EXAMENSARBETE

GDPR: Securing Personal Data in Compliance with new EU-Regulations

Hadi Bitar
Björn Jakobsson

Informationssäkerhet, master
2017

Luleå tekniska universitet
Institutionen för system- och rymdteknik

LULEÅ TEKNISKA UNIVERSITET

A7009N

Information Security Master Program

MASTER THESIS – 30 credits

GDPR: Securing Personal Data in Compliance with new EU-Regulations

Luleå University of Technology

Spring Semester 2017

Supervisor:

Prof. Tero Päivärinta

Authors:

Hadi Bitar (hadbit-1)

Björn Jakobsson (bjjakh-2)

Foreword

We would like to thank our supervisors Lars G Magnusson at Tieto AB, Professor Tero Päivärinta and all our opponents at Luleå University of Technology for their feedback, constructive critique, and overall support during the entire process. We also would like to thank everyone at Tieto AB Luleå, and especially Pasi Hautamäki for his support and for providing us with the opportunity to write this thesis. A huge thank you goes out to all the security professionals that participated in our study and to Tieto and Verisec for the support and information we received. Finally, we would like to thank our families and friends for their continued support in our quest for education and learning.

Abstract

New privacy regulations bring new challenges to organizations that are handling and processing personal data regarding persons within the EU. These challenges come mainly in the form of policies and procedures but also with some opportunities to use technology often used in other sectors to solve problems. In this thesis, we look at the new General Data Protection Regulation (GDPR) in the EU that comes into full effect in May of 2018, we analyze what some of the requirements of the regulation means for the industry of processing personal data, and we look at the possible solution of using hardware security modules (HSMs) to reach compliance with the regulation. We also conduct an empirical study using the Delphi method to ask security professionals what they think the most important aspects of securing personal data, and put that data in relation to the identified compliance requirements of the GDPR to see what organizations should focus on in their quest for compliance with the new regulation. We found that a successful implementation of HSMs based on industry standards and best practices address four of the 35 identified GDPR compliance requirements, mainly the aspects concerning compliance with anonymization through encryption, and access control. We also deduced that the most important aspect of securing personal data according to the experts of the Delphi study is access control followed by data inventory and classification.

Table of Contents

List of Abbreviations

Abbreviation	*Explanation*
A29WP	Article 29 Data Protection Working Party
AES	Advanced Encryption Standard
CIS	Center for Internet Security
CISSP	Certified Information Systems Security Professional
CMAC	Cipher-based message Authentication Code
CMVP	Cryptographic Module Validation Program
CSC	Critical Security Control
CSP	Critical Security Parameter
DES	Data Encryption Standard
DPA	Data Protection Authority
DPD	Data Protection Directive
DPO	Data Protection Officer
DSA	Digital Signature Algorithm
DSM	Data Security Manager
ECDSA	Elliptic Curve Digital Signature
ECJ	European Court of Justice
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EFTA	European Free Trade Association
EU	European Union
EuroPriSe	European Privacy Seal
FIPS	Federal Information Processing Standard
GDPR	General Data Protection Regulation
GMAC	Galois Message Authentication Code
HMAC	Hash-based Message Authentication Code
HSM	Hardware Security Module
IDS	Intrusion Detection System
IPsec	Internet Protocol Security
ISO	International Organization for Standardization
IT	Information Technology
KEK	Key Encryption Key
KMS	Key-Management System
MAC	Message Authentication Code
MDS	Manipulation Detection Code
MS-CAPI	Microsoft Crypto Application Programming Interface
NIST	National Institute of Standards and Technology
OWASP	Open Web Application Security Project
PC	Personal Computer
PCI DSS	Payment Card Industry Data Security Standard
PKCS	Public-Key Cryptography Standard
PKI	Public Key Infrastructure
RNG	Random Number Generator
RNG	Random Number Generator
RQ	Research Question
RSA	Rivest, Shamir, Adleman
SA	Supervisory Authority
SHA	Secure Hash Algorithm
SME	Small and Medium-sized Enterprises
SNIA	Storage Networking Industry Association
SSH	Secure Shell
TLS	Transport Layer Security
VPN	Virtual Private Network

Table of Figures

1 Introduction

Up until 2016 all 28 European Union (EU) member states had their own laws regarding the collection, storing and processing of personal information about its citizens in conjunction with the Data Protection Directive 95/46/EC (DPD) issued by the EU on October 24th, 1995. On April 27th, 2016, a new General Data Protection Regulation (henceforth GDPR) was adopted by the European Commission and will be in full effect on May 25th, 2018, replacing all local and national data protection laws in the EU's member states as well as replacing the DPD (European Commission, 2015). This new regulation includes many new rules for organizations and enterprises operating in the EU to adhere to and understand. Among the most discussed news in GDPR is the introduction of a new fining system that is part of the new regulation. This system contains the clause that any organization not in compliance[1] with the new regulation may be fined up to 4% of their annual global profit, which hopefully will work as an effective deterrent and encouragement for organizations and enterprises to be compliant with GDPR as soon as possible (European Commission, 2015).

Other than introducing powerful fines, the GDPR contain many new and interesting regulations and rules for organizations and enterprises to adopt and adhere to. One major positive change is the introduction of the "one stop shop", which means that the organizations and enterprises operating in the EU only needs to be in contact with one data protection authority instead of one in each EU country. The regulation states that this primary data protection authority is to be selected based on where the organization or enterprise's main base of operations within the EU is located (European Commission, 2015).

All organizations and enterprises that are processing personal data must appoint their own "Data Protection Officer" (DPO). The DPO may be employed or contracted as a service and the DPO must have the corresponding expertise to the processed data in question. There are some exceptions to this rule based on the size of the organization and the amount of data that is being processed, for example, Small and Medium Enterprises (SME's) (European Commission, 2016), need not appoint a DPO if they are not processing enough personal data (European Commission, 2015).

Data protection by design and per default will become the new norm in the EU. All products and services aimed or used in the European market must be designed with data protection in mind from the earliest stages of development. And products and services shall by default have the privacy settings in a privacy-friendly mode (European Commission, 2015).

Every citizen and visitor in the EU is the legal owner of any data about them originating within the union, and has the right to be forgotten under the new regulations. This means that the owner of the privacy data, has the right to have their data removed from any platform or service if there are no legitimate grounds for keeping it (European Commission, 2015).

Individuals has the right to move their own personal information from one service provider to another of their own choice. This is done for smaller companies to be able to compete with bigger companies

[1] "Compliance" in the context of this paper is defined as "conformity in fulfilling official requirements" (Merriam-Webster, n.d) It will be up to the European courts with the assistance of the GDPR Supervisory Authorities to determine what those requirements actually entails in future court cases.

for the customer data and for making it clear that the data always is owned by the individual, and that they themselves decides where it is used and stored. This also means that the individual must be better informed on how the data provided is being used and that they have actual access to it as well as to information on how it is being used and for what (European Commission, 2015).

Another big impact that the GDPR will have is on the information about data breaches and data leaks. Under GDPR, breach notification will be mandatory and failure to inform the individuals and the supervisory authorities[2] about any loss of data as fast as possible will result in fines for the organization or enterprise in question (European Commission, 2015).

1.1 Problem Area

This thesis is commissioned by Tieto AB as a study to find if Hardware Security Modules (HSMs) can be utilized as a tool to reach some level of compliance with GDPR and if they then can be part of an offer to customers, simplifying the compliance process somewhat.

The new regulations introduced under GDPR will have implication for almost all organizations and enterprises that collect, stores or processes personal data and that operates with EU citizens' data. Since the regulation comes into effect on May 25th, 2018 there is a lot of work to do for organizations and enterprises to become fully compliant before that. What measures are needed, what technologies should be used and how is compliance with the regulations achieved as "easily" and smoothly as possible? The regulation itself only briefly mentions methods for organizations and enterprises to use and apply on the data that is to be controlled under the new regulations, these proposed methods are encryption and/or to apply pseudonymisation[3] to the data to render it unreadable or unintelligible if stolen or lost.

Is there a way for organizations to reach compliance with GDPR using some technology or method? The regulation states that using encryption, if used properly, means that notice to data owner at a breach no longer is necessary (Article 34 paragraph 3a) and that encrypting the data at rest and in transit should mean that the organization is in compliance with GDPR regulation and should not face any fines or issues if data is lost or leaked since the data maintains its confidentiality if properly encrypted (Article 83 paragraph 2c and 2d) (European Union, 2016). This also means that the key-management within the organization must be properly applied and utilized since encrypted data with a poor key, or even a lost key means that the data loses its confidential status (Chandramouli, et al., 2014).

This is where the use of hardware encryption comes into the picture, there are devices and systems on the market that is designed to protect data both in storage and in transit by applying powerful encryption schemes to it using a specific hardware device called Hardware Security Module or HSM (other common names include: Tamper Resistant Security Device/Module, Cryptographic Accelerator, Secure Application Module, Hardware Cryptographic Module). The HSM contains the hardware necessary to encrypt and de-crypt data without putting any additional strain on the storage server's CPU or other resources, it also takes care of the key-management and does all this within a tamper-proof unit designed to react to any attempt of malicious intrusion or modification.

[2] More on supervisory authorities in chapter 2.2.1

[3] The processing of personal data in such a way that it cannot be associated with the specific data subject without needing to resort to additional information. This additional information is stored and secured separately where the necessary measures are taken to keep the information from being linked to a specific individual (Bolognini & Bistolfi, 2016).

1.2 About Tieto AB

Tieto is an IT service company active in more than 20 countries with approximately 13000 employees. The organization is one of the largest IT service providers in Europe and the largest one in the Nordic region, giving it a noticeable global presence through its product development business and global delivery centers. The organization provides full life-cycle services for both the private and public sectors, in the field of communications and embedded technology, including financial services, healthcare and welfare, industrial, consumer services and industry solutions.

1.3 Aim of Study and Research Question

We aim to examine if compliance with GDPR can be achieved through the implementation and use of HSMs, and what the residual risks of such an approach are with regards to accountability to provisions of the GDPR. GDPR is written in a purely legal format and lacks real suggestions for ways to achieve compliance and the only technical suggestions mentioned in the regulation to achieve data protection is either encryption or pseudonymisation.

We will investigate how using hardware security modules (HSMs) to comply with the regulation with regards to encryption can help, and try to determine if HSM is a viable way to become compliant with the GDPR encryption articles.

We will also look at different methods of encryption available and briefly describe them to try to further increase the readers understanding of the issue. In addition to that, we will also briefly describe the problems with key-exchange and key-management when dealing with cryptography.

We will also conduct a Delphi-study with a panel of experts within the field of information security to find out what they consider to be the most important aspects when dealing with the security of personal data. In the end, we hope to provide the reader with a list of aspects that are not mitigated or addressed by using HSMs.

The research questions (RQ) for this work are:

- How can the use of HSM aid in achieving compliance with GDPR?
- What GDPR requirements would be left un-addressed by using such an approach?

1.4 Delimitations

The thesis will mainly focus on the technical suggestions mentioned in the GDPR, specifically encryption, therefore, pseudonymisation will not be discussed in the thesis. This also means that the scope of this thesis will focus on the aspects of the GDPR that can be addressed by technology and technical measures. Many parts of the GDPR is addressed by purely managerial methods, such as request for consent and lawful reasons for processing, these aspects will not be covered in the thesis.

Since the study focuses on the protection of personal data processed and/or used by organizations with a focus on Tieto AB, there will be no differentiating between protection of data in development state or in operational state. Meaning that regardless if the organization that is processing and using personal data is doing it for IT-development or for already operational IT is insignificant, the same law will affect both cases in the same way.

1.5 Limitations

Due to the sensitivity of the topic at hand it might be difficult to gather expert panel members for the Delphi study where these experts might have a slight fear of disclosing information that might affect their company or organization and put them at risk. The same would apply when trying to find

informants for the interviews. This could result in the lack of empirical data which would jeopardize the validity and reliability of the study. The solution would be to create questions that discuss the research topic in general matter where the informants are more comfortable in answering them, additionally giving the research valuable data that can be further analyzed and discussed.

1.6 Structure

The fundamentals of the thesis are described in chapter 2 in the form of a theoretical framework. Chapter 3 describes the techniques used to gather the data for this text, in the form of empirical studies as well as the literature study. The results and analysis from the literature- and empirical study are then presented in chapter 4, and finally the discussion about the findings is found in chapter 5.

2 Theoretical Framework

This chapter describes the different theories and key concepts contained in the thesis, starting with an explanation of the new general data protection regulation (GDPR), and continuing with the description of cryptography and hardware security modules.

2.1 EU-GDPR

The history of data protection regulations, directives and conventions in the EU dates to the early 1970's following rapid advancements in the field of information technology and increased debate about privacy issues that followed the increased processing of personal data in computers, when the federal state of Hessen in Germany instituted the first national data protection law in the world (Wilhelm, 2016). In 1985 the Council of Europe Data Protection Convention came into effect that contained the first international legally binding principles regarding data protection (de Hert & Papakonstantinou, 2014).

In 1995 the DPD 95/46/EC was released, directing member states in how individuals within the EU shall be protected with regards to the processing of personal data as well as the free movement of such data between EU member states.

In 2009 the EU Commission launched a review of DPD 95/46/EC and found several aspects that could be improved upon, such as the creation of an EU internal market with coherent legislation for multinational companies to adhere to instead of different laws in different EU member states, this way globalization issues and the enforcement of the data protection rules could also be addressed and streamlined (European Commission, 2010). The first proposal for the new regulation was released in January of 2012 (European Commission, 2012) and was then discussed and changed in various instances of the European Union and its member states. Finally, in April of 2016 the Council of the European Union and the European Parliament adopted the proposal and it became a regulation and entered into force on May 4th, 2016. Article 99 of the regulation states that it applies to all member states starting from 25th of May 2018 (European Union, 2016).

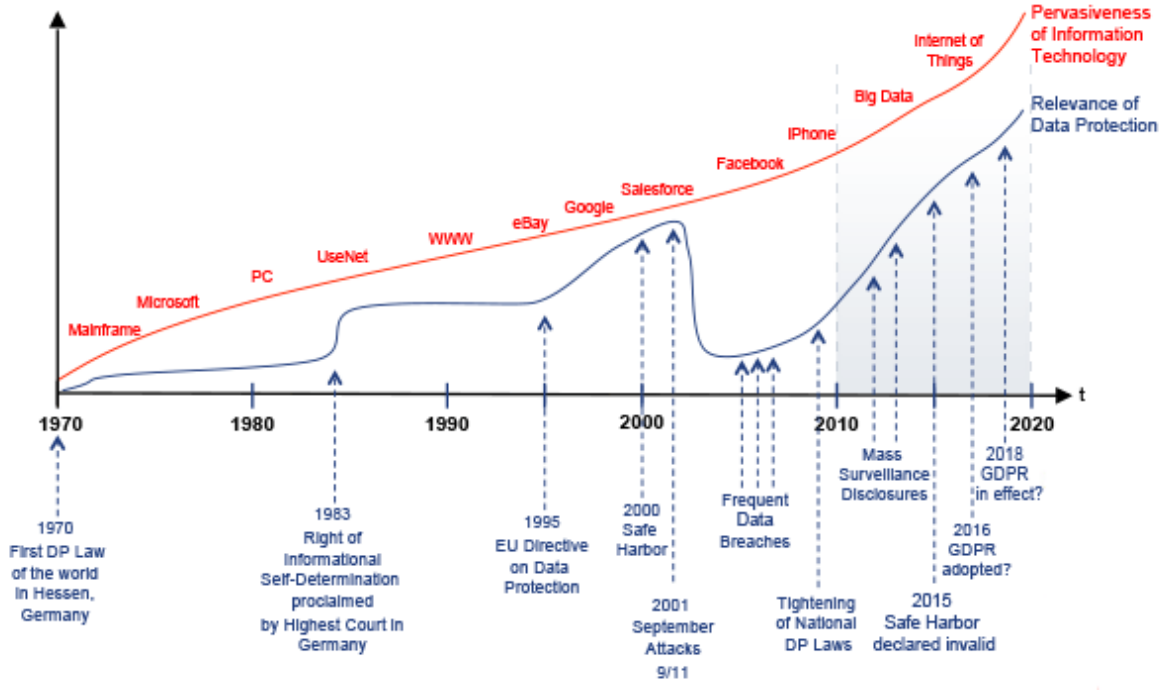

Figure 1 – History of the GDPR (Wilhelm, 2016)

The main changes introduced to the new Regulation (GDPR) is that a right to be forgotten has been introduced (article 17 of the GDPR) and individuals will have easier access to their data and the right to understand how their data is being processed (article 15 of the GDPR). Individuals will also have a right to move their data between service providers (article 20 of the GDPR) and to know when a data controller[4] or data processor[5] has lost data due to an intrusion or hack (article 34 of the GDPR). There are also provisions in the GDPR that states that data protection is to be part of products and services from the earliest stages of development ("data protection by design" article 25 of the GDPR) and that privacy settings per default are set to levels that ensures and prioritizes data protection ("data protection by default" article 25 of the GDPR) (European Commission, 2015; European Union, 2016).

2.1.1 Personal Data

The GDPR defines personal data in Article 4 paragraph 1 of the GDPR as; "any information relating to an identified or identifiable natural person" (data subject) (European Union, 2016). This means that all data processed and stored that may be linked to an actual citizen of the European Union falls under the application of the GDPR. The same definition also states that direct or indirect identification of a person, by using data references such as a name, an identification number, location data, or even factors such as gender, economic and cultural identity constitutes personal data.

There is a debate regarding the width of the definition of personal data in the GDPR, and the question is how it will be implemented in the future when the GDPR comes into effect since no court cases exist yet to provide precedence for the interpretation of the legal text. Two different approaches to the definition of personal data, an absolute- and a relative approach, have been discussed by Gerald Spindler and Philipp Schmechel in their 2016 article titled "Personal Data and Encryption in the European General Data Protection Regulation" (Spindler & Schmechel, 2016).

- Absolute approach – The absolute approach means that data that is encrypted still would be considered as personal data and would still be subjected to the application of the GDPR. The reasoning for this approach is that the encrypted data is basically only a form of pseudonymized data and that is still is possible to convert it to readable data by using the cryptographic key or by cracking the encryption algorithm used to encrypt the personal data. This approach does not take issues such as time and cost of breaking the algorithm or to gain unauthorized access to the keys into account at all and even theoretical chances of advertently accessing the protected personal data is included.
- Relative approach – The relative approach on the other hand does take the aforementioned issues into account, meaning that it takes the effort required to be able to read the personal data and identify the data subject into account as well. This means that personal data that is protected by encryption requiring the use of a secured key or otherwise substantial time and cost to crack can be regarded as anonymized data and therefore possibly be exempt from the application of the GDPR.

[4] The entity using the results from the processing of personal data, i.e. the entity that collects the personal data from its users and/or customers and who has an interest in the processing of the data (Treacy, 2010).
[5] The entity that carries out the actual processing of personal data, on behalf of - and based on requirements from the controller (Treacy, 2010).

The position within the European Court of Justice (ECJ) at the time of this study seems to be leaning towards an absolute approach, according to Spindler and Schmechel, based on opinions from Article 29 Working Party[6] (A29WP) and the ECJ Advocate General Campos Sánchez-Bordona[7] (Spindler & Schmechel, 2016).

2.2 Primary Effects of GDPR

The new regulation demands that action be taken to protect the data owner's personal and private data in a sufficient way. The data shall not be accessed by un-authorized users and personnel shall have access granted based on least privileges (meaning that root and system administrators should not have access to the encryption keys). There also needs to be measures in place that limits the effects of a possible data breach, meaning that data that is lost or stolen is un-readable (European Union, 2016).

2.2.1 Supervisory Authority

A result of the implementation of the GDPR is the creation of supervisory authorities (SA) with the task of regulating and supervise the processing of personal data in the EU. These authorities are the ones responsible for compliance validation.

Article 51 of the GDPR states that independent SA shall be established, at least one in each member state and if there are multiple SA one shall be designated as the lead SA. Businesses that have multiple establishments in the EU need only to report to the one SA that is based where the business "central administration" is located according to article 4 paragraph 16 of the GDPR (European Union, 2016, p. 34) – This means that organizations only need to deal with one SA for their GDPR compliance issues and this is what the term "one stop shop" in the regulation means (GDPR Recitals; 124-128) (European Union, 2016, p. 7).

Article 58 of the GDPR states that each SA shall have investigative powers to perform data audits on data processors and data controllers and to obtain all pertinent information the SA requires to perform its task. They are also granted access to any of the processors and controllers premises to carry out such auditing tasks. The same article also grants the SA corrective powers such as the power to issue warnings and reprimands to the processor and controller, and to order the same to comply with requests from data subjects, and to comply with the GDPR regulation and to rectify or erase personal data.

In addition to this the SA have authorization and advisory powers, such as advising the data controller through consultation and to authorize data processing if the member states law requires that.

According to the EU justice website there will be several different data protection authorities in place, on national, union, EFTA and in "third countries" (Countries outside of the EU) (European Commission, 2016). There is however no clear explanation available about the connection between SA and these other authorities such as the Data Protection Authorities (DPA), European Data Protection Supervisor (EDPS), European Data Protection Board (EDPB) and Data Protection Officers (DPO) at the time of this thesis, we therefore assume that they are all types of SA.

[6] The Article 29 Working Party is an independent body set up under article 29 of the DPD to advice on data protection issues. Its members are national DPAs and the EDPS. The A29WP's opinions are not legally binding but are very influential (European Commission, 2016).

[7] Opinion of Advocate General Campos Sánchez-Bordona, delivered on 12 May 2016, Case C-582/14 – Patrick Breyer v Bundesrepublik Deutschland.

Data Protection Authorities (DPA) is a supervisory authority which oversees monitoring the processing of personal data within its jurisdiction, providing advice to the data controllers regarding legislative and administrative measures relating to the processing of personal data and hearing complaints lodged by citizens regarding the protection of their data protection rights. It is also the DPAs role to determine if data controllers and data processors have done a good risk analysis and impact assessment prior to starting the processing as well as assess the same after a data breach (European Union, 2016, pp. 17-18).

European Data Protection Supervisor (EDPS) is an independent EU body responsible for monitoring the data processing of citizens done within the context of the EU institutions and bodies. The EDPS has a similar mission to the DPA but aimed at EU internal processing for different purposes. The EDPS keeps a record of all data processing that poses potential risks to individual privacy and investigates complaints lodged by people whose data are being processed within the EU institutions and bodies. They also conduct inspections and offer consultations on all matters of personal data processing (European Data Protection Supervisor, 2017).

European Data Protection Board (EDPB) is described in the GDPR as a "body of the Union" (Article 68), and is describes as the coordinating entity between DPAs in Europe. The board is composed of the head of one SA of each member state and of the EDPS and will act as a means of consistency in the ruling and application of the GDPR as well as an advisory organ to the Commission. The EDPB will also be responsible for issuing guidelines, recommendations and best practices on procedures and measures of the GDPR to all member states (See article 70 of the GDPR for full list of tasks). Disputes between DPAs will be resolved in the EDPB according to article 65.

Data Protection Officer (DPO) is described in the GDPR article 37-39 and is a position that data controllers and data processors must designate if certain conditions on the size and scope of the processing of personal data is met (article 37 of the GDPR), basically this means all public-sector bodies, organizations with more than 250 employees (article 30 of the GDPR) and those organizations where the monitoring of data subjects is a core activity (article 37 of the GDPR). The data protection officer can be shared between different organizations or be hired as a consultant (article 37 of the GDPR), but must be free from influences from the data controllers and processors on how to do its job (article 38 of the GDPR). The DPO tasks are stated in article 39 of the GDPR as the following:

- Advisor to the data controllers or data processor on the obligations of the GDPR
- Monitor the data controllers and data processors compliance with GDPR
- Advisor on data impact assessments
- Cooperate with – and act as point of contact for the SA's

2.3 Non-compliance

Organizations not in compliance with the GDPR may be subject to extensive administrative fines according to a new infringement system introduced with the GDPR. The fines that may be imposed on organizations in violation with the regulation are substantial, up to 4% of global revenue or 20 million Euro, whichever is higher, for serious breaches against for example non-compliance with an order from an SA, the overall lawfulness of the processing, and consent of the data subjects. The lower level fines are set to up to 2% of global revenue or 10 million Euro, whichever is higher for breaches against data protection by design and default, breach notifications to the data subject and designation of a DPO (Article 83 of the GDPR (European Union, 2016)).

2.3.1 Breach Notification

ISO/IEC 27040 defines a data breach as a "compromise of security that leads to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to protected data transmitted, stored, or otherwise processed" (ISO, 2015, p. 2).

Article 33 of the GDPR states that when a breach of confidentiality is detected it must "without undue delay" be reported to the supervisory authority, "…unless the data breach is unlikely to result in a risk to the rights and freedoms of a natural person" (European Union, 2016).

In article 34 of the GDPR it is stated that the data subject must be informed of the breach if "the data breach is likely to result in a high risk to the rights and freedoms of natural persons" (European Union, 2016), article 34 paragraph 3a then states that this communication is not required if the data controller has "implemented appropriate technical and organizational protection measures, and those measures were applied to the personal data affected by the personal data breach, in particular those that render the personal data unintelligible to any person who is not authorized to access it, such as encryption" (European Union, 2016).

It should be noted that all data breaches regarding personal data, encrypted or otherwise, always shall be reported to the appropriate DPA (Article 29 Data Protection Working Party, 2014).

2.3.2 Privacy by Design and Default

To mitigate the risks involved with processing and storing personal data there is a concept of privacy by design and default included in the GDPR. This concept is there to provide the framework for developing integrity safeguarded systems and designs throughout the entire project lifecycle. From the conceptualization, all the way through development and deployment to the decommissioning of systems.

Some such methods for privacy are described in the GDPR:

Data minimization (article 5 and 25 of the GDPR)*,* the concept of not storing more data than is necessary for the task at hand. Also, means to minimize the actual data that might identify a person in the database for instance (The Swedish Data Protection Authority , 2012).

Access controls for personal data (article 29 of the GDPR) is an integral part of any sensitive system*,* and making sure that only users with a need to know may access the personal and sensitive parts of the data. Making sure that systems have access controls in place so that access rights cannot be elevated, transferred or faked (The Swedish Data Protection Authority , 2012).

Data protection, IT-systems that deal with personal information must be secured throughout its lifecycle. To employ such functions after the fact, when a system already is deployed, is both difficult and likely to be expensive. Instead the systems should be designed from the beginning with the security of its data in mind. This means that functionality for encryption (recital 83 of the GDPR) should be built in for communication and storage (The Swedish Data Protection Authority , 2012). There also need to be clear rules and policies in place to ensure that the users of the systems are aware and trained to react to data breaches and other incidents. There should also be an audit trail built into the system, with logs and traceability of all access made to the system. The system also requires a safe backup method so that data and its audit trails are recoverable after a disastrous incident. Finally, there needs to be a method and a process for the safe destruction of the data when it is no longer useful or required (The Swedish Data Protection Authority , 2012). Chapter 2.4 below will further expand in the data protection theory.

User friendly systems can be utilized to guide the users of the systems to work in a way that promotes privacy by default, for example by not gathering excessive data and by not displaying data that is not necessary. The systems can have an easy function for the removal of data after its use and to automatically remove sensitive and unnecessary data before archiving. In addition to this, the system can have privacy default functions for creating presentations, diagrams and statistics where it automatically anonymizes or removes the sensitive data from the end report (The Swedish Data Protection Authority , 2012). Systems designed for the data subject to use should have clear and understandable information describing what the information the data subject enters will be used for and a function that clearly asks the user for consent (article 4, paragraph 11 of the GDPR) (European Union, 2016).

2.3.3 Impact Assessment

Article 35 of the GDPR states that a data protection impact assessment should be carried out prior to the processing of personal data, if the processing is likely to result in a high risk to the rights and freedoms of natural persons. Basically, this means that the controller must analyze the risks of the processing and address the identified risks with technical or organizational measures (recital 84, 90-94 and article 35 of the GDPR) (European Union, 2016). By doing this impact assessment, the controller also gets tangible proof that the processing has been assessed and that risks have been addressed (Wright, 2013). The DPAs shall together with the EDPB publish lists of processing that require impact assessment and if new technologies are used to process personal data the controller must perform an impact assessment before the actual processing starts (article 35 of the GDPR) (European Union, 2016).

2.4 Data Protection

The term data protection relates to the process of safeguarding data from both internal and external threats, whether it is at rest or in motion. The key to data protection, in general, is to work accordingly with the three core principals of information security known as the CIA-triad (confidentiality, integrity and availability) (Agarwal & Agarwal, 2011).

- **Confidentiality:** To ensure that the data is not disclosed or made available to unauthorized entities.
- **Integrity:** To ensure that the data remains in its original state, meaning that it has not been manipulated or tampered with by any unauthorized entity.
- **Availability:** To ensure that the data is available when needed and that the system hosting the data is fully functional without any faults.

As mentioned previously, article 34 of the GDPR states that if a breach were to occur, the controller of the data will not need to notify the individuals of the data breach in case sufficient technical security measures ensuring the confidentiality of the data have been implemented, such as encryption. Article 83 of the regulation continues with stating that by ensuring the confidentiality of the data, both at rest and in transit, the controller of the data will not be subjected to the fines stated in the regulation in case of a data breach (European Union, 2016). This indicates that encryption should be an initial solution for protecting data within organizations (Tankard, 2016) and an essential countermeasure against various threats and vulnerabilities (Solterbeck, 2006).

Other than implementing encryption, it is important that the organization has an appropriate key-management solution where the keys for encrypting and decrypting data are stored and handled using appropriate security controls and measures (Tankard, 2016). The loss or mishandling of an access key would jeopardize the confidentiality of the data (Chandramouli, et al., 2014), this could lead to the allegations that the organization did not apply the sufficient technical controls to protect their data,

forcing them to pay the fines set by the regulation (Tankard, 2016). Even though encryption sets a strong foundation for protecting the data within organizations, it alone will not suffice as a solution. There are other aspects that should be considered and implemented to work in harmony within the organization's information security infrastructure to mitigate the risks of data disclosure as much as possible, such as access controls, role management and auditing (Solterbeck, 2006; Tankard, 2016).

2.4.1 Protecting Data at Rest and in Motion

When speaking of data at rest, we speak of information that are stored in different types of physical media whether the media is optical, magnetic or on a piece of paper. When speaking of data in motion, we speak of information that is being transferred between different components, nodes, programs, locations and during an input/output process.

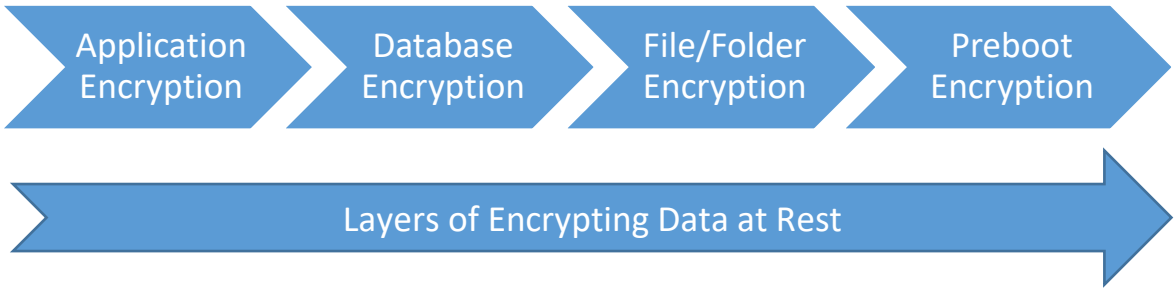

Figure 2 – The four stages of encrypting data at rest (Solterbeck, 2006)

Figure 2 illustrates the different categories one should keep in mind when protecting data at rest (Solterbeck, 2006):

- **Application Encryption:** Encrypting application data based on the fields within that data, such as username, password etc. then mapping these fields to each user's privileges.
- **Database Encryption:** To encrypt database fields or columns along with assigning access rights to the data contained within the database giving access only to authorized users.
- **File/Folder Encryption:** To manage and control access to individual files and folders within an organization based on the organizational policies.
- **Preboot Encryption:** To encrypt data within servers and require proper authentication and authorization of users before booting up devices and granting access to any corporate data.

These four layers can be considered essential when protecting data at rest and covering all four layers properly would heavily mitigate the risk of data getting compromised (Solterbeck, 2006).

To protect data in motion, virtual private networks (VPN) were developed for a more secure data transfer. There are different VPN tunnels used for encrypting and authorizing traffic such as Transport Layer Security (TLS) which uses a combination of symmetric and asymmetric encryption, which will be explained in the following section of this chapter. Other known secure tunnels are Secure Shell (SSH) which use the Diffie-Hellman key exchange and verifies data integrity with the use of message authentication code (MAC), and Internet Protocol Security (IPsec) that uses hash algorithms integrity and authenticity along with symmetric key algorithms for confidentiality (Prowse, 2015; Solterbeck, 2006). These terms are explained in detail in *2.4.2 Encryption*.

Encrypting data at rest and in motion along with having a suitable key-management solution is essential for protecting and ensuring the confidentiality and integrity of all data managed in organizations. If a breach were to occur, the organization will be investigated, in accordance with the GDPR, to check what safeguards were applied before the breach occurred. Having applied the proper

encryption and protection measures for the data might potentially reduce the sanctions placed on the organization (Tankard, 2017).

2.4.2 Cryptography

Cryptography is the study of implementing different techniques for protecting data and securing communication. Encryption is only a process of cryptography that aims at ensuring the confidentiality of the data, however cryptography as a whole is intended to cover several security aspects related to data protection (Saha, 2015). Other than ensuring the confidentiality of the data, the purpose of cryptography is to ensure that the following security aspects are fulfilled (Saha, 2015):

- **Authentication:** To ensure that the data received is sent from an authorized party.
- **Integrity:** To ensure that the data has not been manipulated, and that it has remained in its original state.
- **Non-repudiation:** To ensure that the parties involved in the data transmission should not be able to deny sending or receiving data.
- **Access-control:** Regulating access to data by authenticating the party requesting access.

These security aspects are met by the combination of several cryptography concepts such as encryption, message authentication and key-management. These concepts are described in the coming sections of this chapter.

2.4.3 Encryption

In the world of cryptography, encryption is defined as the process in which information is changed from a comprehensible form, known as plaintext, to an incomprehensible form known as a cipher text. The entire process of encrypting and decrypting data is done using a preset algorithm, also known as a cipher, with the help of a so-called key. This key is the fundamental part of the entire encryption process where it holds the blueprint to how the information is encrypted and how to decrypt it. The strength of the key is determined based on its size in bits, the bigger the key is the harder it is for unauthorized entities to decrypt the data (Prowse, 2015).

Keys are either private or public. A private key is kept secret and is only known to a specific entity or entities, whereas a public key is known and publicly distributed to all involved entities to exchange data over a secured connection. The use of private and public keys may differ depending on the type of encryption algorithm used. These algorithms are classified into two types, symmetric and asymmetric (Prowse, 2015).

2.4.3.1 Symmetric Encryption

Symmetric algorithms are known for using a single shared private key between the sender and the receiver. This key is often referred to as a secret key or symmetric key since the same key is used for both encrypting and decrypting data (Acosta, et al., 2016; Chandramouli, et al., 2014). The symmetric key algorithms are classified into two types:

- **Stream Cipher:** This type of symmetric algorithm is used to encrypt each binary digit in the data stream, one bit at a time (Prowse, 2015). The algorithm generates a pseudorandom random stream, known as a keystream, which is combined with the plaintext one bit at a time (Acosta, et al., 2016).
- **Block Cipher:** This type of symmetric algorithm encrypts the plain text by processing it into different fixed sized blocks where each block is made up of a group of data bits. All blocks are then individually encrypted using the same key data (Acosta, et al., 2016; Chandramouli, et al., 2014).

The National Institute of Standards and Technology (NIST)[8] have up until now validated and approved two symmetric encryption algorithms that can be implemented as security functions. The algorithms are, AES (Advanced Encryption Standard) and Three-key Triple-DES (Data Encryption Algorithm) (NIST, 2015).

2.4.3.2 Asymmetric Encryption

Asymmetric encryption, also referred to as public-key cryptography, generates a pair of non-identical keys where one is public and the other is private. The keys are however related mathematically where one key is used to encrypt the data and the other paired key is used to decrypt the data (Acosta, et al., 2016; Chandramouli, et al., 2014).

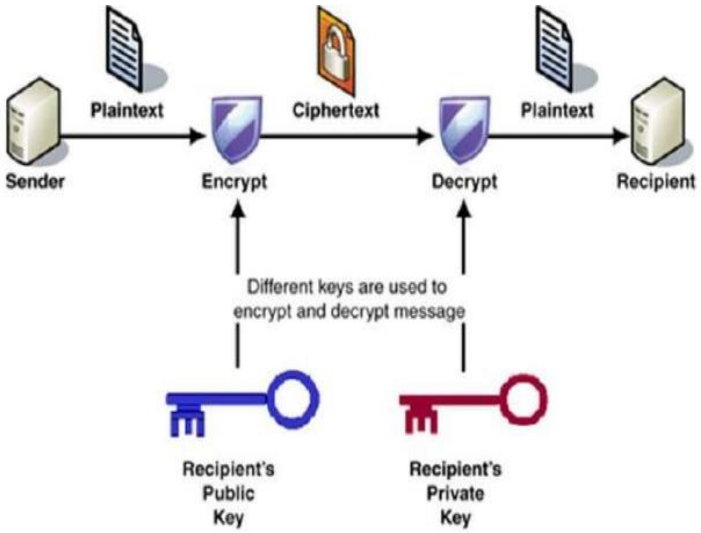

Figure 3 – Illustration of Public Key Cryptography (Tutorialspoint, 2017)

Figure 3 illustrates the basic public key cryptography process where the sender intends to transfer data to a recipient. The sender starts by encrypting the data with the recipient's public key and transfers the data to the recipient, the recipient then uses his private key to decrypt the data. In more complex public key cryptography designs, the sender wants the recipient to be assured that the data sent is from him. To achieve this, the sender signs the data using his private key and the recipient can check the signature using the sender's public key. This is referred to as a digital signature that ensures the integrity of the encrypted data and protects it from being manipulated by an unauthorized third party (Prowse, 2015; Stallings & Brown, 2012).

[8] The National Institute of Standards and Technology (NIST) is a non-regulatory federal agency of the United States Department of Commerce holding and is directed towards promoting and maintaining measurement standards (NIST, 2016) The focus on NIST standards in this thesis is mainly based on the fact that Tieto AB uses those standards. NIST has a long experience with HSMs and have a certification and validation program for HSMs and encryption algorithms as well as a testing standards. PCI-DSS, CIS and OWASP also refers to NIST standards. ISO has standards for HSMs (ISO 19790) and for key-management (ISO 11770) but have less adaptors as of yet (Pattinson, 2012).

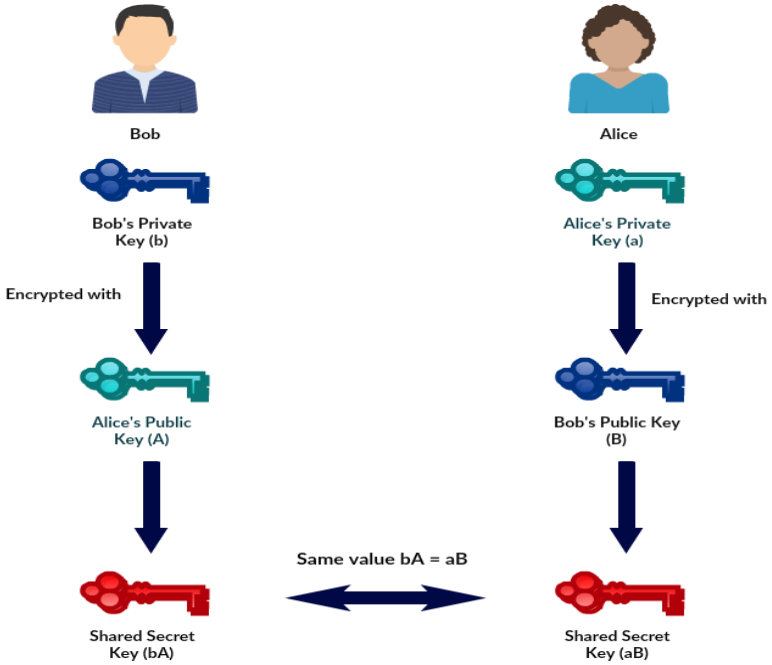

Figure 4 – Diffie -Hellman Key Exchange

Another common public key cryptography design is the implementation of the Diffie-Hellman key exchange process, which is intended for securing the key exchange process between the involved parties over a public network (Stallings & Brown, 2012). This process combines both asymmetric keys and symmetric keys, where each user involved in the exchange process generates a public/private key pair and distributes the public key to the involved parties which will be used to create a secret key shared between them (Prowse, 2015; Stallings & Brown, 2012). In order to clarify the concept as much as possible, *Figure 4* illustrates a simplified form of the Diffie-Hellman key exchange process. Both Bob and Alice have shared their respective public keys with each other, Bob encrypts his private key (b) with Alice's public key (A) to form the shared secret key (bA). Alice encrypts her private key (a) with Bob's public key (B) to form the shared secret key (aB). So, both Alice and Bob have now obtained a secret key with a value equal to the other (bA = aB), this key can now be used for encryption and decryption of the data transmitted between both parties (Prowse, 2015; Stallings & Brown, 2012).

Up until now the approved asymmetric key algorithms are, DSA (Digital Signature Algorithm), ECDSA (Elliptic Curve Digital Signature Algorithm) and RSA (Rivest, Shamir, Adleman) (NIST, 2015).

2.4.4 Data Authentication

As already established, encryption is used to maintain the confidentiality of data. However, data authentication, also referred to as message authentication, is used to maintain the integrity of the data whether it is at rest or in motion. To ensure the integrity of the data, it is important to have implemented mechanisms or functions that can verify that the data has not been tampered with, that the data is from an authentic source, and check the data's timestamp to validate that the data has not been excessively delayed beyond what is considered to be the normal data transmission time for the network. This can be achieved using the so-called hash functions (Stallings & Brown, 2012).

A hash is known to be a summary of data in a string or numerical form, and is used for protecting the integrity of data at rest and in motion. A hash is generated through the implementation of a hash function, which is a procedure that takes an arbitrary block from the data and converts it into a fixed-sized hash value (Prowse, 2015). When the data is in transit, a hash value is generated at the source, after the data arrives at the destination, an algorithm is applied to the hash value which generates a

second hash value, the latter is compared to the first value to determine that the data has not been tampered with, thus verifying the data's integrity (Prowse, 2015; Stallings & Brown, 2012). Hash functions are classified into two types, un-keyed hash function and keyed hash function (Ariwibowo & Windarta, 2016; Tiwari & Asawa, 2012).

- **Un-keyed Hash Functions:** hash functions that require one parameter, which is the message, to generate a hash. This can also be referred to as Manipulation Detection Code (MDC) (Tiwari & Asawa, 2012), message digest or classified more generally as one-way hash functions (AlAhmad & Alshaikhli, 2013; Stallings & Brown, 2012). The term "one-way" is used to describe the hash as irreversible, meaning that one should not be able to recreate the hashed message (Prowse, 2015). These hash functions are mostly used in creating digital signatures for identifying the sender and authenticating the data. The hash function can be encrypted using symmetric encryption, in this case authenticity is guaranteed if one assumes that the symmetric key is only known to the sender and receiver. It can also be encrypted using public key cryptography, as previously described in *2.4.3.2 Asymmetric Encryption,* where the sender encrypts the hash with his private key creating a digital signature. After the data reaches the receiver, a hash value is calculated for the data, the receiver decrypts the digital signature using the sender's public key, and then the calculated hash value is compared alongside the decrypted hash value. If both hash values match, it should be clear that the data is sent from the intended source assuring the authenticity of the data source, and the fact that the data cannot be altered without having access to the private key of the sender assures the integrity of the data (Stallings & Brown, 2012).
- **Keyed Hash Functions:** hash functions that require two parameters, which are the message and a key, to generate the hash. These types of hash functions are used to construct variations of the so-called message authentication code (MAC), which is used for both ensuring the integrity of the data along with authenticating the source of the data (Tiwari & Asawa, 2012). One of the most widely used MAC type is HMAC (Hash-based Message Authentication Code), which involves using a secret key in conjunction with a hash function to produce a hash value (or MAC) before transmitting the data (Prowse, 2015). The receiver performs the same process on the received data (secret key + hash function) to obtain a new MAC, the calculated MAC is then compared to the received one, if they match the receiver is assured that the message has not been altered. If the data was to be manipulated during transmission, the received MAC would have differed from the receivers calculated MAC, since the unauthorized party in this case is unable to modify the MAC appended to the data to concur with the modifications made to the data without knowing the secret key. This assures both the data integrity and the authenticity of the sender (Stallings & Brown, 2012).

Hash functions are also used alongside salt values[9] to safeguard the stored passwords by hashing them. The password and salt value are used in conjunction with a hash function to produce a fixed-length hash code, the hash value is then stored alongside a plaintext copy of the salt value in the column or password file for the corresponding user (Stallings & Brown, 2012).

Hashing can also be used in intrusion detection systems (IDS). The IDS create a hash value or a checksum for the data that it is configured to monitor, this hash value is based on the different attributes of the stored data such as size, modification date etc. This hash value is then periodically

[9] A random value used as an additional parameter to a hash function to produce a hashed password (Prowse, 2015)

compared to the stored data to determine if any modification has taken place (Stallings & Brown, 2012).

The following hash functions and message authentication codes are approved at the time of writing this thesis, SHA-2, SHA-3, HMAC[10], CMAC (Cipher-based message Authentication Code), and GMAC (Galois Message Authentication Code). SHA-1 is known to have security issues but is still acceptable to be used in all hash function applications, except in generating digital signatures (NIST, 2015).

2.4.5 Key Management

To make encrypted data protected even if lost or accessed by unauthorized entities the keys to the cipher must be kept safe (Prowse, 2015). But what happens when data is needed elsewhere or when it is shared with other authorized systems? The data can either be decrypted prior to transfer so that the data recipient receives the data in clear text, or the keys to decrypt the data can be assigned to the authorized systems so that the data can stay in its protected encrypted format when moving to the recipient. But how are the keys transmitted in a safe and protected way, and how can one make sure that only the authorized systems get them? To do this in a safe way and mitigate as many risks as possible a key-management system, or KMS, can be used (Prowse, 2015).

First, let's look at the encryption algorithms discussed in chapter 2.4.1 and how they deal with their keys. The symmetric encryption algorithms such as AES use the same key to encrypt and decrypt the data, which means that if the data is transmitted somewhere else to be decrypted and used, the key must exist in multiple places as well. This mean that the key becomes more vulnerable to attacks as it is transmitted and stored in more than one location. If the key is stolen it can be used to decrypt all the data that was encrypted using that key. The asymmetric encryption algorithms however use different keys for encryption and decryption. One key is always considered private and should never exist in more than one place at a time, but the other key(s) are called public keys and can be distributed openly to everyone as necessary, (hence the name public key infrastructure or PKI).

A KMS exist to help with the implementation and use of cryptographic keys in a secure manner and deals with the policies, documentation and practices for said keys (NIST, 2016).

NIST offers a comprehensive documentation of key management in its special publications 800-57 series which consist of three parts as described in Table 1 below.

Publication	Contents
NIST SP 800-57 Part 1	General key management guidance. Intended for system developers and system administrators.
NIST SP 800-57 Part 2	Focuses on organizational key management infrastructure and key management policies, practices and plans. Intended for system and/or application owners.
NIST SP 800-57 Part 3	Focused on key management issues related to the available cryptographic methods. Intended for system installers, system administrators and end users.

Table 1 - NIST Special Publications on Key Management

[10] Can be used only with a key length greater than or equal to 112 bits (NIST, 2015)

2.4.5.1 Key Management Life Cycle

One of the aspects of a KMS is to deal with the entire life cycle of a cryptographic key. A good and useful key must be generated in a secure and trusted environment, registered to a cipher, distributed, implemented, used, suspended after its planned lifetime, and finally destroyed or stored securely for future use (NIST, 2016). In addition to these principles, there needs to be procedures in place on rotating keys, how to deal with potentially compromised keys and, if necessary, to recover lost or damaged keys (OWASP, 2016).

Both OWASP[11] and NIST describes a key lifecycle with four states.

Key lifecycle states according to:		State description
NIST	**OWASP**	
Active	Current	The key is active and in service both encrypting and decrypting data
Deactivated	Retired	The key is no longer used for encrypting data, just for decrypting data previously encrypted by it.
Compromised	Expired	Key is compromised and is only used for decryption of data previously encrypted by it so that it can be re-encrypted using a new and active key.
Destroyed	Deleted	The key no longer exists anywhere. Any data still encrypted by the key is considered lost.

Table 2 - Key lifecycle states

For key management, NIST adds four phases and two states which creates a model as can be seen in Figure 5 below.

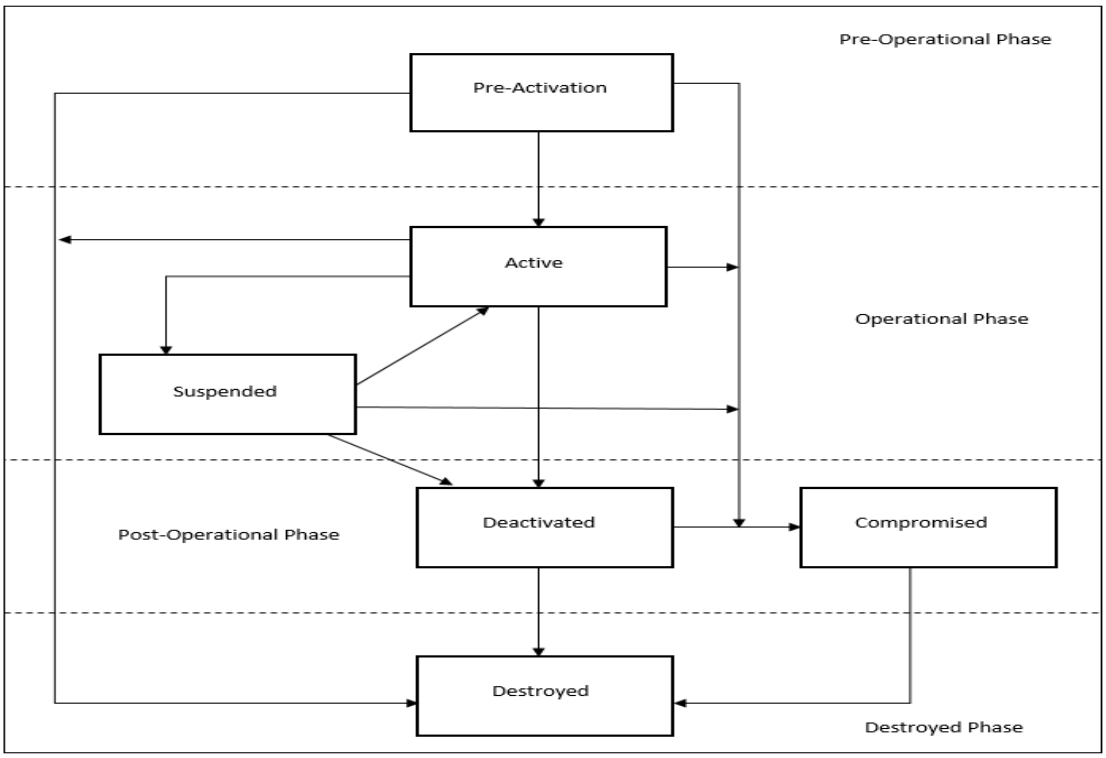

Figure 5 – Key Management States and Phases, modeled after NIST SP 800-57

[11] OWASP, or Open Web Application Security Project is a not-for-profit charitable organization focused on improving security of software. More information can be found at http://www.owasp.org

As Figure 5 shows, keys are never able to go back to a previous phase if it has transitioned to a new phase. The figure is read as follows from the top. A key is in the pre-operational phase when it has been created but has not been registered to a user, system, application and so forth. If the key never becomes registered it can move directly to the destroyed phase. If the key prior to registration becomes compromised it moves to the post-operational phase. When a key has been registered, and becomes an active key it moves to the operational phase at the time of its activation. If an active key becomes compromised it moves to the post-operational phase. If the key is no longer needed or if it has reached the end of its planned life time it moves to the post-operational phase. A deactivated key may stay in the post-operational phase for as long as the key is needed for decryption. When the key is no longer needed for any operations it is moved to the destroyed phase (NIST, 2016).

For a KMS to be valid it needs to have all of the above key operations and phases managed. If any keys are un-accounted for or if their whereabouts are unclear, the KMS becomes invalid and the protected data is in serious risk of compromise. A valid KMS is also very useful to maintain traceability for auditing, as events usually are can be recorded as logs whenever a key pass through a key operation and when the key transitions to a new phase (NIST, 2016).

Below is a useful explanation of some KMS-terms.

Key generation	A new key is generated using a random number generation process to produce an unpredictable key. NIST SP 800-133 “Recommendation for Cryptographic Key Generation” states that all key generation shall be performed within a FIPS 140-2 compliant HSM (NIST, 2012).
Key registration	The key becomes associated with a user, system, application or policy. It can be registered as a signing key, encryption or decryption key, etc.
Key storage	The key is kept safe in storage within the HSM, whether it is in use or not. By storing it in a HSM the keys are kept separate from the data it is meant to protect. If the keys are stored outside of the HSM they are usually kept encrypted with another key that is stored within the HSM, these keys are usually called key encryption keys (KEK).
Key distribution	A key must have a way to be securely transmitted from the safe storage to the application or physical device that needs to use it. This can be done in many ways, one is to set up secure link between the key storage (HSM) and the application that needs the key, but often this isn’t enough because the KMS should also know that the application requesting the key is trustworthy and that their identity can be validated, and vice versa, that the HSM is trustworthy and identifiable.
Key use	A key in active use. A key should only be used for one purpose, such as encryption or authentication. Using one key to perform many tasks may seriously jeopardize the security of the system.
Key rotation	All keys should have a limited lifetime as the longer it exists and the more data that is attached to it, the more important it becomes. This raises its value to intruders and hackers. To mitigate this risk all keys should be rotated or refreshed periodically

Key backup	If a key is lost all the data encrypted by that key is lost as well. Therefore, the should exist a key backup procedure to take backups of the primary keys to another safe storage, perhaps located offsite to limit effects of for example fires or other events that may trigger a HSM to clear its key storage.
Key recovery	When a key is lost there needs to be a procedure in place to find and restore that key from the key backup storage, and to recover that key safely so that it is not exposed while in transit or while waiting for implementation in an encryption system. It is also vital that the recovery process clearly states who within the organization that can order such a recovery and how that recovery is carried out. The auditing trace must be defined as well to ensure that traceability is achieved.
Key revocation	When a key is compromised or even just suspected of compromise it must be revoked immediately. This requires a clear and well tested policy and procedure so that the revocation is communicated and so that all instances that use that key is informed and provided with a new key.
Key suspension	A key that is at the end of its operational life cannot usually be destroyed if it has been used to encrypt a lot of data, instead it needs to be stored or put in key suspension so that it can be accessed to decrypt the data it belongs to. Of course, all data can be re-keyed following a key rotation, but that may not always be feasible or economic. The storage of suspended keys must be just as safe as the storage for active keys.
Key destruction	When there is no use for a key anymore it should be destroyed. All instances of the key must be destroyed, such as backups, and there needs to be traceability for future audits that the destruction actually took place. It is important to understand that when a key is destroyed, all data encrypted with that key is lost as well.

Table 3 – KMS terminology

2.5 Hardware Security Module (HSM)

Hardware Security Modules or HSMs are cryptographic modules based on a combination of hardware and software to implement cryptographic functions in an IT-environment (NIST, 2002). HSMs come in various forms and formats, ranging from smartcards, PCI plugin cards, and the standalone network based HSM. This study focuses on the standalone HSM as they can be accessed and utilized by multiple servers and clients, regardless of platform, and due to their processing power, which is required in applications with requirements on the performance aspect of the HSM.

The purpose of the HSM is to safely generate and store cryptographic keys and to act as the trusted crypto anchor in an encrypted system. The basic use of an HSM is to let it perform all cryptographic processing on the protected data. As an example, it can store all encryption keys and perform all encryption and decryption of data on the request from client systems, as the following example shows:

1. A client with access to a server with symmetrically encrypted data fetches the desired data in its encrypted form.
2. The client transfers the data to the HSM for decryption.

3. The HSM checks the access rights of the client and decrypts the data if the client has the required rights.
4. The HSM transmits the decrypted data back to the client using a secure encrypted transmission method, usually through the use of VPN, TLS or IPsec.
5. The client decrypts the information from the HSM and can then use it.

This way ensures that all data encryption keys are always kept safe within the HSM.

The HSM functionality is contained within a physical perimeter called a cryptographic boundary, usually it takes the form of a box in which the entire cryptographic process takes place including the key-management and the actual encryption and decryption phases. The HSMs are rated to a specific security level based on requirements from the NIST standard FIPS 140-2.

2.5.1 Security Levels

There are currently[12] four different levels that a HSM can be rated to:

Security Level (FIPS 140-2)	Description
1	At least one approved[13] algorithm or security function shall be used. No specific physical security mechanisms are required. An example of a level 1 HSM is an unprotected PCI encryption board in a PC.
2	Adds the requirement of tamper-evidence to the HSM. Tamper-evidence are for example seals and coatings on the HSM that must be broken to gain access to the physical unit. It can also be represented by pick-resistant locks that protects the unit. In addition to the tamper-evidence, a level 2 HSM is required to utilize role-based authentication. Examples of level 2 HSMs are PCI boards with protective covers.
3	Adds strong enclosures, tamper-detection and response to the physical security. Opening the HSMs enclosure shall be detected and the HSM responds by zeroing all critical security parameters contained within the HSM. Level 3 HSMs shall also require identity-based authentication. These HSMs are typically rack mounted units kept in locked environments.
4	The HSMs at level 4 have a very high probability to detect attempts to breach the physical security and responds by zeroing the critical security parameters contained within the HSM. At this level, the HSMs are also monitoring the environment around the enclosure for signs of tampering, such as changes in temperature or voltage levels. HSMs at level 4 are usually used in environments where they cannot be physically protected and therefore must be able to react to any intrusion attempt on its own volition.

Table 4 – FIPS 140-2 Security levels for HSM

2.5.2 Cryptographic Boundary

All the FIPS 140-2 levels require that the cryptographic functionality be confined within a cryptographic boundary that separates this functionality from other functions of the equipment, for example the communication parts of the HSM. The different security levels also require different levels of protection for the cryptographic boundary, ranging from just being protected by the

[12] The current FIPS 140-2 standard is from 2001 and is currently under revision. A draft of the new standard, FIPS 140-3 exists but has not been released as of spring 2017.

[13] NIST approved algorithms are published in NIST Special Publication 800-131A Revision 1

environment it is installed in (contained within a protected industrial environment) to a self-evaluating system built in to the HSM that reacts to any attempts to tamper with the cryptographic functions.

To facilitate the protection of the cryptographic functions the cryptographic boundary is usually contained within an enclosure. This enclosure may have access doors to allow access to the hardware inside, and from FIPS 140-2 level 2 also have tamper-evidence equipped for technicians and other personnel to see if the HSMs cryptographic functionality may have been interfered with. These tamper-evidence usually comes in various forms of seals that must be broken in order to gain access to the inside of the enclosure (NIST, 2002). HSMs certified according to FIPS 140-2 level 3 must have tamper-resistant functionality built in, meaning that opening the enclosure by using the access doors results in that the tamper-resistant system detects it and responds by zeroing all plaintext cryptographic keys and other critical security parameters (CSPs) (NIST, 2002). HSMs certified at FIPS 140-2 level 4 shall have a tamper-detection system that ensures high probability that any attempt to access the enclosure from any direction shall be detected and that all plaintext CSPs are immediately zeroed. This can be done by using sensors that measure the ambient temperature, voltage levels, pressure and strain on the enclosure and so on (NIST, 2002).

2.5.3 Random Number Generator

Random number generators (RNG) are used in the generation of keys in both public-key encryption algorithms and symmetric encryption algorithms. They are also used for generating session keys, digital envelopes and in the handshaking process during key distribution. The generated sequence of random numbers should distinctively be random and unpredictable (Stallings & Brown, 2012).

One main concern when it comes to randomness is validating if the sequence of numbers is truly random. There are two criteria used for validating the randomness (Stallings & Brown, 2012):

- **Uniform distribution:** Meaning that the frequency of occurrence of the numbers in the sequence should be relatively equal or uniform.
- **Independence:** No number in the sequence can be derived from the other numbers. However, there is no concrete way to determine if independence between the values truly exist. Usually several tests are done where several number sequences are generated, these sequences are examined to see if the values are independent from each other, these tests are repeated several times until one is confident enough to determine that independence exist.

For an RNG to be approved for use in a cryptographic module it must pass the conditional tests, for an RNG, specified in FIPS 140-2. When an RNG generates an *n-bit/n-bit block*, the first *n-bit/n-bit block* generated will be saved and compared to the second *n*-bit/*n*-bit block generated. Each successive *n-bit/n-bit block* is compared to the one generated before it, the RNG would fail if any two-compared *n-bits/n-bit blocks* are equal (NIST, 2002).

2.5.4 HSM Interface

Interfacing with a HSM can be done using many different methods depending on the type of HSM. A non-networked HSM can for example be interfaced with using serial communication, Ethernet crossover cable or physically using the interface panel on the HSM (if such exist). Today however, many HSMs are indeed network based as they need to serve multiple clients, applications and services and are therefore sometimes exposed to a network of devices.

There is however some separation between the cryptographic functionality within the HSM and the physical ports for data input and output, however the actual physical ports do not have to be separated, data input to the HSM and data output from the HSM are however separated logically (NIST, 2002). To reach FIPS 140-2 level 3 and 4 a HSM must disconnect the output data logically while

performing key-generation, manual key input and while performing key zeroing to mitigate the risk of key data inadvertently leaking out of the HSM (NIST, 2002).

Additionally, the HSMs operating on FIPS 140-2 level 3 and 4 requires a division of command structure to let key data out of the HSM or allow changes to be made to the HSM. This usually means that an *m-of-n*[14] structure is employed to mitigate risks of single user access.

2.5.4.1 Remote access and administration

Administration of a networked HSM can be done remotely on many current HSMs. Usually this is done by using encrypted communication and passwords. There are different solutions available on the market, for example using smart cards or other tokens together with a PIN or password to gain administrative access. Remote administrative access to a HSM usually does not entail root- or super user access. Instead the different cryptographic parts of the HMS require specialized *m-of-n* access by the owner of the specific cryptographic function, again, usually employing an *m-of-n* structure. The actual setup of the HSMs cryptographic functionality differs between the different models of HSMs. Gemalto SafeNet Luna for example employs an *m-of-n* structure using tokens of different colors together with a PIN input device connected to the HSM via the network or directly to a USB port on the HSM (SafeNet Gemalto, 2010). Thales nShield products uses different sets of smartcards, based on the access rights of the user, administrator (no access to key data) or operator (access to key data), together with PIN to gain access and to manipulate the HSM (Thales, 2015).

2.5.4.2 Application and programmatic interface

Most HSMs today conforms to the Public-Key Cryptography Standard no. 11 (PKCS#11) for cryptographic operations. PKCS#11 is an open source standard developed by RSA Laboratories for the creation and use of cryptographic tokens (RSA Security Inc., 2004). PKCS#11 is platform independent and is used in many crypto APIs such as Mozilla Firefox, OpenVPN, OpenSSH and OpenSSL (Wikipedia Contributors, 2017).

There is other proprietary software also used to access the functionality of HSMs, one such example is Microsoft Crypto API (MS-CAPI) used by Microsoft Windows.

2.5.5 KMS and HSM

A HSM is an inherently good device to let handle all keys in a KSM since it usually employs strong access controls and is kept in a safe location.

A HSM can be used as a trust anchor in a certificate structure where it can oversee and control all access rights to the encrypted data.

Below is an example of how a HSM can be used as a trust anchor in an encrypted environment:

1. A client needs to access data on encrypted server.
2. The client lacks the required key to be granted access to the encrypted data on the server.
3. The client sends its authorization signature and the required key-identifier to a key server and registers its own public key with the key server.
4. The key server checks the validity of the client's authorization signature and checks what access rights the signature is granted.

[14] An *m-of-n* structure is when a key or access right is split into pieces (*n*) and is divided between multiple users (*m*). It is then decided how many *m* is required to present their piece of *n* to gain access to the system. Sometimes referred to as *k-of-n.*

5. If the client has the required access rights, the key server sends the clients public key and the requested key-identifier to the HSM attached to the key-server. (The HSM is only accessible by the key-server).
6. The HSM wraps the client's public key together with the requested key and returns it to the key-server.
7. The key-server signs the wrapped keys and send them to the client.
8. The client validates the signature and stores the wrapped keys.
9. The client uses its private key to unwrap the desired key.
10. The client now has the desired key and can access the data required on the encrypted server.

Alternatively, the client can ask for the data from the server, the server checks the client's access rights with the HSM and if cleared sends the encrypted data to the HSM that decrypts it and sends the data in clear text via encrypted tunnel to the client, or the server can send the data encrypted to the client that then needs to send it to the HSM for decryption before receiving it in clear text via a secure tunnel from the HSM. The idea with both these alternative approaches is to, besides the HSM's key management functions also utilize the powerful cryptographic processing power of the HSM instead of putting that burden on the server or the client's hardware.

2.5.6 Certification and Validation Program

NIST provides a Cryptographic Module Validation Program (CVMP) that lists all HSMs that have passed the FIPS 140-1 and 2 requirements. The list can be found at the NIST Computer Security Division Computer Security Resource Center online (NIST, 2017).

There are other validation and certifications programs as well, such as the *Common Criteria Evaluation of Information Technology (IT) products* and the *ISO 19790 Information technology – Security techniques – Security requirements for cryptographic modules* and *ISO 24759 Information technology – Security techniques – Test requirements for cryptographic modules*.

2.6 Knowledge Gap

No scientifically reviewed literature regarding the use of encryption, by using HSMs or otherwise, for compliance to GDPR was found during the literature study for the theoretical framework of this thesis. Some whitepapers published by HSM vendors and IT-research organizations on the subject were however found.

Thales E-Security[15] has released a whitepaper on data encryption and key management addressing compliance with the GDPR called "Addressing Key Provisions of the General Data Protection Regulation (GDPR)" (Thales E-Security, 2017). In the whitepaper, they identify article 32 and 34 of the GDPR as the main provisions that can be addressed by encryption technologies and using their products (European Union, 2016). The paper claims that the GDPR calls for a layered approach including access control, encryption, and monitoring, and goes on to suggest how Thales products can help address those layers. The suggestion includes the use of Thales Vormetric Data Security Manager (DSM) that is available in different variants, including one that is FIPS 140-2 Level 3 certified and that

[15] Thales E-Security is part of the Thales Group, a French multinational company that operates within the aerospace, defense, transport, and security industry (Thales Group, 2016).

contains a HSM (Model V6100) (Thales E-Security, 2017). The DSM in this suggestion is used to limit the access to the sensitive data by utilizing access control functions and to store and manage all cryptographic keys while the actual encryption is performed using Thales proprietary "Vormetric" encryption solutions (Thales E-Security, 2016).

Bloor Research International Ltd.[16] has released a whitepaper written by Fran Howarth called "For the EU's new data protection regulation, encryption should be the default and should be seen as a strategic part of the entire security system" (Howarth, 2016). The whitepaper addresses the need for data protection to prepare for the new data regulations that GDPR entails as well as to prepare for a world where data breaches and attacks are becoming more and more common (Howarth, 2016). Part of the solution is according to the whitepaper to protect data using encryption as a default measure. It is also stated in the end of the whitepaper that encryption alone is not enough, and that it needs access controls, auditing, and visibility of who is accessing the sensitive data.

Gemalto N.V.[17] together with DQM GRC Ltd.[18] has released a whitepaper called "Essential Security Technologies for GDPR Compliance" (Gemalto, n.d.). In it they focus on several different actions to implement to become compliant to GDPR. The whitepaper doesn't mention specific technologies but instead focuses on security areas that needs attention. Data discovery and detection is the first objective to address, it addresses article 30 of the GDPR and describes the need for identifying where the sensitive data is located in the system (European Union, 2016; Gemalto, n.d.). Some of the other objectives to address are: access control and restriction, pseudonymisation and encryption, and backup and recovery (Gemalto, n.d.).

Although the whitepapers from these two vendors and the one IT research company are all focused on the GDPR and encryption as a means to address the compliance requirements it is only Thales that explicitly states that their products can address specific articles of the GDPR (articles 32 and 34 (Thales E-Security, 2017)). This could lead to the conclusion that there is a lack of research on the technology specifically developed for addressing data protection issues and specifically on data protection requirements of the GDPR and there is therefore a knowledge gap about HSMs and their potential role in achieving compliance with GDPR.

[16] Bloor is an independent research and analyst company based in the UK. They focus in IT related research and consulting (Bloor, 2017).
[17] Gemalto is an international digital security company based in the Netherlands. They provide technologies and services for identify authentication and data protection (Gemalto, 2017).
[18] DQM GRC are based in the U.K. and specializes in data governance, compliance advisory and technologies benchmarking i.a. (DQM GRC, 2016).

3 Methodology

This study aims to answer two questions:

- How can the use of HSM aid in achieving compliance with GDPR?
- What GDPR requirements would be left un-addressed by using such an approach?

For the first question, we will mainly utilize the literature study as a tool to analyze the GDPR and the security capabilities of a HSM. The data collected can then be mapped against each other to find the parts of the GDPR that can be addressed by using HSMs and the parts that cannot. For the second question, we will use a Delphi-study to find what aspects security professionals identify as the most important when securing personal data. The information collected in the Delphi-study can then be used together with the identified un-addressed GDPR requirements from research question one, to discuss the validity of such an approach.

3.1 Literature study

The first place to look for information upon starting work with this thesis was the EU database for legal documents. The regulation 2016/679 "…on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)" (European Union, 2016) has become the foundation upon which this thesis builds its purpose and research. This document, that is called GDPR for this thesis includes several keywords that where later used when searching for additional reference material. The keywords and phrases found were: Privacy, Personal data, Data protection, Privacy enhancing tools, Encryption, Pseudonymisation, and Data breach.

Since the thesis is done with an organization's interest in mind that wanted to explore the use of HSMs as a compliance method for the GDPR, the term Hardware Security Module was added to the list of keywords and phrases. As the identified list of keywords and phrases became more and more detailed, it was clear that the literature study became concept-centric and therefore all concurrent data was analyzed primarily based on their concepts (Webster & Watson , 2002).

The search engines used for the data gathering was primarily the PRIMO database provided by Luleå University of Technology. Google Scholar was also extensively used. The search terms used to provide the study with the initial batch of literature was:

Search term	No. of hits in PRIMO
Personal data compliance	164131
Hardware Security Modules AND compliance	3432
GDPR AND Encryption AND Privacy	2498
Key-management AND HSM	1620
Data protection AND HSM	1154
GDPR and data protection	331
Compliance with GDPR	189
Hardware Cryptography AND GDPR	3
Hardware Security Modules AND GDPR	2

Table 5 – Search terms used for gathering literature

All searches in PRIMO were limited to peer-reviewed and articles published within the last five years. Some searches resulted in a massive number of hits, but since a few documents were chosen from the first couple of pages of hits the search was not narrowed down further. All documents found were assessed by their title, and if that indicated an interesting topic, the next stage was to read the

abstract. If the document was deemed relevant for the study, it was saved and its data was entered into a spread-sheet where the authors could mark it as useful after reading through it briefly. These documents were then analyzed further by a more thorough read through. This analysis looked at the references of the documents to be able to conduct backwards searches to gain access to the original sources. The documents where then assigned to specific groups of topics to make them easier to handle and find. These topics where: Anonymization, Compliance, Cryptography, Encryption, GDPR, Hardware Encryption, HSM, Key Management, Personal data, Privacy, Pseudonymisation, Risk Analysis, Software Encryption, and Encryption Standard which also represent the concepts of the literature study.

In addition to the data gathered from the university databases and Google Scholar there was an internal file share made available to the authors by Tieto AB for whom the study was performed. This file share contained about 70 documents, whitepapers, surveys and other documents that were also added to one of the literature spread-sheets tabs and organized according to their contents.

Different international- and national standards where used as well.

3.2 Empirical study

This research will have a qualitative approach, meaning that it does not focus on quantifying the studied objects (Landrum & Garza, 2015) but rather focus on the different characteristics of these objects from a pluralistic viewpoint (Esaiasson, et al., 2012)

The purpose of this thesis is to determine if the use of HSM is a feasible way to achieve at least some level of compliance to the GDPR, and to determine what, if any, residual GDPR requirements with regards to said compliance that needs to be addressed through technical or other means. This indicates that the study is mainly a qualitative one, since the requirements of the GDPR are so loosely defined in their context as a legal documents and not pure requirement specifications. The input for deciding what measures, technical and otherwise, that constitutes as compliance to the different parts of the GDPR are mostly going to be based on opinions from people involved with data security and data processing on both the technical level and the legal level. There is no precedence on the actual accountability of compliance to the GDPR as of yet, as the regulation has not come into effect and therefore there are no rulings from the different SA's on what compliance actually entails on a detailed and specified level. The general lack of hard data on how the text of the GDPR should be interpreted leads us to focus on a qualitative approach to try to determine how the general consensus, if any exist, among the interested parties line up when discussing compliance to the regulation.

The specific data gathering for this study, apart from the previously discussed literature review, will consist of a Delphi-study utilizing a panel of experts as the main empirical approach. More on the selection of the professionals and experts in the corresponding data gathering chapters below.

3.2.1 Delphi Study

The Delphi research method is known to be an iterative tool for gathering empirical data with the help of an expert panel or panels. The experts are required to participate and answer several rounds of questionnaires where each consecutive questionnaire is designed based on the results from the previous rounds. The data gathering process ends when the researchers feel that the necessary information has been gathered, consensus between the participants have been reached, and the research question has been answered (Skulmoski, et al., 2007).

The reason for using the Delphi method in this thesis is due to the lack of knowledge regarding the presented problem area and the difficulty to determine the impact of GDPR before it has been implemented. The Delphi method would fit this thesis since it is used in areas where there is a lack of

knowledge in a certain problem (Skulmoski, et al., 2007) or in researching what does not yet exist as done in forecast studies (Halal, et al., 1998; Okoli & Pawlowski, 2004; Whitman & Mattord, 2014).

The Delphi method is preferred in this thesis over other data gathering methods, such as interviews and focus groups, due to traveling and communication cost being kept to a minimum since all communication will be e-mail based. But also, due to the four key features presented by Rowe and Wright (1999):

- **Anonymity**: providing anonymity to the participants, giving them a platform to express their opinions freely without feeling pressured to adapt or adjust their opinions based on the other participants.
- **Iteration**: Delphi also aims at involving the participants as much as possible throughout the empirical data gathering process by allowing the participants to clarify their opinions and views after progressing from round to round.
- **Controlled feedback**: Sharing a summary of the responses to the panel members giving them an additional opportunity to clarify their opinions or views if they deem it necessary.
- **Statistical aggregation of obtained responses:** Basically, to quantitatively analyze and interpret the obtained data.

Since this research has a qualitative approach, the last key feature might not be suitable with our research. However Skulmoski, Hartman & Krahn (2007) argues that one can modify these features to adapt to one's own research, and proposed the term "Classical Delphi" to describe the type of Delphi method with characteristics that would comply with the features summarized by Rowe and Wright (1999).

3.2.1.1 Expertise Criteria and number of participants

The Delphi-study participants should be selected based on their expertise with the issue studied in the thesis, and this expertise can be broken down into four requirements (Skulmoski, et al., 2007):

1. Knowledge and experience with the issue
2. Capacity and willingness to participate
3. Sufficient time to participate
4. Effective communication skills

Based on these requirements we decided to create a profile for the experts needed for this specific study. We decided to put emphasis on the first requirement specifically by focusing on project leaders, security architects, security consultants or other security professionals with experience from at least one major project that included privacy or data protection. They also need to be familiar with GDPR or other similar privacy/data protection regulations.

The reason we find that familiarity with GDPR, or similar regulations, is sufficient for the Delphi study is because we consider it to be difficult to label an individual as an "expert" in GDPR. Based on the literature and discussions regarding GDPR, many seem to agree that different parts and concepts in the regulation are still vague or ambiguous when addressing the compliance requirements and needs further definition (Gilbert, 2016; Spindler & Schmechel, 2016). At the time of conducting this research, there is yet no official certificate or type of testimonial that would grant an individual the term "expert in GDPR" and we cannot depend on an individual's personal experience with GDPR either, since the regulation is yet to be implemented and there is yet no documented case regarding breaches to the regulation, for an individual, to gain experience from. Therefore, it would compromise the validity of the Delphi study if a participant is labeled as a GDPR expert. However, data protection is a main concept in GDPR and it seems more plausible to identify experts when it comes to the concept of data

protection, whether it is on a technical level or managerial. Certifications such as *Certified Information Systems Security Professional (CISSP)* exists to attest to an individual's expertise within the information security field. Also, as Okoli & Pawlowski (2004) describes, a certified expert can through his/her contacts or network gather other experts, in our case within data protection, which are not necessarily certified but their expertise is validated by years of working experience along with their achievements and merits within the field, in our case within the field of information security (Okoli & Pawlowski, 2004). These data protection experts, both certified and non-certified, will help in identifying critical aspects related to protection of personal data, we can then map these critical aspects to the information about data protection in the GDPR to then be able to answer the research questions.

The experts for the Delphi panel will be gathered with the help of our certified CISSP supervisor at Tieto. Through his contact network at the organization, he will gather the experts which best fit the profiles and requirements previously stated.

For the 2nd to 4th requirement we simply specified that the expert also must be willing, and have the time, to participate in the panel as well as have the communication means to be effective in receiving information and respond in an efficient manner (an internet connection and access to e-mail).

Based on the criteria for experts we aim to get a sample size of at least 10 panel participants. This sample size should be seen in relation to the total number of experts that fit the requirements mentioned earlier within the organization where the study is performed. The group selected will also be considered as a homogenous group who share similar knowledge and expertise, and thereby the sample size hopefully can produce sufficient results (Skulmoski, et al., 2007).

3.2.1.2 Delphi process description

Once the participants of the study have been chosen and informed about the background of the study and the part they are to play in its development it is time to start the actual Delphi-study. The first round of the study is meant to be a kind of brainstorming session to catch as much data as possible for the next step(s) in the study. The question for the first round is based on the research questions for the study but was generalized in order to cast a wide net and catch as much as possible from the panel participants as mentioned earlier.

Figure 6 below shows the process used to conduct the Delphi study for this thesis.

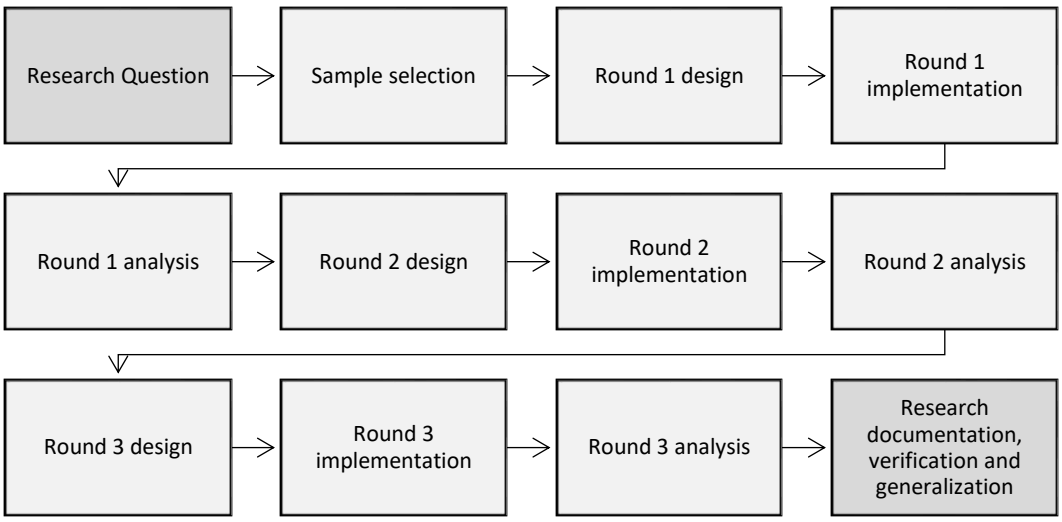

Figure 6 – Three round Delphi study process, based on concept from (Skulmoski, et al., 2007)

Delphi round 1:

The first round of the Delphi study will be done as a brainstorming session (Okoli & Pawlowski, 2004; Skulmoski, et al., 2007). All panel participants are asked to list aspects that the panel participant deems important for the topic of the study, based on their own expertise and experience. The topic of the study will be stated as a question and the panel participants are asked to list as many issues as they can and to briefly explain their context and possible consequence.

The question that defines the topic of the study is:

Identify the most critical aspects[19] of securing an information system with regards to personal data protection? (Personal data is defined by GDPR as any information relating to an identified or identifiable natural person). Name at least 10 aspects and describe the corresponding issues briefly with context and problem.

This question is meant to give the researchers information on what aspects the experts associate with personal data protection and to create a foundation based on the problems that the experts have encountered when implementing data protection and privacy in real world applications and projects.

At this stage of the study there is no real connection to the GDPR specifically, that is a choice we have made because of the lack of known specific GDPR knowledge among security professionals that we have found during the initial research stages of the thesis. We have instead chosen to focus on a more open and wide question about privacy, personal data and data protection since those are all among the key concepts we have identified for the GDPR during the literature study, see chapter 3.1. By gathering data about these concepts in the Delphi study, we are hoping to gain insight into the problems and issues that IT projects usually encounter when dealing with such concepts and thereby gain information for the RQs of this thesis.

The results from the first round is expected to contain upwards of 100 identified aspects, however many are probably going to be similar or even identical, and just described and worded differently. All the identified aspects from responses to round 1 are analyzed mainly by noting how many of the panel participants have suggested each aspect (Okoli & Pawlowski, 2004). Then an initial narrowed down and consolidated list, based on the similarities within the different responses and where identical answers are removed, is created and aspects are categorized to create a basis for a more focused result that can be used in round 2. This consolidated list is sent to the panel participants for validation prior to the next round of the Delphi study. This validation step is crucial according to Schmidt, as "without this step, there is no basis to claim that a valid, consolidated list has been produced" (Schmidt, 1997, p. 769).

Delphi round 2:

The consolidated list of categorized aspects from round 1, is sent out and the panel participants are asked to narrow the list down further by selecting the most important issues for further evaluation. Again, they are asked to explain their reasoning. This is done to basically reduce the size of the list to focus on the most important aspects according to the panel participants. The responses from round 2 are again analyzed and the new list of aspects are consolidated and sent back to the panel participants for validation and reflection.

[19] Aspect is used in the context of this thesis as:" a particular status or phase in which something appears or may be regarded" (Merriam-Webster.com, n.d.).

Delphi round 3:

The third round will consist of the narrowed down list of aspects which the panel participants are asked to rank to in accordance of their importance. The ranking may be done using a simple scale between 1 and 10. This is done to try to measure the actual consensus within the panel on what the important aspects are when they are subjected to rating them on for example a 10-point scale, where 1 is unimportant and 10 is the most important. The actual consensus of the panel can then be calculated using the method called Kendall's coefficient of concordance or Kendall W (Salkind, 2010). If the panel has not reached consensus on the ranking of the list after this round the list may be sent out again for re-evaluation by the panel participants in order to try to reach consensus by having them subjected to the other panel members reasoning and arguments (Skulmoski, et al., 2007).

3.3 Expected results

The results will give us data on what aspects actual security professionals value the highest today with regards to privacy and data protection. The result expected will not be connected to GDPR or requirements thereof specifically but will instead be connected to the important concepts of GDPR identified in the literature study, as mentioned earlier. The data will consist of a ranked list of aspects that can be addressed theoretically by the functionality of HSMs, meaning that can the aspects be addressed at all using HSMs? And if they can, how effectively are the HSM addressing the aspects? The list can also be mapped directly to the various requirements of the GDPR identified through the literature study, giving us, a much more fine-grained view of the aspects involved with the different GDPR requirements. The combined result can then answer RQ1 on a more specific level than the granular level of the concepts of the literature study.

The list of aspects that are left un-addressed will be used as the foundation for providing answers for RQ2, by identifying GDPR requirements that are left un-addressed by the implementation of HSMs and how that will affect the effort of reaching compliance with GDPR.

4 Result & Analysis

The results and analysis of the literature study and the Delphi study are presented in this chapter.

4.1 Compliance in the GDPR

The articles in the GDPR has been analyzed with regards to requirements in need of specific compliance actions from the data controller and processors and the findings are presented in Table 6 below (European Union, 2016).

Article	Title
5	Principles relating to processing of personal data
6	Lawfulness of processing
7	Conditions for consent
8	Conditions applicable to child's consent in relation to information society services
9	Processing of special categories of personal data
10	Processing of personal data relating to criminal convictions and offences
12	Transparent information, communication and modalities for the exercise of the rights of the data subject
13	Information to be provided where personal data are collected from the data subject
14	Information to be provided where personal data have not been obtained from the data subject
15	Right of access by the data subject
16	Right to rectification
17	Right to erasure ("Right to be forgotten")
18	Right to restriction of processing
19	Notification obligation regarding rectification or erasure of personal data or restriction of processing
20	Right to data portability
21	Right to object
22	Automated individual decision-making, including profiling
24	Responsibility of the controller
25	Data protection by design and by default
27	Representatives of controllers or processors not established in the union
28	Processor
30	Records of processing activities
32	Security of processing
33	Notification of a personal data breach to the supervisory authority
34	Communication of a personal data breach to the data subject
35	Data protection impact assessment
36	Prior consultation
37	Designation of the data protection officer
38	Position of the data protection officer
39	Tasks of the data protection officer
45	Transfer on the basis of an adequacy decision
46	Transfer subject to appropriate safeguards
47	Binding corporate rules
49	Derogations for specific situations
89	Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes

Table 6 - Compliance requirements in GDPR

4.2 Data Protection

Articles 24, 25, and 32 of the GDPR is requiring that technical measures are taken in accordance with the risk of the data processing. This means that the level of technical measures to take as a data processor/controller is based on the risk that the processing results in to the owner of the data. Article 32 further states that these measures need to have processes in place for "regularly testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing" (European Commission, 2012). The GDPR states that these measures need to be documented to demonstrate compliance (article 5(2) and 30), and requires organizations to implement data protection principles (article 25(1) such as data minimization (European Union, 2016).

As the wording of the regulation is somewhat vague and offers no insight on how data protection can be achieved other sources have been utilized to gather data on the principles and best practices of data protection. The Center for Internet Security[20] has issued a guide called "CIS Critical Security Controls for Effective Cyber Defense"[21] to aid organizations to achieve cyber security and compliance with security requirements (Center for Internet Security, 2017). This guide contains 20 controls that if implemented should help protect the organization against cyber-attacks. Table 7 below presents the 20 controls in the CIS guide:

1	Inventory of Authorized and Unauthorized Devices
2	Inventory of Authorized and Unauthorized Software
3	Secure Configurations for Hardware and Software
4	Continuous Vulnerability Assessment and Remediation
5	Controlled Use of Administrative Privileges
6	Maintenance, Monitoring, and Analysis of Audit Logs
7	Email and Web Browser Protections
8	Malware Defenses
9	Limitation and Control of Network Ports
10	Data Recovery Capability
11	Secure Configurations for Network Devices
12	Boundary Defense
13	Data Protection
14	Controlled Access Based on the Need to Know
15	Wireless Access Control
16	Account Monitoring and Control
17	Security Skills Assessment and Appropriate Training to Fill Gaps
18	Application Software Security
19	Incident Response and Management
20	Penetration Tests and Red Team Exercises

Table 7 - CIS Controls Overview

[20] Center for Internet Security is a nonprofit organization with the mission to safeguard private and public organizations against cyber threats.

[21] The CIS Controls are referenced by the National Institute of Standards and Technology (NIST) Cybersecurity Framework as a recommended implementation approach, and the European Telecommunications Standards Institute (ETSI) has adopted and published the CIS Controls guides.

Data protection is described in Critical Security Control (CSC) no. 13 with two actions that are categorized as "foundational"[22] (Center for Internet Security, 2016, p. 47). These actions are:

- Perform an assessment of data to identify sensitive information that requires the application of encryption and integrity controls.
- Deploy approved hard drive encryption software to mobile devices and systems that hold sensitive data.

CSC 13 also describes actions such as continuous scanning for unauthorized documentation containing sensitive data, and for sensitive data that is stored in plain text. CSC 13 also recommends storing encryption keys within HSMs (Center for Internet Security, 2016).

CSC 14 includes the recommendation that organizations need to know what sensitive information it possess, where it resides and who has access to it and proposes a data classification scheme with at least two levels: public (unclassified) and private (classified). CSC 14 also describes the need for access control based on need-to-know with the following steps (Center for Internet Security, 2016, pp. 50-51):

- Segment the network based on the label or classification level of the information stored on the servers. Locate all sensitive information on separated VLANS with firewall filtering to ensure that only authorized individuals are only able to communicate with systems necessary to fulfill their specific responsibilities.
- All communication of sensitive information over less trusted networks should be encrypted. Whenever information flows over a network with a lower trust level, the information should be encrypted.
- All information stored on systems shall be protected with file system, network share, claims, application, or database specific access control lists. These controls will enforce the principle that only authorized individuals should have access to the information based on their need to access the information as part of their responsibilities.
- Sensitive information stored on systems shall be encrypted at rest and require a secondary authentication mechanism, not integrated into the operating system, in order to access the information.
- Enforced detailed audit logging for access to nonpublic data and special authentication for sensitive data.

The Payment Card Industry (PCI) has issued its own data security standard called PCI DSS. It specifies requirements, procedures and best practices for organizations involved in payment card processing. These requirements, procedures and best practices can in many ways be transferred to the processing of personal data as well and help with compliance to GDPR. The PCI DSS has a set of 12 high level requirements presented in Table 8.

[22] Defined by CIS as "… essential improvements to the process, architecture, and technical capabilities of organizations to monitor their networks and computer systems to detect attack attempts, locate points of entry, identify already compromised machines, interrupt infiltrated attackers' activities, and gain information about the sources of an attack." (Center for Internet Security, 2016, p. 92).

Build and Maintain a Secure Network and Systems	1	Install and maintain a firewall configuration to protect cardholder data
	2	Do not use vendor-supplied defaults for system passwords and other security parameters
Protect Cardholder Data	3	Protect stored cardholder data
	4	Encrypt transmission of cardholder data across open, public networks
Maintain a Vulnerability Management Program	5	Protect all systems against malware and regularly update anti-virus software or programs
	6	Develop and maintain secure systems and applications
Implement Strong Access Control Measures	7	Restrict access to cardholder data by business need-to-know
	8	Identify and authenticate access to system components
	9	Restrict physical access to cardholder data
Regularly Monitor and Test Networks	10	Track and monitor all access to network resources and cardholder data
	11	Regularly test security systems and procedures
Maintain an Information Security Policy	12	Maintain a policy that addresses information security for all personnel

Table 8 - PCI DSS Overview

PCI DSS requirement no. 3 states the following for all cardholder data storage (PCI Security Standards Council, 2016) :

- (3.1) Keep cardholder data storage to a minimum by implementing data retention and disposal policies, procedures and processes that include at least the following for all cardholder data storage:
    - Limiting data storage amount and retention time to that which is required for legal, regulatory, and/or business requirements
    - Specific retention requirements for cardholder data
    - Processes for secure deletion of data when no longer needed
    - A quarterly process for identifying and securely deleting stored cardholder data that exceeds defined retention
- (3.4) Render Primary Account Number unreadable anywhere it is stored (including on portable digital media, backup media, and in logs) by using any of the following approaches:
    - One-way hashes based on strong cryptography
    - Truncation
    - Index tokens and pads (pads must be securely stored)
    - Strong cryptography with associated key-management processes and procedures
- (3.5) Document and implement procedures to protect keys used to secure stored cardholder data against disclosure and misuse
- (3.5.2) Restrict access to cryptographic keys to the fewest number of custodians' necessary
- (3.5.3) Store secret and private keys used to encrypt cardholder data in on (or more) of the following forms at all times:
    - Encrypted with a key-encrypting key that is at least as strong as the data-encrypting key, and that is stored separately from the data-encrypting key
    - Within a secure cryptographic device (such as a hardware (host) security module (HSM) or PTS-approved point-of interaction device)

- As at least two full-length key components or key shares, in accordance with an industry accepted method

The NIST Cybersecurity Framework provides organizations with guidance on how to better understand, handle and mitigate cybersecurity risks. It is a voluntary framework based on a number of existing standards, guidelines and practices. Version 1.0 of the framework was released in 2014 and version 1.1 in currently in development in and more than 3000 participants from industry, academia and the US government participate in the development of the framework (NIST, 2016).

The framework divides the categories of actions into five categories, identify, protect, detect, respond and recover. The data protection categories can be found in the protect function, under data security which is defined as follows: "Data Security (PR.DS): Information and records (data) are managed consistent with the organization's risk strategy to protect the confidentiality, integrity, and availability of information" (NIST, 2014). The data security category has the following subcategories among others:

- PR.DS-1 Data-at-rest is protected
- PR.DS-2 Data-in-transit is protected

The subcategories are then not further described but are instead supplied with the corresponding standard or procedural guide to adhere to. PR.DS-1 and PR-DS-2 refers to the above described CIS CSC, and to NIST SP-800-53.

NIST SP-800-53 "Security and Privacy Controls for Federal Information Systems and Organizations" provides security and privacy controls to protect organizations and their assets and individuals against cyber-attacks, natural disasters, structural failures and human errors (NIST Joint Task Force Transformation Initiative, 2013). NIST SP-800-53 covers a huge amount of aspects, access controls to threat awareness programs, the parts that the Cybersecurity Framework links to in PR.DS-1 and PR.DS-2 can be found in section Systems and Communication Protection (SC).

- SC-8 deals with transmission confidentiality and integrity and transmitted information. The control function here is to use cryptographic mechanisms to protect the data, protect the data headers and routing information, and protect the data pattern from disclosing clues about the contents based on frequency, size and amount.
- SC-12 describes how cryptographic keys are established and how to manage them. The control function states that NIST FIPS approved key management technology and processes and that the organization maintains information availability even when users lose keys.
- SC-13 describes cryptographic protection. NIST FIPS validated equipment and algorithms are suggested.
- SC-28 deals with protection of information at rest. Both when stored in on-line, meaning that the protected data is used and maintained regularly and when stored off-line, meaning stored as backups or archived.

4.3 Delphi Study

As previously mentioned, the expected number of participants for the Delphi study was set to 10, however, six experts accepted to participate and only five of them followed through with the study. This should not however affect the Delphi research in any considerable way, there are published studies with four participants (Gustafson, et al., 1973) and as low as three participants (Lam, et al., 2000). These experts were fully qualified to participate in the Delphi study based on the requirements presented in chapter *3.2.1 Delphi.*

Table 9 presents an overview of the participants' roles and the certifications included in each role.

Role	No. of participants	Certifications
Security Consultant	1	CISM (Certified Information Security Manager), RHCE (Redhat Certified Engineer)
Security Architect	2	CISSP, ITIL v2 (Information Technology Infrastructure Library),
Security Team leader	1	ITIL v3 Foundation, Associate of ICS2 (Information Security Certification)
Chief Sales Officer (CSO)	1	PCSE (Payshield, Thales HSM, Certified Systems Engineer)
Sum	5	

Table 9 - The number of participants with their current roles and certifications

It should be noted that one of the participants also have a Master of Laws in International Private Law from Stockholm University which adds a legal perspective to the study as well.

In the first round of the Delphi study, the participants were asked to state at least 10 critical aspects of securing an information system with regards to personal data protection. The brainstorming session resulted in a list of 49 aspects due to one participant only stating nine. Before creating a consolidating list of the identified aspects, an overview was obtained by separating the aspects into different categories. These categories are based on general concepts used in the information security/information system area to describe the different aspects related to data protection. The identified aspects were consolidated resulting in a list of 32 aspects that was returned to the participants for validation prior to the second round of the Delphi, *see Table 10 and Appendix A1 – Round 1 of the Delphi Study.*

Categories	No. of identified aspects	Description
InfoSec management	8	Includes aspects such as, creating and implementing policies, controls and guidelines for securing and restricting access to data. Also, involves the handling of security budgets and mandating security efforts within the organization.
Role management	4	Includes aspects such as setting access levels based on the employee's role within the organization, document access rights and verifying users.
Risk management	1	Having a clear process for analyzing risks and managing them within the organization.
Key management	1	Implementing a secure KMS that is supported by policies and processes within the organization
Education	4	To educate all employees, with access to personal data, about policies and privacy regulations to raise awareness. Training employees in different aspects of information security.

Incident response plan	1	Having a specific incident response plan in case of any type of breach regarding personal data.
Data inventory	3	To take inventory, classify data regularly, evaluate the mapping of data and document data and the legality of it.
Privacy by Design	1	Taking privacy, as defined in the GDPR, into account throughout the entire design process.
Technical	4	Properly protect and configure systems, devices and networks. Implementing and maintain proper password management with focus on usability.
Encryption	2	Centralized encryption management to have encryption coherence within the organization along with implementing strong encryption technology for protecting sensitive data.
Audit	3	Focus on designing systems for logging, auditing and use tools to support and facilitate forensic analysis in order to clearly demonstrate compliance.
Total	**32**	

Table 10 - Categories used for consolidating the identified aspects

After the validation process ended, the consolidate list was sent back out to the participants where they were asked to choose, according to their opinion, the 10 most critical aspects from the listed 32 aspects. If the list is less than 100 aspects the participants are required to select more than 10 percent of these aspects (Schmidt, 1997). All five participants managed to complete this phase of the Delphi study.

After receiving the responses from all participants, the researchers can eliminate all answers that are not selected by a simple majority (Schmidt, 1997). Since this study included five participants, all the aspects that were selected by less than three participants were eliminated from the list, this reduced the list from 32 aspects to seven aspects and was sent back to the participants for a validation prior to the third round. Table 11 below presents the categories that moved on the final round and the number of aspects in each. See *Appendix A2 – Round 2 of the Delphi Study* for further details.

Categories	No. of aspects
InfoSec Management	3
Education	1
Incident response plan	1
Data inventory	1
Audit	1
Total	**7**

Table 11 - Categories chosen by the simple majority and the number of aspects in each.

After the validation process, the list of the seven aspects was sent back to all the participants asking them to rank these aspects from 1 to 7 in order of importance regarding personal data protection, where 1 is considered the most critical aspect and 7 being the least critical one. The experts were also required to motivate their rankings since it is suggested that the experts would reach consensus more quickly if they could take part of each other's justifications (Okoli & Pawlowski, 2004).

The consensus between the experts was measured using Kendall's *W* coefficient of concordance since it is seen as the best way for measuring non-parametric tests (Okoli & Pawlowski, 2004; Schmidt, 1997). The coefficient *W* ranges from 0 to 1 where 0 indicates that there is no consensus and 1 indicates that there is perfect consensus between the panel participants (Okoli & Pawlowski, 2004). If the value of *W* is 0.7 or greater it would be considered as a strong agreement (Schmidt, 1997) whereas anything less than the aforementioned value would require the list, with each participant's justification, to be resent to the experts in order to try to achieve a stronger agreement between them (Okoli & Pawlowski, 2004).

The first ranking round provided a *W* value of 0.363 which suggests a weak agreement (Schmidt, 1997). This result indicates that a second-round need to be conducted in order to try and elevate the level of consensus between the participants. As suggested by Okoli and Pawlowski (2004), the aspects from the first ranking round were listed in order of their mean value before sending the list back to the experts. Based on O'Neill, Scott and Conboy (2009), for each aspect, the following was sent back to each panel member for the second ranking round:

- The current level of consensus based on the value of *W*
- The mean rank of each aspect
- How the expert himself had ranked the aspect (This was an addition of our own and not stated by O'Neill, Scott and Conboy)
- The comments and views of the other experts on each aspect

The experts were asked to go through this information to see if they would like to modify their rankings and answers from the previous ranking round, and to try providing an explanation to why they might have changed their opinion. See *Appendix A3 – Round 3 of the Delphi Study* for further details.

In the second ranking round, only two out of the five experts decided to change their previous ranking and only one of them gave a motivation to why he decided to change it. One of the participants who did not want to change his answer provided a justification for his reasoning. See *Appendix A3 – Round 3 of the Delphi Study* for further details. After this ranking round, Kendall's *W* improved to 0.620 which is considered to be somewhere between a moderate agreement and a strong agreement (Schmidt, 1997). Since most the panel experts did not want to change their answers and the minority were satisfied with their change, we decided to end the Delphi study at this point.

Aspect Rank	Aspect	Mean Rank
1	Create and implement policies and controls for access to sensitive data.	**1,2**
2	Map data flows of classified and sensitive data. Take inventory and create classification of all data regularly. Have policies in place to guide evaluation and mapping of data.	**3**
3	Appoint a professional and validated DPO.	**3,2**
4	Make Information Security an integral part of the organization, and train personnel in the different aspects of Information Security.	**4**
5	Create, use and maintain a specific data incident/breach plan for personal data.	**4,8**
6	Use auditing to secure the organization (Internal and External)	**5,8**
7	Change management needs to be integral to the organization and changes must be analyzed for risks and effects prior to implementation.	**6**

Table 12 - Final Result of the Delphi Study

Table 12 presents the final results of the Delphi study showing the final rank and the mean rank of each of the seven most critical aspects.

4.4 Answering the Research Questions

The analysis of the results provided answers for the research questions and are described below.

4.4.1 How can the use of HSM aid in achieving compliance with GDPR?

On the 25th of May 2018, the GDPR comes into effect and a lot of organizations will have some work to do before reaching compliance with the regulation. This thesis set out looking for how encryption, specifically through the use of HSMs, could influence the work towards said compliance.

Compliance with GDPR will require changes in many organizations, especially in how they work with data that is classified as personal data according to GDPR. The different standards and checklists studied for this thesis have a couple of things in common: Identify and classify data in levels of sensitivity, protect sensitive data with encryption, and use a KMS.

So, can HSMs help reach some compliance with GDPR? Let's start to look at what HSMs are good at. They are great at performing advanced cryptographic functions such as encryption and decryption of data and they have built in key-management systems that are designed to keep the keys used for encryption and decryption safe. They are also equipped to validate users and requests for access to encrypted data by different identification methods and tokens which means that HSMs can assist with access controls such as role-based-authentication and identity-based-authentication.

The proper implementation (according to industry standards and best practices) of HSMs for encryption purposes also entails performing some activities and actions that are required by the GDPR as well. First off, organizations that want to protect data by encrypting it must know what data to protect. This means that the organization needs to start mapping all data, classify it according to sensitivity and value and then to find out exactly where the data is stored, used and transmitted. This data mapping or data inventory is critical to become compliant with GDPR and to be able to utilize

encryption in a valuable way. By doing the work of data inventory, the organization also gains knowledge of what data it actually gathers, and may start to minimize the data collected to only the absolute necessary amount for the purpose of their processing. This would hopefully lead to data minimization, a review of the lawfulness of the gathered data, as well as audit trails and logging.

In addition to all this, the data inventory is the foundation of the organizations ability to respond to requests from the data subject on information on what data the organization has stored regarding that specific data subject, and consequently also be able to respond to requests by the data subject to be forgotten.

So, by choosing HSMs as the primary method to safeguard data through encryption and by implementing it according to the industry standards and best practices, organizations should also achieve partly compliance with the following articles of the GDPR:

- Article 5 – Principles relating to processing of personal data
    - States the requirement of “accountability”.
        - HSMs would here address the logging of the access to the personal data and thus create and maintain an audit trail of the processing.
- Article 24 – Responsibility of the controller
    - States the requirement of “technical and organizational measures” to ensure the “rights and freedoms of natural persons”.
        - The implementation of HSMs according to the industry best practices and standards would be considered such a technical and organizational measure.
- Article 25 – Data protection by design and default
    - States the requirement of “technical and organizational measures” to ensure the “rights and freedoms of natural persons”.
        - The implementation of HSMs according to the industry best practices and standards would be considers such a technical and organizational measure.
- Article 32 – Security of processing
    - States the requirement of “technical and organizational measures” to ensure the “rights and freedoms of natural persons”.
        - The implementation of HSMs according to the industry best practices and standards would be considers such a technical and organizational measure.
    - Encryption of personal data.
        - The implementation of HSMs according to the industry best practices and standards would provide security by encryption.
    - Ability to restore personal data in the event of incident.
        - HSMs can backup keys and restore them if necessary and would therefore be part of a restoration system.
    - Prevent disclosure, loss, alteration and accidental or unlawful destruction.
        - These requirements are the core function of HSMs.
    - Access controls.
        - The implementation of HSMs according to the industry best practices and standards would provide the means and measures for access control of sensitive data.
- Article 34 – Communication breach of a personal data breach to the data subject, would be avoided altogether by successful implementation of article 32.

The approach with HSMs alone would result in partly compliance with articles 5, 24, 25, and 32 of the GDPR. However, encryption alone does not mean compliance with GDPR aside from parts of article 32 of the regulation.

The implementations of HSMs should also push the organization to develop a strategy regarding the key management to protect the valuable and encrypted data in a way that ensures that only the authorized gains access and that the data is not inadvertently lost due to lost or corrupted keys. This would also help with GDPR in the sense that the organization can show due diligence in their efforts to safeguard the personal data.

4.4.2 What GDPR requirements would be left un-addressed by using such an approach?

The result from the Delphi panel is a list of the seven most important aspects of securing personal data.

Rank	Aspect	Mapping to GDPR
1	Create and implement policies and controls for access to sensitive data.	**Article 32** – Security of processing
2	Map data flows of classified and sensitive data. Take inventory and create classification of all data regularly. Have policies in place to guide evaluation and mapping of data.	**Article 30** – Records of processing activities
3	Appoint a professional and validated DPO.	**Article 37** – Designation of the data protection officer. **Article 38** – Position of the data protection officer
4	Make Information Security an integral part of the organization, and train personnel in the different aspects of Information Security.	**Not considered as a requirement by the GDPR but relates to:** **Article 5** – Principles relating to processing of personal data **Article 25** – Data protection by design and by default **Article 32** – Security of processing
5	Create, use and maintain a specific data incident/breach plan for personal data.	**Article 33** – Notification of a personal breach to the supervisory authority **Article 34** – Communication of a personal data breach to the data subject
6	Use auditing to secure the organization (Internal and External)	**Article 5** – Principles relating to processing of personal data **Article 24** – Responsibility of the controller
7	Change management needs to be integral to the organization and changes must be analyzed for risks and effects prior to implementation.	**Article 24** – Responsibility of the controller

Table 13 - The final aspect ranking

As can be seen in *Table 13*, the panel did not include encryption in their top seven. Instead the highest ranked aspect is a call for policies and controls on how access to sensitive data is handled. The consensus of the panel is rated as at least “moderate agreement” with a Kendall W value of 0.62 (Schmidt, 1997) which shows that the focus of becoming compliant with the GDPR should be on policies and procedures primarily since the bulk of requirements of the GDPR can be addressed by managerial actions.

The Delphi-panel ranks the data inventory and classification as the second most important aspect to address when dealing with sensitive data like personal information. The GDPR requirement of appointing a DPO is at number three, (directly linked to article 37 of the GDPR). The need for an overall organizational security policy and “mindset” is placed as number four, and the need for planning for the worst is put at number five in the form of incident and breach planning. Logging and auditing is selected as number six on the top seven list, and is also a powerful tool and requirement in the GDPR. Finally, the experts of the Delphi panel address the need for change control and management within organizations in order to prevent and mitigate risks born out of new developments and updates.

The resulting list of compliance aspects that would remain even after a HSM implementation is presented below in *Table 14*. Articles 5, 24, 25, and 32 are highlighted in yellow as they are in part addressed by HSMs. Aside from these four articles the list of compliance aspects and their corresponding risks are mostly left un-addressed.

Article	Title
5	Principles relating to processing of personal data
6	Lawfulness of processing
7	Conditions for consent
8	Conditions applicable to child's consent in relation to information society services
9	Processing of special categories of personal data
10	Processing of personal data relating to criminal convictions and offences
12	Transparent information, communication and modalities for the exercise of the rights of the data subject
13	Information to be provided where personal data are collected from the data subject
14	Information to be provided where personal data have not been obtained from the data subject
15	Right of access by the data subject
16	Right to rectification
17	Right to erasure ("Right to be forgotten")
18	Right to restriction of processing
19	Notification obligation regarding rectification or erasure of personal data or restriction of processing
20	Right to data portability
21	Right to object
22	Automated individual decision-making, including profiling
24	Responsibility of the controller
25	Data protection by design and by default
27	Representatives of controllers or processors not established in the union
28	Processor
30	Records of processing activities
32	Security of processing
33	Notification of a personal data breach to the supervisory authority
34	Communication of a personal data breach to the data subject

35	Data protection impact assessment
36	Prior consultation
37	Designation of the data protection officer
38	Position of the data protection officer
39	Tasks of the data protection officer
45	Transfer on the basis of an adequacy decision
46	Transfer subject to appropriate safeguards
47	Binding corporate rules
49	Derogations for specific situations
89	Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes

Table 14 - List of residual aspects

As mentioned earlier in the study, article 34 - Communication of a personal data breach to the data subject, can be eliminated totally if the requirements of article 32 are fulfilled.

5 Discussion & Conclusion

Reflecting on the choices made during the process of writing this thesis there are a few that might warrant some discussion.

Firstly, the choice to limit the list from round 2 of the Delphi study to just seven aspects. This was done because of the decision to choose the aspects based on simple majority, and since there were five participants in the panel that meant that more than two would have to choose an aspect for it to make its way to round 3. The choice to let simple majority rule was done because of not wanting to add bias to the study by setting a lower bar just to allow for more aspects to pass on to the third round. After all, only one participant had "encryption" as an aspect in round one and only two had "key management". Selecting aspects chosen by at least two participants would have added seven additional aspects to the third round, including those that was focusing on encryption as data protection and key management, and deviating from the selected method of major majority to get these aspects that corresponded well with the thesis core purpose was decided against mainly because of the previously stated reason of bias. In addition to that discussion the choice to use simple majority as the deciding factor to reduce selections in a Delphi-panel was also suggested in the 1997 article "Managing Delphi Surveys Using Nonparametric Statistical Techniques" by Roy C. Schmidt (Schmidt, 1997).

Secondly, the choice to continue with the Delphi study even when the initial goal of 10 participants could not be met. This was a choice partly based on studying previous Delphi studies and determining that having six (later five) participants was not unheard of. As mentioned in chapter 4.3 there are numerous studies with fewer than 10 participants. Gene Rowe and George Wright in their 1999 article "The Delphi technique as a forecasting tool: issues and analysis" presents a table of scientific work using Delphi-panels and lists 15 studies that have had less than 10 participants and 10 studies with five or less participants. After all, it is the quality of the participants that is the core of the Delphi study (Skulmoski, et al., 2007).

5.1 Contribution

This research will contribute in giving organizations an overview of what GDPR articles are addressed and what are not with the use of hardware security modules. We also think that the aspects provided by the experts through our Delphi study could give the researchers and practitioners an indication on what the most critical points to consider are, when it comes to data protection.

The final list of the top seven most critical aspects could provide a small sample of what to focus on the most regarding data protection, and the list of 49 aspects from the first Delphi-round provides a wider picture of what the experts deem to be the most critical. Researchers and practitioners could use this information, and justification from the panel experts, to gain further insight and understanding of the importance of each aspect and what role each aspect has in data protection. Our Delphi study could also be used as a basis for further research with Delphi panels of similar expertise to determine if the outcome would be similar to this study.

Organizations that are aiming at reaching compliance with GDPR and are considering implementing an HSM as a solution could benefit from this study by getting an unbiased overview of what the HSM can do and what not with regards to complying with the GDPR. This research could also benefit the HSM vendors where they could use this study as a source or refer to it when making claims regarding HSMs and their role in complying with the GDPR.

Overall, this research contributes to the information security field of study in general, and data protection with GDPR specifically, where there is yet no published academic literature covering this topic.

5.2 Final Thoughts and Conclusion

The intent of this study was to research how implementing hardware security modules would aid in reaching compliance with the GDPR and which GDPR requirements would be left unaddressed. The literature study showed us that the HSM alone would only partly comply with four articles of the GDPR given that it is a technical solution, whereas the rest of the articles would require a more management oriented approach such as conducting impact assessment, designate a DPO, formulating new user agreements that would include the new user rights such as the "Right to be forgotten" and the list goes on. It is obvious that a HSM would not be able to contribute to these types of requirements. The results of the Delphi panel reinforce this idea even more where we could see that the experts did not consider a main technical solution to be part of the final list of the seven most critical aspects for personal data protection. The focus was mainly on managerial/policy aspects where a HSM could act more as a complementary tool, not as a main solution, to help facilitate the implementation of the policies.

This does not however mean that the implementation of a HSM is a bad idea, on the contrary, it is still considered to be one of the most powerful tools when it comes to encryption and key management solutions, and could still be considered as a big piece of the GDPR compliance puzzle when it comes to personal data protection. The HSM would after all address a big part of the requirements of article 32 of the GDPR which is in line with the reasoning of Thales E-Security in their whitepaper discussed in chapter 2.6 Knowledge Gap (Thales E-Security, 2017) and also aid in facilitating the number one ranked aspect of securing personal data according to the Delphi panel.

Finally, it should be stated that the lack of focus on encryption and key management from the Delphi-panel was surprising to us as researchers. Upon reflection, this lack of technical focus is also evident in the studied literature on the GDPR and on data protection, but going in to this study with the aim to study how a technical solution, such as how HSMs, could be used to address compliance requirements of the GDPR probably meant that we had a slight bias and initially thought that the abilities of HSMs would be more sought after by the security professionals in the Delphi study as well as in the literature.

Further research on these subjects needs to be done when the GDPR has come into full force in May of 2018 to analyze the impact it will have on the industry that is processing personal data, on organizations within and outside of the EU, on the citizens within and outside the EU, and on the information security field of study as a whole.

References

Acosta, A. J., Addabbo, T. & Tena-Sanchez, E., 2016. Embedded eclectronic circuits for cryptography, hardware security and true random number generation: an overview. *International Journal of Circuit Theory and Applications,* Issue Special Issue Paper.

Agarwal, A. & Agarwal, A., 2011. The Security Risks Associated with Cloud Computing. *International Journal of Computer Applications in Engineering Sciences,* 1(Special Issue On CNS), pp. 257-259.

AlAhmad, M. A. & Alshaikhli, I. F., 2013. Broad View of Cryptographic Hash functions. *International Journal of Computer Science Issues,* 10(4), pp. 239-246.

Ariwibowo, S. & Windarta, S., 2016. *Distinguishing Attack and Second-Preimage Attack on Encrypted Message Authentication Codes (EMAC).* Yogyakarta, AIP Publishing LLC..

Article 29 Data Protection Working Party, 2014. *Opinion 03/2014 on Personal Data Breach Notification.* [Online]
Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp213_en.pdf
[Accessed 02 March 2017].

Bloor, 2017. *Welcome to Bloor.* [Online]
Available at: http://www.bloor.eu/
[Accessed 14 June 2017].

Bolognini, L. & Bistolfi, C., 2016. Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation. *The Internation Journal of Technology Law and Practice.*

Center for Internet Security, 2016. *The CIS Critical Security Controls for Effective Cyber Defence Version 6.1,* New York: Center for Internet Security.

Center for Internet Security, 2017. *CIS Control.* [Online]
Available at: https://www.cisecurity.org/controls/
[Accessed 31 May 2017].

Chandramouli, R., Iorga, M. & Chokani, S., 2014. Cryptographic Key Management Issues & Challenges in Cloud Services. In: *Secure Cloud Computing.* New York: Springer, pp. 1-30.

de Hert, P. & Papakonstantinou, V., 2014. The Council of Europe Data Protection Convention Reform: Analysus of the new text and critical comment on its global ambition. *Computer Law & Security Review,* 30(6), pp. 633-642.

DQM GRC, 2016. *About us.* [Online]
Available at: http://www.dqmgrc.com/about-us
[Accessed 14 June 2017].

Esaiasson, P., Giljam, M., Oscarsson, H. & Wängerud, L., 2012. *Metodpraktikan: Konsten att studera samhälle, individ och marknad.* 4th ed. Stockholm: Norstedts Juridik AB.

European Commission, 2012. *Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) Com(2012) 11 final.* Brussels: European Commission.

European Commission, 2015. *European Commission Press Release Database.* [Online]
Available at: http://europa.eu/rapid/press-release_MEMO-15-6385_en.htm
[Accessed 18 November 2016].

European Commission, 2016. *Article 29 Working Pary.* [Online]
Available at: http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=50083
[Accessed 02 March 2017].

European Commission, 2016. *Data Protection.* [Online]
Available at: http://ec.europa.eu/justice/data-protection/bodies/index_en.htm
[Accessed 10 February 2017].

European Commission, 2016. *What is an SME?.* [Online]
Available at: http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition_en
[Accessed 18 November 2016].

European Commission, C. f. t. C. t. t. E. P. t. C. t. E. a. S. C. a. t. C. o. t. R., 2010. *A comprehensive approach on personal data protection in the European Union, COM(2010) 609 final.* Brussels: European Commission.

European Data Protection Supervisor, 2017. *Duties.* [Online]
Available at:
https://secure.edps.europa.eu/EDPSWEB/edps/cache/offonce/EDPS/Membersmission/Duties
[Accessed 10 February 2017].

European Union, 2016. *Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).* Brussels: Official Journal of the European Union.

Gemalto, 2017. *Who we are.* [Online]
Available at: http://www.gemalto.com/companyinfo/about
[Accessed 14 June 2017].

Gemalto, n.d. *EU Compliance: General Data Protection Regulation (GDPR).* [Online]
Available at: https://safenet.gemalto.com/data-protection/data-compliance/european-union-eu-compliance/
[Accessed 13 June 2017].

Gilbert, F., 2016. EU General Data Protection Regulation: What Impact for Businesses Established Outside The European Union. *Journal of Internet Law,* pp. 3-8.

Gustafson, D. H., Shukla, R. K., Delbecq, A. & Walster, G. W., 1973. A comparison study of difference in subjective likelihood estimate made by individuals , interacting groups, Delphi groups and nominal groups.. *Organizational Behavior and Human Performance,* 9(2), pp. 280-291.

Halal, W. E., Kull, M. D. & Leffmann, A., 1998. The Goeorge Washington University Forecast of Emerging Technologies: A Continuous Assessment of the Technology Revolution. *Technological Forecasting and Social Change,* Volume 59(1), pp. 89-110.

Howarth, F., 2016. *For the EU's new data protection regulation, encryption should be the default,* London: Bloor Research Ltd..

ISO, 2015. *ISO Standards Catalogue.* [Online]
Available at:
http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=44404
[Accessed 08 February 2017].

Lam, S. S. Y., Petri, K. L. & Smith, A. E., 2000. Prediction and optimization of a ceramic casting process using a hierarchical hybrid system of neural networks and fuzzy logic.. *IIE Transactions,* 32(1), pp. 83-92.

Landrum, B. & Garza, G., 2015. Mending Fences: Defining the Domains and Approaches of Quantitative and Qualitative Research. *Qualitative Psychology,* 2(2), pp. 199-209.

Leedy, P. D. & Ormrod, J. E., 2015. *Practical Research: Planning and Design.* 11th ed. Harlow: Pearson Education Limited.

Merriam-Webster.com, n.d. *Aspect.* [Online]
Available at: https://www.merriam-webster.com/dictionary/aspect
[Accessed 5 June 2017].

Merriam-Webster, n.d. *Compliance.* [Online]
Available at: https://www.merriam-webster.com/dictionary/compliance
[Accessed 03 March 2017].

NIST Joint Task Force Transformation Initiative, 2013. *NIST Special Publication 800-53 Revision 4, Security and Privacy Controls for Federal Information Systems and Organizations,* Gaithersburg: National Institute of Standards and Technology.

NIST, 2002. *NIST Computer Security Resource Center.* [Online]
Available at: http://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf
[Accessed 10 January 2017].

NIST, 2012. *NIST Special Publications.* [Online]
Available at: http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-133.pdf
[Accessed 23 March 2017].

NIST, 2014. *Framework for Improving Critical Infrastructure Cybersecurity Version 1.0,* n.a.: National Institute of Standards and Technology.

NIST, 2015. *NIST Special Publications.* [Online]
Available at: http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-131Ar1.pdf
[Accessed 10 February 2017].

NIST, 2016. *Cybersecurity Framework FAQS Framework Basics.* [Online]
Available at: https://www.nist.gov/cyberframework/cybersecurity-framework-faqs-framework-basics
[Accessed 1 June 2017].

NIST, 2016. *NIST.* [Online]
Available at: https://www.nist.gov/about-nist
[Accessed 10 February 2017].

NIST, 2016. *NIST Special Publications.* [Online]
Available at: http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-57pt1r4.pdf
[Accessed 23 March 2017].

NIST, 2017. *Resource Center.* [Online]
Available at: http://csrc.nist.gov/groups/STM/cmvp/validation.html
[Accessed March 2017].

Okoli, C. & Pawlowski, S. D., 2004. The Delphi method as a research tool: an example, design considerations and applications. *Information & Management,* 42(1), pp. 15-29.

O'Neill, S., Murray, S. & Conboy, K., 2009. *A Delphi study on collaborative learning distance education.* s.l., European Conference on Information Systems (ECIS).

OWASP, 2016. *Key Management Cheat Sheet.* [Online]
Available at: https://www.owasp.org/index.php/Key_Management_Cheat_Sheet
[Accessed 23 March 2017].

Pattinson, F., 2012. *ISO's Cryptographic Module Work,* Austin: Atsec Information Security Corporation.

PCI Security Standards Council, 2016. *Data Security Standard - Requirements and Security Assessment Procedures Version 3.2,* Wakefield: PCI Security Standards Council, LLC.

Prowse, D. L., 2015. *CompTIA Security+ SY0-401 Cert Guide, Deluxe Edition.* 3rd ed. Indianapolis: Pearson Education, Inc.

Rowe, G. & Wright, G., 1999. The Delphi technique as a forecasting tool, issues and analysis. *International Jorunal of Forecasting,* 15(4), pp. 353-375.

RSA Security Inc., 2004. *PKCS#11 v.2.20: Cryptographic Token Interface Standard,* s.l.: RSA Laboratories.

SafeNet Gemalto, 2010. *Securing Network-Attached HSMs: The SafeNet Luna SA Three-Layer Authentication Model,* s.l.: SafeNet Inc..

Saha, T., 2015. An Enhanced Approach to Secure Message Using Combination of Symmetric and Asymmetric Cryptography and Triangulation. *International Journal of Latest Trends in Engineering and Technology,* 5(1), pp. 296-301.

Salkind, N. J., 2010. *Encyclopedia of Research Design.* Thousand Oakes, CA: Sage Publications Ltd.

Schmidt, R. C., 1997. Managing Delphi Surveys Using Nonparametric Statistical Techniques. *Decision Sciences,* 28(3), pp. 763-774.

Skulmoski, G. J., Hartman, F. T. & Krahn, J., 2007. The Delphi Method for Graduate Research. *Journal of Information Technology Education,* Volume 6, pp. 1-21.

Solterbeck, A., 2006. Data Encryption: Protecting data at rest and in motion. *Network Security,* 13(9), pp. 14-17.

Spindler, G. & Schmechel, P., 2016. Personal Data and Encryption in the European General Data Protection Regulation. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law,* 7(19), pp. 163-177.

Stallings, W. & Brown, L., 2012. *Computer Security Principles and Practice.* 2nd ed. Harlow: Pearson Education Gate.

Tankard, C., 2016. What GDPR means for business. *Network Security,* 23(6), pp. 5-8.

Tankard, C., 2017. Encryption as the cornerstone of big data security. *Network Security,* 24(3), pp. 5-7.

Thales E-Security, 2016. *Vormetric Transparent Encryption.* [Online]
Available at: https://www.thalesesecurity.com/sites/default/files/2017-02/sb-vormetric-transparent-encryption.pdf
[Accessed 14 June 2017].

Thales E-Security, 2017. *General Data Protection Regulation (GDPR) Compliance.* [Online]
Available at: http://go.thalesesecurity.com/rs/480-LWA-970/images/General_Data_Protection_Regulation_WP_INT.pdf
[Accessed 13 June 2017].

Thales Group, 2016. *About us.* [Online]
Available at: https://www.thalesgroup.com/en/about-us
[Accessed 14 June 2017].

Thales, 2015. *Thales Security World,* s.l.: Thales.

The Swedish Data Protection Authority , 2012. *Inbyggd integritet.* [Online]
Available at: http://www.datainspektionen.se/Documents/faktablad-inbyggd-integritet.pdf
[Accessed 16 February 2017].

Tiwari, H. & Asawa, K., 2012. A secure and efficient cryptographic hash function based on NewFORK-256. *Egyptian Informatics Journal,* 13(3), pp. 199-208.

Treacy, B., 2010. Working Party confirms 'controller' and 'processor' distinction. *Privacy & Data Protection,* 10(5), pp. 3-5.

Tutorialspoint, 2017. *Tutorialspoint.* [Online]
Available at: https://www.tutorialspoint.com/cryptography/public_key_encryption.htm
[Accessed 13 February 2017].

Webster, J. & Watson , R. T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly,* 26(2), pp. 13-23.

Whitman, M. E. & Mattord, H. J., 2014. *Management of Information Security.* 4th ed. Stamford: Cengage Learning.

Wikipedia Contributors, 2017. *List of applications using PKCS 11.* [Online]
Available at:
https://en.wikipedia.org/w/index.php?title=List_of_applications_using_PKCS_11&oldid=761734703
[Accessed 27 April 2017].

Wilhelm, E.-O., 2016. *A brief History of the General Data Protection Regulation.* [Online]
Available at: https://iapp.org/resources/article/a-brief-history-of-the-general-data-protection-regulation/
[Accessed 10 February 2017].

Wright, D., 2013. Making Privacy Impact Assessment More Effective. *The Information Society,* 29(5), pp. 307-315.

Appendix A1 – Round 1 of the Delphi Study

This appendix includes all the 49 aspects gathered from the first round along with the consolidated list of 32 aspects.

Appendix A1.1 – All the Aspects from Round 1

Aspect No.	Aspect	Context
1.	**An Information Security Office and IT security office that has mandate or budget to drive security efforts**	It's common to have a security officer with neither budget or a mandate, which means no efforts will ever be carried through
2.	**A security program sponsored by Executive level management**	If senior management does not see the value of a governance model/program for Information security that in turn drives IT security, then no security effort can achieve its goals. The goal of security is securing resources at a cost that is acceptable to the business – and this cannot be determined by IT. Security must be driven by a cost-vs-risk valuation, and that is a business decision.
3.	**Access control**	Our access models are often faulty concerning granularity, defining right access for everyone. It is compromises. There exist some tools saying they solved this, but they are very expensive.
4.	**Access process**	I only seen one organization fulfilling SOX access control process, being the same as for GDPR. IT was not an IAM or IAAM system, but an administrative support for who to get access to what and signals to end access. IAM was the step after.
5.	**Admin access**	We need "strong" users handling functional issues. But how to limit these access, so they cannot misuse (remember 80% of all IT crime is still internal).
6.	**Appoint a DPO**	You need to have a DPO that can have the oversight of the personal data in an organization, the DPO is also a resource to get support in privacy questions

7.	**Audit support**	Too many systems have faulty log functions and bad configs, making it hard for an auditor to ID who did what.
8.	**Awareness of personal data**	often system owners/administrators/users are not aware of what by definition personal data is
9.	**Change management**	All changes – including deployment of new systems – should be subject to change management where a change is vetted by a change management board who is responsible for assuring that all processes are respected
10.	**Control of who has access to personal data**	Often the system owner is the role that should have the knowledge of where the personal data is accessed, but often some other role is determining the access rights.
11.	**Data analysis and flow mappings for technical systems**	Every system must have documentation regarding what information it stores or processes so that proper protections can be selected.
12.	**Document data flows of personal data**	You need to know how personal data is flowing to and from your information system.
13.	**Document legal basis for processing personal data**	You need to be able to demonstrate under what legal rights you are processing subject's personal data.
14.	**Education/Training**	in my experience, often the education and training of privacy is something that's regarded as a cost that one can cut, but there is actually a requirement in GDPR
15.	**Information Security policies**	Information Security needs to produce policies to guide the evaluation of data resources
16.	**IT Security standards, guidelines etc.**	IT Security needs to have a set of "off the shelf" guides to secure resources according to the value of the resources

17.	**IT technical solutions for supporting functions**	Common security tasks that are common between systems should be in place and maintained and well secured, for example Identity management, authentication services, network security, VPN/Remote Access, Endpoint security etc.
18.	**Logging of CRUD (Create, Read, Update, Delete) actions on Personal data**	Logging of who have done what with the personal data is a critical point to be able to demonstrate that you have control over who has done what with the personal data
19.	**Logging system**	We need better log analysis, from servers, networks and maybe even from NiPS systems. We need to extract info easier to get warnings.
20.	**Maintain a data privacy incident/breach response plan**	As the requirements of notify regulators and data subjects of incidents/breach of the personal data, you need to have a plan about how and what you are going to communicate with data subjects and regulators.
21.	**Maintain procedures for responding to data subjects' requests**	You need to be able to respond to data subject's requests of their personal data, then there needs to be procedures about how you comply with such requests, you need to know/validate that the information system can handle for example the right to be for gotten.
22.	**Missing DPO, or officer without knowledge or power to do a good work.**	There is always a big need to have a responsible that has the knowledge and strength to implement the needed security.
23.	**Missing or poor main security policy.**	If a security policy is not good and implemented, most of following security work is at same poor quality.
24.	**Mitigating**	OWASP Top 10 and SANS Top 25 vulnerabilities. Some of these is +30 years old and still being used by attackers.
25.	**Network Configuration**	See No 4 Server, the same concerns. Layered network security - A bit like the Tor Onion model. The open flat network is dead.

26.	**No educated employees, or employees that do not get time to spend on security activities.**	If no one can do the work, it will not be done.
27.	**No process for following up on security activities.**	Internal audits, external audits can help keeping the organization on its toes.
28.	**No/poor handling of Configuration Items, i.e. equipment.**	Do be able to secure something you need to know what you have, what it is worth if compromised.
29.	**No/poor process for Risk Assessments.**	In the risk assessments, the need for security is found.
30.	**Only acquire solutions and software that can demonstrate a secure-by-design and secure-by-default approach**	Security needs to be an integral part in the requirements specification for acquisition of software or solutions.
31.	**Passwords**	Reached end of life. A 2-phase system, personally I would like Yubikey
32.	**Poor management of employees and subcontractors.**	Insider crime/theft/ etc. can be the result if the staff is not made happy, and is divided with separation of duties and other measures to prevent security events.
33.	**Poor or no key management for crypto.**	It does not matter how thick your door is if you leave the keys in the door lock.
34.	**Poor process/knowledge to buy security competence and resources**	Mot organization might need to find professional help to secure its resources
35.	**Privacy by design**	You need to make sure that services and products are handled and maintained with privacy by design, the services and products should be default secure for personal data
36.	**Protection has to be such that people still can work**	Too hard security is a risk, people finding unauthorized routes.
37.	**Server Configurations**	Too many servers do not have a good and secure config, they use manufacturer's default that are to relax for GDPR. Reducing protocols to secure ones, remove a lot of today open attack vectors.
38.	**The organization does not understand difference between information & IT security.**	It is important to understand that IT-security is just one part of information security.

39.	**User and administrator education**	Both users and admins need to be up-to-date with the security program, especially regarding their own responsibilities etc.
40.	**Creating and maintaining a comprehensive inventory of information assets, Information classification policies and corresponding information security policies.**	And understanding that these are not static “manuals" but living documents, so coupling these inventories and policies to a constantly ongoing change management process is the key here.
41.	**Doing regular risk analyses of the systems and information assets, in order to adapt to a changing threat landscape**	
42.	**Architect the system facilitate forensic analysis, and to demonstrate compliance.**	Not doing so is like not having done the work of inventory and classification of the data assets. It creates inefficiencies due to poor control of security.
43.	**Architect systems according to GDPR principles and rights of subjects, e.g. data minimization principles.**	This is not about compliance but about common sense. Data that is not needed should be cleansed on a regular basis. Only data which is necessary to communicate should be communicated.
44.	**Minimize un-structured in-data. Such data is a big threat to personal data protection because it is much more difficult to monitor.**	And GDPR does not exempt un-structured data from data protection requirements. A classic example is free text opinions in health care, school or social care systems, where very sensitive information can be recorded.
45.	**Change management needs to be a part of the culture of any organization.**	People, systems and the world around us are always undergoing change.
46.	**Strong encryption technology and centralized management of encryption.**	It goes without saying that poor encryption will not keep sensitive personal data safe, and having islands of encryption just because Microsoft, Oracle and the rest have not managed to build joined up standards, creates obvious weaknesses in security design and a higher probability of human error.

47.	**Good Key Management processes and systems**	It is still common to see strong encryption which is not coupled with good processes and systems for enforcing key management policies and access controls. Locking a super secure vault with a key and then leaving the key lying around outside the vault is just bad security.
48.	**IAM processes and systems (2FA, PAM)- tightly linked to the point about key management.**	Unless we introduce policies and controls for privileged access to sensitive data and introduce multi factor authentication to ensure the "technical" security of a digital identity, protecting the data itself will do little good.
49.	**Logging and Security analytics, to support forensic exercises as mentioned above.**	Good analytics driven SIEM like splunk or similar is a great tool to assist in this work

The above table presents all the 49 aspects gathered from the brainstorming session in the first round of the Delphi study.

Appendix A1.2 – Consolidated List

The following is the consolidated list that is based on the 49 aspects above, the list resulted in 32 aspects.

1. Use auditing to secure the organization (Internal and External)
2. Design systems for auditing, logging and to facilitate forensic analysis and to clearly demonstrate compliance.
3. Use analytical tools to support forensic activities and to analyze logs.
4. Change management needs to be integral to the organization and changes must be analyzed for risks and effects prior to implementation.
5. Map data flows of classified and sensitive data. Take inventory and create classification of all data regularly. Have policies in place to guide evaluation and mapping of data.
6. Document legality of all data.
7. Documentation of data that are stored or processed.
8. Properly protect and configure systems processing or storing sensitive data (based on the data inventory).
9. Educate all personnel with access to personal data about privacy regulations, both internal and external personnel to raise awareness.
10. Make Information Security an integral part of the organization, and train personnel in the different aspects of Information Security.
11. Validate the effects, usefulness and usability of Information Security prior to implementation.
12. Utilize useful and tested mitigation techniques, (found in OWASP Top 10 and SANS Top 25 vulnerabilities for example), and train personnel accordingly.
13. Protect sensitive data using strong encryption technology.
14. Centralize encryption management to create encryption coherence within organization.
15. Create, use and maintain a specific data incident/breach plan for personal data.

16. Create and implement policies and controls for access to sensitive data.
17. Create and implement policies and guides for resource securing.
18. Information Security Officers with budget and mandate to drive security efforts in the organization.
19. Information Security program sponsored and supported by executive level management.
20. Create and implement procedures describing how to respond and act to requests from data subjects.
21. Map skill and knowledge level of personnel responsible for securing resources, and make access to professional support available.
22. Appoint a professional and validated DPO.
23. Create/select and implement Key Management system. Make it supported by policies and processes.
24. Design according to Privacy by Design as defined in the GDPR.
25. Implement policies and processes for risk analysis and risk management.
26. Define and implement policies and processes regarding responsibility for granting access to sensitive data. Educate and train the person/persons responsible for granting access.
27. The different access levels granted need to be based on a data inventory and classification. Grant access based on role within the organization, and use multi person-control to limit threats from disgruntled personnel and other abuse from a single user.
28. Verifying the identity of users.
29. Document the access rights management procedure and make it auditable.
30. Create and implement policy for password management and maintenance. Make it usable.
31. Standardize secure support functions for common systems.
32. Verify and validated configuration of devices, networks and systems.

Appendix A2 – Round 2 of the Delphi Study

This appendix includes the 10 critical aspects chosen by each participant and the final seven aspects that were chosen based on the simple majority.

Appendix A2.1 – The critical aspects chosen by the participants

Aspect no.	Participant A	Participant B	Participant C	Participant D	Participant E	Total
1		1		1	1	**3**
2			2			**1**
3				3		**1**
4		4	4	4	4	**4**
5	5	5	5	5		**4**
6	6					**1**
7						**0**
8	8		8			**2**
9	9					**1**
10		10	10	10	10	**4**
11						**0**
12						**0**
13			13		13	**2**
14						**0**
15	15	15	15			**3**
16	16		16	16		**3**
17				17		**1**
18		18		18		**2**
19				19	19	**2**
20	20					**1**
21						**0**
22	22	22			22	**3**
23			23		23	**2**
24		24			24	**2**
25	25				25	**2**
26						**0**
27				27		**1**
28		28	28			**2**
29						**0**
30					30	**1**
31		31				**1**
32						**0**

This table presents which aspect each participant chose, then every aspect that was chosen by the simple majority of three or more participants was highlighted in green and the ones highlighted in red were either chosen by a minority or not chosen at all.

Appendix A2.2 – The Final Seven Aspects

Category	Aspect
Audit	Use auditing to secure the organization (Internal and External)
InfoSec Mgmt.	Change management needs to be integral to the organization and changes must be analyzed for risks and effects prior to implementation.
Data Inventory	Map data flows of classified and sensitive data. Take inventory and create classification of all data regularly. Have policies in place to guide evaluation and mapping of data.
Education	Make Information Security an integral part of the organization, and train personnel in the different aspects of Information Security.
Incident Response	Create, use and maintain a specific data incident/breach plan for personal data.
InfoSec Mgmt.	Create and implement policies and controls for access to sensitive data.
InfoSec Mgmt.	Appoint a professional and validated DPO.

This table presents the final seven aspects that were chosen by a simple majority and in which category they belong to.

Appendix A3 – Round 3 of the Delphi Study

This appendix includes the ranking of the aspects based on the first ranking round of round 3 and the experts reasoning for each aspect. It also includes the justification given by some of the participants in the second ranking round.

Appendix A3.1 – The First Ranking Round

Aspect Rank	Aspect	Combined Reasoning from all participants	Mean value
1	Create and implement policies and controls for access to sensitive data.	Controlling data access is key. Without rules no action is defendable, we have to have a foundation to base actions on. First you need to have control over whom have access to sensitive data, forcefully if needed, you need to protect the sensitive data before you can even receive the sensitive data. Good policies are important, and that they are supporting, and not just very secure but difficult to use. I consider a top-down approach to security to be critical – Security begins with a set of values that the organization holds which guides the assignment of value (both financial and intangible) to information resources. This in turn guides the writing of policies regarding Information security. These policies steers IT security to deploy proper protection to the resources.	**1,8**
2	Make Information Security an integral part of the organization, and train personnel in the different aspects of Information Security.	Pretty obviously, a good idea. The rules and processes needed to do the work, a bad foundation will always topple the daily operations in the end, as seen in many security breaches. Like military or rescue processes, trained till you puke, but in the bone marrow when the day comes. The organization needs to have the understanding and culture of handling of sensitive data, without the understanding of information security the rest of the points are mute. Security for personal data, needs to keep the entire organization educated and on their toes to keep all systems secure. If some systems can be compromised, it can spread to the systems that have access to sensitive personal data.	**3,4**

3	Map data flows of classified and sensitive data. Take inventory and create classification of all data regularly. Have policies in place to guide evaluation and mapping of data.	Inventory of data stores and data flows is critical, without it personal data protection makes little sense. We need to know who accessed what and when and why. Data not visible cannot be analyzed but can be found by an auditor, costing 4% in fines. Also, answering, “we had no idea this data existed” is a 4%:er. Therefore, the DPO need to drive the mapping process. When you have an understanding of information security and have control over the access of the data, you also need to map the data flows to be able to maintain access control of the data and this point is also a cornerstone to handle the change management regarding the personal data, you need to take informed decisions.	**3,6**
4	Appoint a professional and validated DPO.	DPO is probably a good idea, but I think some organizations can manage to implement good security even without a dedicated person! I repeat, not having is a sure thing getting 4% fines, it is the one driving above, but also the guardian of the rules. A key in supporting an organization with handling personal data and a requirement where applicable in law. This is an important starting point to create continuity, clear responsibility and skills and knowledge to be successful in the long run.	**3,6**
5	Create, use and maintain a specific data incident/breach plan for personal data.	Managing breaches needs to be fast and consistent. GDPR 4% fines of the global turnover, this demands an incident plan as a priority action, not to pay those 4%. Such plan is also the factual DRP action for personal data. Again, without it, we lack a foundation to control the work. If the worst happens and in a stressful situation there is good practice to already beforehand know how to handle such incidents, instead of starting to look to whom should be doing what and whom to inform.	**4,8**

6	Use auditing to secure the organization (Internal and External)	Important because without auditing it is difficult to “sharpen up” for the future and learn from mistakes. The audit process is the best quality control an IT department can wish for. It is transparent and effective, but it needs a number of prerequisites. The auditing is the instrument of following up that the processes and requirements are met, and also suggestions to improvement in the organization. Goes to maintaining the security and handling of personal data. With the risk to be audited that we keep our knifes sharp. Audit makes sure processes are existing and in use, and makes it possible to improve.	**5,2**
7	Change management needs to be integral to the organization and changes must be analyzed for risks and effects prior to implementation.	Change management is a must in order to not just do things right to begin with but to ensure that it becomes an iterative process. An auditor not finding a change record is to lose an audit remark. Change management is the documentation base for all work done, not working or fragmented, it will leave us wide open for auditors finding flaws, for there will be such. The change system is the foundation of the auditor’s quality control, but in most cases, it need a lengthy refresh and tune-up. To maintain protection of the personal data when changes to the environment and or processes that handle personal data. Without good change management, there is no control of the systems, and this can open up security vulnerabilities that can compromise systems that used to be audited and considered secure.	**5,6**

This table includes the ranking of the aspects based on the first ranking round which had a *W* value of 0.363, and the participants reasoning for each aspect. As seen in the table, aspect 3 and aspect 4 have the same mean rank of 3.6. Secondary ranking done by calculating median values for the aspects tied with regards to the mean ranking value. Aspect 3 had a median of 3 and aspect 4 had a median of 4. The aspect with the lowest median is ranked higher.

Appendix A3.2 – The Reasoning of the Participants in the Second Ranking Round

The table below presents the participants' reasoning behind their choices in the second ranking round.

Participant Number	Ranking Reasoning (Numbering based on result from first ranking round)
1	1. Same 2. Change to 3. I agree that less security but as a part of the corporate culture and backbone, is better than more policy which is not practiced. 3. Same. I cannot see the point in implementing any data security measures unless we understand if we have any sensitive data, where it is, and have classified the data according to sensitivity. If we don't start there it is as if we are assuming that we can put all data in a vault and secure everything. But we know that this is impossible because data is stored in many places, needs to be communicated across different networks and is consumed on different devices. There is no single vault in the digital era. 4. Change to 4. I agree that having someone represent the interests of personal data is a priority. 5. Change to 7. I have revised my view on this a bit. If information security is ingrained in the corporate culture, then this point is really a function of that. It just helps to speed up incident handling. Had security not been an integral part of the company, then the incident plan becomes more critical 6. Same 7. Change to 5. I still feel that if we cannot manage change by getting new resources and systems and data to be covered by the other processes we have covered in the other points, then we risk creating islands of bad practice in an otherwise well-functioning organization.
2	Have reviewed the account from round 3, but my evaluation stands, though the low consensus. One reason for this, might be our differenced backgrounds. Myself with a background not only from security, but as CIO, TIO and info Security/technical security. It seems that the controls are where we have the highest scores or consensus, but that answers about the means to do it, maybe more reflects our experiences.
3	(No response to the second ranking round)
4	(Changes to the list made, but no reasoning provided due to time constraints of the participant)
5	(No changes made)