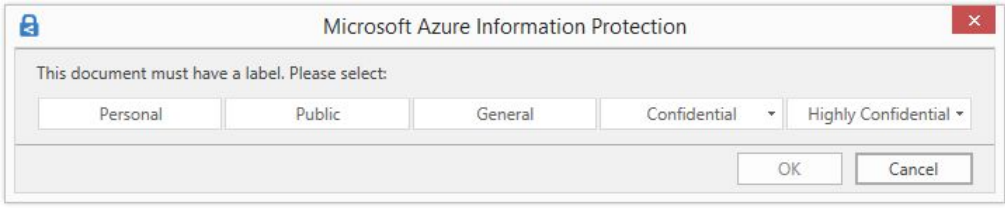
1. Recommend classification level



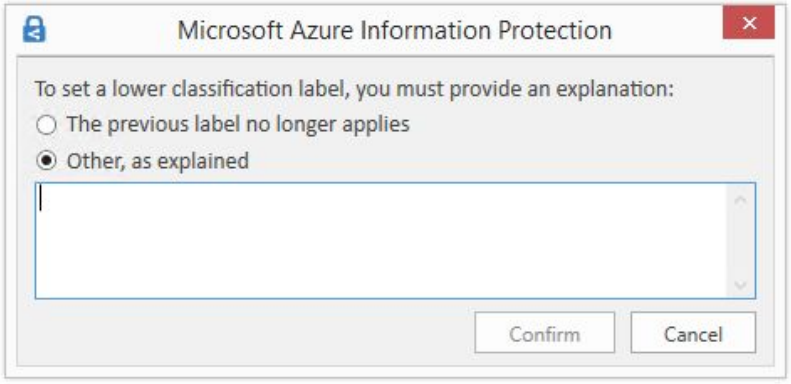
2. Default value



3. Mandatory classification



Scenario 2 - Override classification *[5 min]*
If you want to perform an action that is in conflict with a rule set by the organization, a user override must be set and a justification for the action must be given. *Questions* - In what circumstances is this function useful in your daily work? - In what circumstances is this function an obstacle in your daily work? - Should it be mandatory or optional to give a justification? ❏ Mandatory ❏ Optional - Would you use this functionality?
Comment:



Scenario 3 - Templates *[5 min]*
The organization has several Word templates. Create a new Word document from a template you normally use. *Questions* - How do you prefer that the classification is managed with templates? - Which templates do you want to classify yourself and which do you want to be classified in advance?
Comment:

Scenario 4 - Managing customers classification schemes *[5 min]*
Imagine a case where you work on a project for a big company. The company use a different classification scheme from that used by your organization. *Questions* - Based on what you have seen far, how should the customer's classification scheme be integrated in the solution? E.g. additional buttons or features? - How do you manage this situation today?
Comment:

Scenario 5 - Email *[10 min]*

- Open Outlook. An organization policy states that documents and other content with the classification *Highly Confidential* is not allowed to be sent by email.

  Questions
  - Should this classification level be visible among the other classification levels or should it be hidden?
    - ❏ Visible
    - ❏ HIdden
    - ❏ Other

  - What kind of system action do you prefer in cases when you try to send an email or document classified as *Highly Confidential*? Why?
    - ❏ No pop-up warning
    - ❏ Pop-up warning
    - ❏ Prevent email from being sent
    - ❏ Other

- Imagine a situation where the classification of the attachments does not match the classification of the email.
  Questions
  - How would you prefer this is handled?
  - Can this be a problem in certain situations? Please support your answers with examples.
  - How should this be handled? A Pop-Up?

- The email subject field can potentially leak information about the content.

  Questions
  - Is this something you often keep in mind?
  - How would you prefer that this is managed by the system?
    - ❏ No pop-up warning
    - ❏ Pop-up warning
    - ❏ Prevent email from being sent

Comment:

Appendix C
Questionnaire

Attached below is the questionnaire created in SurveyMonkey. It received 36 responses and was available for a period of two weeks.

Security and Usability

This survey is part of my master thesis study in Communication Technology at NTNU. The participation is anonymous and the data will be kept confidential.

The questions asked are related to your classification routines and perception of a security tool. Your response is valuable information in the process of implementing the tool in the organization.

It will take approximately 5-10 minutes to complete the survey.

Thank you for your participation!

If you have any questions, please send an email to vildeih@stud.ntnu.no.

Vilde Innset Hurum

Security and Usability

Classification of documents

To begin with, some questions about you and your current classification routines will be asked.

1. Do you have a personnel manager role?

- ◯ Yes
- ◯ No

2. What kind of documents do you classify today?

- ☐ None
- ☐ Documents internal to the organization
- ☐ Customer specific documents
- ☐ Other (please specify)

3. I must meet additional classification requirements to the general requirements for the organization (e.g. customer classification schemes)

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

4. I work with projects that may be exposed to information security risks, such as malicious attacks and industrial espionage.

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

5. Does or would working with projects exposed to information security risks affect your awareness regarding information security and the organization’s security policy?

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

6. How often do you classify documents?

- ◯ Every day
- ◯ A few times a week
- ◯ About once a week
- ◯ A few times a month
- ◯ About once a month
- ◯ Less than once a month

7. I believe classification of documents is important

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

8. Getting work done fast has a higher priority than following the security policy

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

9. How often do you create documents based on the organization's templates?

- ◯ Always
- ◯ Usually
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Security and Usability

Classification Policy

This section will ask questions related to your knowledge and use of the organization's classification policy. The classification policy mentioned in some of the questions refers to the organization's policy X.

10. I am familiar with the organization's classification policy

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

11. I am aware of the consequences of classifying wrong

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

12. I am often unsure about which classification level to apply

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

13. When I am unsure about which classification level to apply, I...

- ◯ Classify the document at a low level
- ◯ Classify the document at a high level
- ◯ Ask the project leader
- ◯ Ask the customer
- ◯ Other (please specify)

14. I believe applying the organization's classification scheme is ...

- ☐ a necessity
- ☐ difficult
- ☐ time consuming
- ☐ easy
- ☐ crucial
- ☐ Other (please specify)

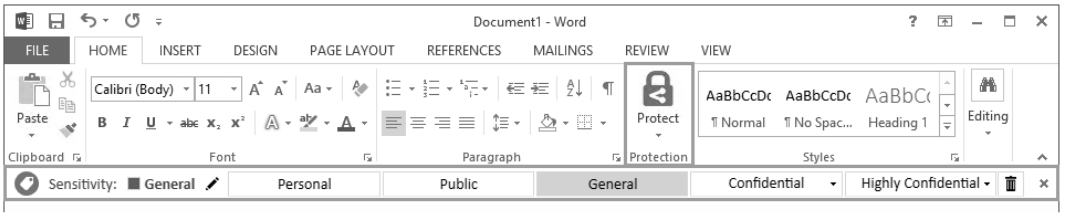
Security and Usability

Applying a Security Tool

This section presents some features of the solution to be implemented and asks questions about

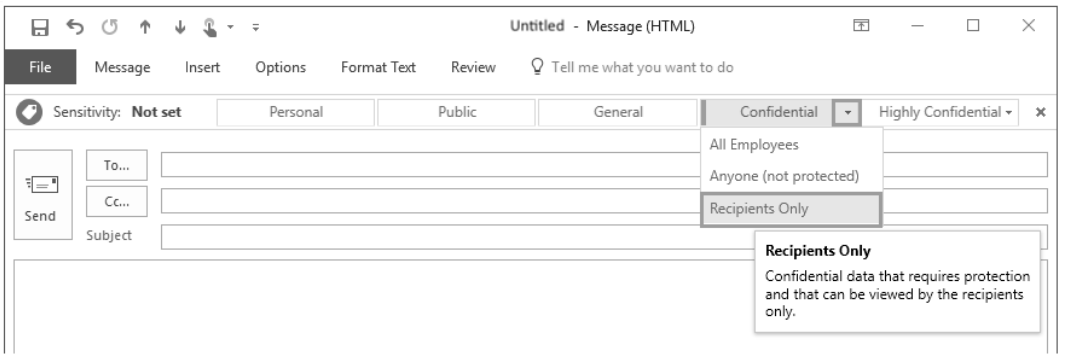
how you think it will impact your daily work.

One of the main features is the use of labels. Labels represent the organization's classification levels and can be applied to documents and emails, adding restrictions to further actions. When a label is set on a document, the data is protected, only people with the correct access priviledges can read and edit it and it is possible to track its location.



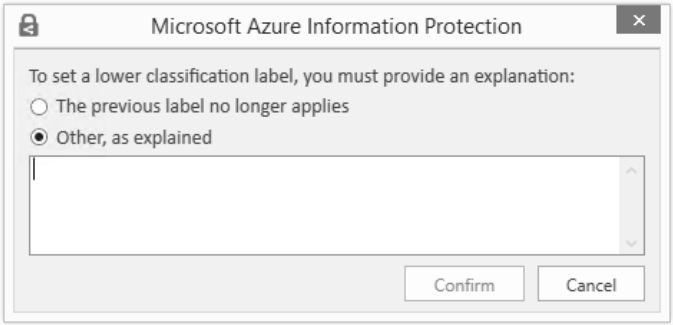
15. The toolbar outlined above shows the organization's classification levels options and will be added to Office programs (Word, Excel, PowerPoint and Outlook). In order to classify a document you choose the appropriate label from the toolbar. How do you perceive the toolbar?

- [ ] Appropriate location
- [ ] Appropriate size
- [ ] Easy to ignore
- [ ] Easy to understand
- [ ] Other (please specify concerns)



16. Above you see the toolbar as it appears in Outlook. A policy in the organization is that content classified as "Highly Confidential" is not allowed to send by email. If a user tries to send an email with content classified as "Highly Confidential", how should this situation be handled in the solution?

- ◯ The solution should prevent the email from being sent
- ◯ The solution should show a warning with options and prevent the email from being sent
- ◯ The solution should show a warning with options and let the user decide whether to send it
- ◯ Other (please specify)



17. In some cases you may want to lower the classification level of a template. However, the organization might have a restriction saying that the template needs to be labeled as "Confidential". To solve this, a justification box similar to the one shown above can be implemented. Would you find this feature usable?

- ◯ Yes
- ◯ No

18. Given that the justification feature above is implemented in the solution, in what cases should it be used?

- ◯ Only when changing the label from the two most confidential levels to a lower classification level
- ◯ Only when changing from levels more than one classification level apart
- ◯ Only when changing to a lower level
- ◯ In all cases when changing classification level
- ◯ Other (please specify)

19. I believe that the security tool should be applied to ...

- ◯ Both emails and documents
- ◯ Documents only
- ◯ Emails only
- ◯ Neither documents nor emails

20. The security tool seems clear and understandable

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

21. Using the security tool will require low effort

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

22. The security tool will enable me to more easily practice the organization's classification policy

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

23. The security tool will help me better protect customer data

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

24. The security tool will decrease my job productivity

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

25. This tool will be useful in my job

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

26. The security tool will make me more aware of the organization's classification policy

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

27. The security tool will improve my classification routines

- ◯ Strongly agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly disagree

28. What factors would prevent you from using of the tool?

- [ ] Lack of information from the organization
- [ ] Unnecessary features
- [ ] Lack of enforcement from the organization
- [ ] I do not use Office365
- [ ] I do not see the usefulness
- [ ] Lack of experience
- [ ] Lack of time
- [ ] Annoying pop-ups
- [ ] Other (please specify)

29. There is a need for this security tool in the organization

- ○ Strongly agree
- ○ Agree
- ○ Neither agree nor disagree
- ○ Disagree
- ○ Strongly disagree

30. I intend to use the security tool

- ○ Likely
- ○ Unlikely

* 31. What do you expect from the organization before the solution is implemented?

- [ ] Enforcement by the organization
- [ ] Workshop/Training
- [ ] Online user guide
- [ ] Information (meeting, email)
- [ ] Other (please specify)

Security and Usability

End of Survey

Thank you for completing this survey!

32. Please write down any additional comments you might have.

Appendix D
NSD

A request was sent to the Norwegian Center for Research Data (NSD). Attached is a PDF in Norwegian of the received confirmation to collect and process personal data.

NSD NORSK SENTER FOR FORSKNINGSDATA

NSD sin vurdering

Prosjekttittel

Implementation of Data Loss Prevention in an organization with focus on user experience and its effect on classification routines

Referansenummer

827012

Registrert

18.11.2018 av Vilde Innset Hurum - vildeih@stud.ntnu.no

Behandlingsansvarlig institusjon

NTNU Norges teknisk-naturvitenskapelige universitet / Fakultet for informasjonsteknologi og elektroteknikk (IE) / Institutt for informasjonssikkerhet og kommunikasjonsteknologi

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Maria Bartnes, maria.bartnes@sintef.no, tlf: 45218102

Felles behandlingsansvarlige institusjoner

SINTEF AS / SINTEF Digital

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Vilde Innset Hurum, vildeih@stud.ntnu.no, tlf: 91894462

Prosjektperiode

09.01.2019 - 05.06.2019

Status

29.01.2019 - Vurdert

Vurdering (2)

29.01.2019 - Vurdert

ENDRINGSVURDERING

Vi viser til endringer i meldeskjema og dialog på meldinger. Vi forstår det slik at endringen kun omfattet behandling av anonyme opplysninger. Vi ser derfor ikke at det er behov for ny vurdering og vurderingen fra 17.01.2019 gjelder fortsatt.

Vennlig hilsen
Marianne H. Myhren

17.01.2019 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 17.01.2019, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD ENDRINGER
Dersom behandlingen av personopplysninger endrer seg, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. På våre nettsider informerer vi om hvilke endringer som må meldes. Vent på svar før endringer gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 05.06.2019.

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er

avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Marianne Høgetveit Myhren
Tlf. Personverntjenester: 55 58 21 17 (tast 1)