TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

JARMO MULTANEN

ANALYSIS OF THE GDPR'S EFFECTS ON A MEDICAL APPLICATION

Master of Science Thesis

Examiner: Marko Helenius
Examiner and topic approved on 29.08.2018

ABSTRACT

JARMO MULTANEN: Analysis of the GDPR's effects on a medical application
Tampere University of technology
Master of Science Thesis, 71 pages
October 2018
Master's Degree Programme in Information Technology
Major: Information Security
Examiner: Marko Helenius

Keywords: GDPR, privacy, data protection, information security, General Data Protection Regulation, privacy by design, data concerning health, medical application, ISO/IEC 25010

The European General Data Protection Regulation (GDPR) came into full effect in May 2018 after a two-year transition period. The regulation aims to improve the data protection of the citizens of the European Union. The regulation also affects the rest of the world. Although not all the rules introduced by the GDPR are new, the regulation contains novel requirements both regarding data protection and information security level. One of these new requirements is the right of a natural person to be forgotten in certain circumstances.

The novelty of the GDPR and in some parts the general wording of the rules contained in the regulation may create difficulties in interpretation for the entities that have to conform to the regulation's rules. This thesis examines through the analysis of a medical application, the impact of the regulation on data controllers and software developers dealing with data concerning health. The data protection and information security requirements presented by the GDPR are applied to the analysed application. The application is analysed against the requirements derived from the GDPR with the help of the Software product quality model of the ISO/IEC 25010 standard.

Based on the conducted analysis, the application is in a good state regarding the GDPR even when some changes need to be implemented. At this stage, the impact of the GDPR on applications containing data concerning health is not significant if best practices were used to develop the application. The impact of the GDPR lies more in the general approach to managing risks directed at the software since the content and the amount of personal data should be considered in risk management.

In addition to the analysis of a medical application, this thesis contains an analysis of the previously existing privacy legislations of the United States, Finland and France. The related privacy laws of these countries are compared to the GDPR so that the content and new additions of the new GDPR would be more apparent.

TIIVISTELMÄ

JARMO MULTANEN: Analyysi GDPR:n vaikutuksista lääkinnälliseen sovellukseen
Tampereen teknillinen yliopisto
Diplomityö, 71 sivua
Lokakuu 2018
Tietotekniikan diplomi-insinöörin tutkinto-ohjelma
Pääaine: Tietoturva
Tarkastaja: Marko Helenius

Avainsanat: GDPR, yksityisyys, yleinen tietosuoja-asetus, tietoturva, sisäänrakennettu tietosuoja, terveyttä koskeva henkilötieto, lääketieteellinen sovellus, ISO/IEC 25010

Euroopan yleinen tietosuoja-asetus (GDPR) tuli voimaan toukokuussa 2018 kahden vuoden siirtymäkauden jälkeen. Asetuksen päämääränä on parantaa Euroopan Unionin kansalaisten tietosuojaa yhdenmukaistamalla käytäntöjä ja vaikuttaa samalla myös muuhun maailmaan. Vaikka kaikki asetuksen esittelemät säännöt eivät ole uusia, niin esitys sisältää uudenlaisia vaatimuksia niin tietosuojan kuin tietoturvankin tason suhteen. Yksi näistä uusista vaatimuksista on luonnollisen henkilön oikeus tulla unohdetuksi tietyissä olosuhteissa.

Tietosuoja-asetuksen uutuus ja paikoin yleisluontoinen esitystapa saattavat aiheuttaa tulkintavaikeuksia tahoille, jotka joutuvat muuttamaan toimintatapojaan asetuksen tulon myötä. Tämä diplomityö tutkii lääketieteellisen sovelluksen analyysin kautta sitä, miten asetus vaikuttaa terveyttä koskevia henkilötietoja käsitteleviin rekisterinpitäjiin ja ennen kaikkea sovelluskehittäjiin. Asetuksen tietosuojaan ja tietoturvaan liittyvät vaatimukset käsitellään analysoidun lääketieteellisen sovelluksen kautta. Sovellusta analysoidaan tietosuoja-asetuksesta johdettuja vaatimuksia vastaan käyttäen ISO/IEC 25010 standardin ohjelmistotuotteen laatumallin toiminnallisen sopivuuden piirteiden avulla.

Analyysin perusteella kyseinen sovellus on tietosuoja-asetuksen huomioon ottaen hyvässä tilassa, vaikka muutoksia tarvitseekin tehdä. Tässä vaiheessa tietosuoja-asetuksen vaikutus terveyttä koskevia henkilötietoja käsitteleviin sovelluksiin ei ole suuri, mikäli sovellusta kehitettäessä on käytetty parhaita käytäntöjä. Tietosuoja-asetuksen vaikutus tuntuu enemmän yleisessä lähestymistavassa ohjelmistoon kohdistuvien riskien hallintaan, sillä henkilötietojen sisältö ja määrä tulee ottaa huomioon riskinhallinnassa.

Lääketieteellisen sovelluksen analyysin lisäksi työssä käsitellään jo olemassa olevia tietosuojalakeja Yhdysvalloissa, Suomessa ja Ranskassa. Näiden maiden lainsäädäntöä verrataan uuteen tietosuoja-asetukseen, jotta asetuksen sisältö ja lisäykset vertautuisivat lainsäädäntöjen aikaisempaan tilaan.

PREFACE

I would like to thank the examiner and instructor Marko Helenius for offering much needed advice during the writing process of this thesis, my employer Atostek Oy for giving me the chance to work on this thesis and Jani Heininen for providing guidance related to the writing process. I would also like to thank the company developing the software application analysed in this thesis for giving me the opportunity to analyse the application.

I am grateful for the support of my family, my girlfriend and friends during the writing process of this thesis.

Tampere, 29.10.2018

Jarmo Multanen

CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

CIA	Confidentiality-Integrity-Availability
CNIL	Commission Nationale de l' Informatique et des Libertés (National Commission on Informatics and Liberty)
CSV	Comma-separated values
DMZ	Demilitarised Zone
e-PHI	Electronic Protected Health Information
FTC	Federal Trade Commission
GDPR	The European General Data Protection Regulation
HHS	U.S. Department of Health & Human Services
HIPAA	The Health Insurance Portability and Accountability Act
HITECH	Health Information Technology for Economic and Clinical Health
HTTPS	Hypertext Transfer Protocol Secure
NHS	The National Health Service
OWASP	The Open Web Application Security Project
PbD	Privacy by Design
SQuaRE	System and software Quality Requirements and Evaluation
SSL	Secure Socket Layer
TDE	Transparent Data Encryption
TLS	Transport-Layer Security
VPN	Virtual Private Network

1. INTRODUCTION

Today there is a vast number of different services and devices that impact the lives of ordinary people. Information about individuals is gathered continuously through these services, like the one's Google and Facebook provide and used for numerous purposes, such as marketing [12]. The scope of information gathering creates several situations where there is a possibility of mishandling information and where privacy issues arise. Esteve [12] mentions lack of proper consent for information usage, user's inadequate access to their information and risk of anonymised data becoming personalised as privacy issues arising from the business practices of Google and Facebook.

Botha et al. [4] note that today privacy and information security are essential to the digital economy. Incidents, where individuals' sensitive data is exposed, happen frequently. Botha et al. analysed data breaches made public in 2015 and 2016 and noted that some of the world's largest data breaches happened during those years. Frequent occurrences of data breaches might lead to cynicism and a feeling of futility among individuals in what is described as "privacy fatigue" that Choi et al. [8] further studied. They implied that service providers and governments need to be aware of the effect of the users' privacy fatigue since high privacy fatigue can cause people to become dissatisfied and reluctant to use online services such as social networks. Choi et al. [8] suggest that governments should discuss privacy issues from the consumers' viewpoint and enact better policies since these policies can be a way of increasing privacy protection level. That, in turn, could increase people's trust in privacy protections and make them more engaged with their privacy so that they would follow the best practices related to privacy and information security.

One policy that tries to consider the consumer's perspective was formed in Europe over several years. The General Data Protection Regulation (GDPR) aspires to improve data protection and is aimed to be as all-encompassing as possible. It has been under work for many years in the European Union (EU), and it has come to full effect in 25th of May of 2018. The new regulation tries to unify the way user data is handled in the EU and to force companies from other locations to conform to these new requirements. The regulation applies to all information handling in the EU and forces companies from other locations to conform to the regulation while working inside the EU handling Union citizens' data. The regulation gives new rights to individuals, such as the right to request erasure of personal data and enforces "data protection by design and data protection by default" principles. It also tries to ensure that information security is considered adequately during each step of personal data handling. [13]

The GDPR replaces the previous EU directive from 1995 titled 95/46/EC [13]. While directives are goal setting legislative acts that EU nations must achieve, each nation devises their laws to reach the goal set by the directive [15]. Regulation, however, is binding and is applied outright replacing the corresponding member state law [15]. The regulation mentions that while objectives and principles of the previous directive are still relevant, the technological advances and other developments from the time when the previous directive published have caused new challenges that demand a new regulation [13]. Not all details the GDPR presents are new, however, and have already been applied in member states like Finland and the previous EU directive. For example, the Finnish Personal Data Act already contains some specifications that are present in the regulation, such as an individual's right to be informed on what kind of information a registry keeper has stored of that individual in chapter 6, section 24 [19]. The regulation adds more specifications to the directive it replaces.

The GDPR has been in a transitional period from May 2016, meaning that the member states and companies operating in the EU have had time to comply with the new requirements the regulation brings alongside it. However, the extent and impact of the regulation were not completely clear in the transitional period. This is because the regulation is ambiguous in places. The GDPR is applied as is until a new member state law is prepared to add more precise measures to the GDPR. For example, in Finland, the new national law was not yet ready when the GDPR came into full effect. Missing national guidelines mean that some parts of the GDPR remain open to interpretation with no legal precedents and qualifications. The ambiguousness can provide a challenge for data controllers and processors since it is not completely clear and specific how parts of the regulation affect them, what are the repercussions of failing to comply and is the current state of their information security policies and models up to date.

One specific area of information handling is medicine and software systems containing patient information. In addition to personal data such as social security numbers, medical applications also contain information about the patient’s diagnosis and health. Even before the GDPR, the handling of such information was under strict regulation since patient information is deemed highly sensitive [22]. However, organisations handling patient data are also subject to the GDPR if they operate in the EU, so they must take the regulation into account. Botha et al. [4] analysed data breach related statistics from Privacy Rights Clearinghouse [38] and conclude that attackers have increasingly targeted the health industry in recent years and the health industry’s percentage in the overall amount of data breaches has increased. That is why the potential new improvements in the privacy regulation and the effect of the regulation should not be disregarded. Thus, the research questions are as follows:

- How does the GDPR affect the software application and registry keepers handling health records?

- What do the GDPR's information security rules mean for organisations handling health related personal data?

These questions will be analysed with the help of an example medical software that is affected by the GDPR. Chapter 2 details the results of a literary analysis regarding previous studies related to the topic of this thesis. In Chapter 3 and Chapter 4 the relevant background for privacy and information security concepts is explained. The most relevant articles of the new GDPR are presented in Chapter 5. Chapter 6 describes the previously existing legislations of chosen example countries. In Chapter 7 the example medical software is presented and analysed regarding the requirements derived from the GDPR. The software is analysed with the help of the Software Product Quality Model presented in ISO/IEC standard 25010 [46]. The required and possible changes for that application are also laid out. Chapter 8 contains discussion about the analysis and its limitations. Finally, Chapter 9 contains concluding remarks.

2. RELATED WORK

The General Data Protection Regulation came to effect during the process of this thesis. Even though the details were decided in 2016, the two-year transitional period meant that the regulation was not enforced until 2018. As such, there has been a period where the GDPR could have been studied, and those studies could have had similar topics as this thesis.

This chapter presents the findings of a related work search conducted on several separate occasions during the writing process of this thesis. The primary focus of the searches were scientific articles written about the GDPR concerning medicine or software applications in medicine. The searched publication databases were IEEE Xplore, ACM Digital Library, SpringerLink and ScienceDirect

The details about the database searches are presented in Table 1. Many of the search terms overlapped with each other as many publications were present in several different searches. While the number of results of many used keywords indicates that the GDPR has been the topic of several studies or was mentioned, the number of relevant studies regarding this thesis and the aims of the related work search was meagre. Overall many studies did not go in depth with the regulation, mentioning it briefly or speculating on its impact.

The GDPR was analysed from many viewpoints in the search results, with big data being one of the most analysed topics. The GDPR was included in several papers that analysed healthcare related laws from all over the world. There were also a few papers discussing privacy regulations and how the GDPR will affect data handled on a global scope. Especially genomic data was the focus of several studies. This focus is understandable since one of the most substantial single influence of the GDPR will be big data as colossal amounts of data from several sources are aggregated and analysed all over the world. While anonymised data is out of the GDPR's scope, the regulation is still bound to affect swaths of data that is currently analysed.

Of the studies that were among the search results, none were similar to the topic and scope of this thesis. The newness of the GDPR likely explains the result of the searches. A few studies discussed some related parts of this thesis. Flaumenhaft and Ben-Assuli [21] reviewed the legislations of several countries regarding personal health records and concluded that the international community has not been able to keep up with the developments of the "health information technology". They see the GDPR as seemingly providing the most extensive protection measures but also mentioned that the regulation contains ambiguousness and room left for interpretation in key sections. Shu and Jahankhani [40] analysed how the GDPR affects information governing of the National

Health Service (NHS) England primary care sector. The analysis is not exhaustive and stays on a general level. They discuss how the GDPR adds specifications and makes changes to previous conventions. An example of change is that the individuals are no longer required to pay for first subject access requests, which will increase operational costs of care organisations.

Table 1. *Publication database search results*

Search date	**Database**	**Keywords and the number of results**
28.06.2018	ScienceDirect	GDPR AND medicine, 54 results
02.07.2018	IEEE Xplore	GDPR AND medicine, 0 results GDPR AND health, 8 results
08.08.2018	SpringerLink	GDPR AND health, 132 results GDPR AND doctor, 37 results
08.08.2018	ACM Digital Library	GDPR AND medicine, 0 results GDPR AND health, 0 results GDPR AND wellness, 4 results GDPR AND doctor, 0 results GDPR AND hospital, 0 results
17.08.2018	SpringerLink	GDPR and medicine, 78 results GDPR AND hospital, 51 results
17.08.2018	IEEE Xplore	GDPR AND doctor, 1 result GDPR AND wellness, 0 results GDPR AND hospital, 1 result
17.08.2018	ScienceDirect	GDPR AND doctor, 37 results GDPR AND wellness, 390 results GDPR AND hospital, 63 results

Lopes and Oliveira [28] surveyed the GDPR preparedness of Portuguese health clinics in their study, and among those that answered the survey, only 14 (25% of the surveyed) clinics responded that they had started or concluded their ratification of the measures. Tikkinen-Piri et al. [48] analysed the differences between the GDPR and the EU directive that preceded it. They developed 12 aspects for data-intensive companies to follow so that they would be able to successfully adopt the measures of the GDPR and follow these measures. The measures include reckoning with the sanctions and considering "data protection by design and data protection by default". The detailed measures remain on a general level and do not focus on specific, concrete actions.

3. PRIVACY

In this chapter, the relevant concepts of privacy are laid out. Further specification of these concepts is needed, as the GDPR does not provide concrete definitions for all the concepts it presents and because these concepts are not unequivocal. Privacy is one of the central themes of the regulation which means that understanding privacy and the various notions related to privacy is worthwhile.

3.1 Definitions of privacy

Privacy can be defined differently depending on the viewpoint and context. The definitions of privacy have also understandably developed as technology has advanced. The definition "right to be left alone" was popularised by Warren and Brandeis in 1890 [55]. They had observed that along with an older notion of physical privacy, intellectual and emotional life was also to be protected from unwanted publicity, from "injury of feelings" [55]. While this definition is too non-specific for this discussion, the idea of people wanting to protect private information about themselves is very relevant today.

Another statement that relates to privacy comes from the Charter of Fundamental Rights of the European Union (2012/C 32/02). While the charter does not discuss privacy in exact terms, Article 7 contains the "Right to respect for private and family life" [14]. While a conclusion can be made that privacy is a fundamental right, the statement does not go into details explaining what privacy might mean in this context. Overall, privacy today is a multifaceted concept and as Spiekermann and Cranor [44] noted, people's data is fragmented into several places with difficult traceability whereas old definitions were made at a time when privacy violation would likely be limited to one person. Solove [41] also remarked that privacy is too complicated "to be boiled down to single essence". Solove [41] discusses risk management and how the balance of power in society affect people's privacy. He examines how a person can control what they reveal to others and what they consent in effectively controlling what information is available about them and what should remain in secret. The discussion Solove presents is a reasonable basis for this examination since the GDPR emphasises the individual’s right to control their information.

As can be seen, privacy is a complicated subject with many definitions and aspects. Another term related to privacy is data protection which is integrated into the name of the regulation. Hoffman et al. [23] note that while in the USA the term privacy is prevalent, that in Europe data protection is a more widely used term. Still, both of those terms are used to the same end which is to protect information from the public [23]. The point about

the regional difference in language seems very plausible considering that the GDPR mostly uses the term data protection.

The GDPR does not define data protection either as was the case with privacy. A Dictionary of Computer science defines data protection as a computer-related version of privacy and is defined alongside privacy [11]. Two concepts are introduced in the dictionary: protection of data about a specific individual or entity and protection of data owned by a specific individual or entity. Data protection legislation entry, on the other hand, discusses the individual's rights to find out what data has been stored of them and how legislation determines how different organisations can use the data they have collected [10]. Based on these definitions it seems that while privacy and data protection might be synonyms in some instances, data protection could be a more specific term. Privacy, as Solove [41] noted, has many facets. The fact that the GDPR incorporates the term data protection and because the Euro-centricity of the regulation, the term data protection might be more relevant in this discussion.

The Charter of Fundamental Rights of the European Union goes on after the “Right to respect for private and family life” to explain the protection of personal data in Article 8 [14]. The article details how everyone has the right to the protection of personal data and how the processing must be based on a given consent or to some other legitimate reason. Personal data can be a multifaceted concept too. The GDPR defines personal data as any information related to an identifiable natural person where the person can be identified directly or indirectly [13]. The regulation singles out identifiers such as name, location data and health-related data. The GDPR's definition suggests that anything could be personal information in the right circumstances. Mai [30] concludes that personal data or personal information is a communicative act and that while controlling or restricting access to said information is a means to protect it, the protection should not be limited to that. One should also think about the usage, analysis, and interpretation of personal data. Mai [30] also notes that the meaning of information ties closely to the context and situation. Mai’s note is in line with the GDPR's notion that personal data is not an absolute term.

Individuals create personal data of themselves directly through their actions. The creation of data is also a side product of the actions an individual makes, such as when they log into a service leaving a log file trace of their interaction. Data is also actively being recorded for different purposes and then saved into a storage place. Data can also only be monitored and not stored anywhere. Then when data is stored, a question arises about how and where it is stored, who has access to it and is it being distributed in some way to other stakeholders that in turn process the data to their ends. Concerns might also arise from the purposes of storing personal data.

The GDPR leaves anonymous data out of its scope [13]. Pfitzmann and Hansen [37] define anonymity as when a subject is not identifiable within a set of subjects where the set

is defined as an anonymity set (the set of all possible subjects). Later they add an angle of an attacker into the definition, meaning that an attacker cannot sufficiently distinguish an individual from the anonymity set [37]. The GDPR is concerned only with the protection of personal data and through that consideration, only places requirements on data that could be identifiable. An anonymised dataset does not warrant special protection. The lack of protection for anonymised data leaves a potential gap since it does not seem reasonable to think that data is automatically worthless or harmless without an identifiable factor in it. Complete anonymisation might not be entirely possible anymore and as Tavrov and Chertov [47] conclude in their study, even if identifying attributes are removed from a data set it is still possible using the right algorithms to violate the anonymity of groups in a data set.

Although Tavrov and Chertov [47] discuss group anonymity, it seems reasonable to conclude that individual anonymity could also be violated in an anonymised dataset. It would also seem that the anonymisation depends on the method used to alter a data set. The problem could be that a data set that is supposed to be anonymised does contain information that is linkable to an individual. However, because such information does not need to be protected under the GDPR, it might be out in the open without needed security measures. Anonymisation can also, therefore, be a way of trying to bypass the GDPR. By anonymising data, a controller can claim to have no data that falls under the effect of the GDPR, therefore, staying out of the scope of it. Then again anonymisation of data is an acceptable way of protecting individual's privacy when done right since it strips the data of all identifiable concepts. That is why the anonymisation of data is not automatically a poor way to increase the privacy protection level of the data. Those that anonymise the data need to be aware of the possible pitfalls anonymisation has.

Anonymised data is altered in such a way that no person can be recognised based on that data. Pseudonymity on the other hand, as defined by the GDPR, has a crucial difference with anonymity, since the data is anonymous but additional information is stored separately from it [13]. If the additional information is linked to the data, then it is once again possible to identify the individual through it [13]. The separately stored data must be stored securely so that the data does not become linkable to an individual. Pfitzmann and Hansen [37] present one definition of pseudonymity as the usage of pseudonyms as identifiers. Thus, pseudonymity can be a weaker state of anonymity.

Pfitzmann and Hansen [37] define linkability as the ability of an attacker to sufficiently distinguish if two or more items of interest are related to each other or not. The GDPR itself does not define linkability. Linkability relates to pseudonymity and personal data in general since pseudonymised data is not pseudonymised if the additional information is linked back to the data set where it was removed. The problem with linkability is that linking the data to other datasets might not be apparent. Sometimes datasets that seem harmless by themselves are suddenly categorised as personal data when linked together.

3.2 Privacy by Design

The article 25 of the GDPR presents the requirement for "data protection by design and data protection by default" [13]. For data protection by design, the regulation states that while considering the nature of the data processing, the controller should implement appropriate technical and organisational measures that in turn implement data-protection principles [13]. Also, necessary safeguards need to be integrated into the processing to meet the requirements of the GDPR and protect data subjects' rights [13]. Pseudonymisation is mentioned as a measure and data minimisation as a data-protection principle. Data protection by default ensures that only personal data that is necessary for specific processing is processed and it applies to collecting, processing and storing data [13]. Accessibility is also mentioned, and data protection by default must ensure that an individual's data is not accessed by anyone who does not have the right to access it. The GDPR does not go more into specifics of what "data protection by design or data protection by default" are, but they seem to be central in the regulation.

"Data protection by design and data protection by default" seem to be related to Privacy by Design (PbD) concept. PbD's goal is to embed privacy to technical specifications from the start and not to add it during or after development [6]. Initially, its primary area of application was information technology, but it has since expanded to other areas too. It is meant to be a technology independent framework that tries to maximise the ability to integrate good information practices to the designs and specifications. [6] The main principles of PbD are:

1. Proactive not reactive; Presentative not remedial
2. Privacy as the default
3. Privacy embedded into design
4. Functionality – Positive-sum, not zero-sum
5. End-to-end lifecycle protection
6. Visibility and transparency
7. Respect for users' privacy [6]

In the following sub-chapters, the main principles are detailed, and their meanings analysed. PbD's principles and demands are not perfect, so it is also helpful to analyse counter arguments made against these principles.

3.2.1 Proactive not reactive; Presentative not remedial

The first principle of PbD suggests that actions related to privacy should anticipate and prevents events that may violate privacy before these events happen. With PbD, the objective is not to wait for a risk to materialise, but instead try to prevent the risk from materialising as well as it is possible. Cavoukian et al. mention as an example of this the ability of individuals to review what information has been stored about them. [6]

Bier et al. [2] note that the principle is easy to understand but hard to apply from the developer's standpoint. It is challenging to predict the future, and it might be impossible to build appropriate proactive measures against an issue. As an example of this Bier et al. [2] mention advances in cryptanalysis and how today's best algorithms might be obsolete in the future. Then again PbD only aims for proactivity and not necessarily to a state where every single possible issue could be known beforehand and then prevented, so the principle is not inherently flawed. The example of encryption algorithms becoming obsolete as time goes on is good, but the principle's aim might be more that the system should not be designed so that only one or two fixed algorithms can be used. Instead, for example, the system should be designed for relatively easy switching of the base algorithm.

3.2.2 Privacy as the default

The second principle explains the notion of the default state in and of itself. Cavoukian et al. [6] argue that personal data of individuals should automatically be protected so that no extra action is required from the individual. The reason given is that the users of a system should not need extra effort for their privacy to remain intact.

The second principle too is understandable but has real-world effects that can be complicating. Bier et al. [2] mention that every subsystem or functionality must be so designed that PbD's principles are accounted for. Not only the core functionality should comply with the PbD. This means that new functionality cannot be directly added to a system, but its effects on the principles of the PbD should be analysed also.

3.2.3 Privacy embedded into design

The next principle relates to the previous one through the embedding of privacy into the design. Privacy is then the default state at least in theory. As privacy is in the design, it is not added or bandaged into the system afterwards. Privacy then becomes an integral feature of the system much any other specified functionality would be. [6] Bier et al. [2] question the principle's idea that privacy would not diminish functionality since often the idea of functionality comes before privacy.

While privacy can be integral to the system, it can complicate or prevent specific functionality. Then again if privacy is thought upon when the functionality is designed, it would less likely cause trouble further on. Whether a functionality that is inherently incompatible with privacy is a good idea is another topic in its entirety, but considering PbD's philosophy, such functionality should not be included into a system. An example of functionality that is incompatible with privacy is collecting and publicly sharing tracking data from a smart watch, like the case when it was found out that Polar's Explore global activity map could be used to track specific individuals and even discover secret locations [27].

3.2.4 Functionality – Positive-sum, not zero-sum

PbD's objective is a win-win situation where the arguments of privacy versus availability, a zero-sum approach, would be left behind. Integrating privacy into a system would benefit all and developers would not have to cut corners further along the development regarding privacy. Cavoukian et al. use an example from healthcare where a patient should not have to choose between the functionality of service and privacy. [6]

Bier et al. [2] note that in the real world there usually are such trade-offs. They point out how this zero-sum approach is not always possible, even though PbD suggests that privacy and functionality should always increase hand in hand and another's growth should not diminish another.

PbD's aim to end the functionality versus privacy debate might not be ultimately achieved, but it is still essential to try and minimise the trade-off by taking privacy issues into account early in the development cycle. As with the Polar's case, there sometimes seems to be a trade-off between an individual's privacy and them wanting to adopt some new technology into their lives. In Polar's case the ability to track and share the routes the users of the application had gone through, that data could be used freely by everyone else too, possibly to nefarious ends. So, the users must choose between the possible and perceived benefits of a technological service and their privacy thereby making it a zero-sum game. Choi et al. [8] analysed privacy fatigue and mentioned how a high level of privacy fatigue could prevent people from using certain services. In that way, the zero-sum game may prevent new technologies from being adopted more widely if their privacy related attributes are not up to standards. It seems that it would be in the long run beneficial for the companies to integrate privacy into their services and applications, which is what PbD tries to achieve.

3.2.5 End-to-end lifecycle protection

Through the fifth principle, PbD tries to ensure that personal data is appropriately handled throughout its lifecycle, from the collecting to the destroying. Cavoukian et al. also mention proper log data files increase flexibility in implementation. [6]

Ensuring privacy depends on adequate information security mechanisms and these must go hand in hand [2]. Bier et al. [2] also note that measuring complex systems concerning their security is difficult since, in addition to useful information security mechanisms, protocol implementation and attacker models also need to be considered. Also, a human factor comes into play as roles and responsibilities need to be assigned to people [2]. It does seem that there are several challenges to end-to-end lifecycle protection that are difficult for one entity to control and think of beforehand. While technical solutions are relatively easy to measure, aspects like the human users of the software systems can be difficult to predict.

3.2.6 Visibility and transparency

The sixth principle seeks to increase transparency so that every stakeholder would know what is goes on with their personal data. It also helps individuals to find out if their data is handled as it should and that the organisations are following the rules. User confidence will likely rise because of transparency. As a healthcare example, a patient should be able to know what information is collected, how the information is used and who can access the information. [6]

According to Bier et al. [2] audits, notifications and information are means to achieve the goals of this principle. They also mention potential conflicts between privacy requirements, like with transparency and unlinkability. Different privacy requirements do not always exist without conflicts with each other.

3.2.7 Respect for users' privacy

The last principle is straightforward and more of a reminder of PbD's goal. The respect for privacy should be an essential interest to software handlers [6]. The other six principles are more standalone requirements whereas this principle is an overarching convention.

Privacy features should be easy to use and user-centric for them to work properly [2]. Data minimisation should be the goal since the user should retain their information self-determination. On the other hand, as Bier et al. [2] point out complete data voidance where no data is stored as a default and it is an unchangeable setting, robs the user from their self-determination. Therefore, a middle road approach should be found, and Bier et al. [2] present data minimisation as the middle road.

3.3 Privacy by Design criticism

PbD's definition provides valuable information since the GDPR's definitions of the "data protection by design and data protection by default" do not go into specifics apart from mentioning pseudonymity and data minimisation. From a legal standpoint this is understandable as the GDPR tries to be technology independent and applicable and as Tsormpatzoudi et al. [49] note, the way the GDPR is worded is flexible because of necessity so that concrete measures can be accommodated for specific cases. Tsormpatzoudi et al. [49] call the GDPR's two concepts more comprehensive than PbD's. Even though PbD is not included in the GDPR word to word, it is still the basis of the regulation's notion of "data protection by design and data protection by default".

PbD is not without issues as can be seen from the counterexamples Bier et al. [2] present to each principle criticising the consistency of PbD as was discussed when the principles were analysed. Tsormpatzoudi et al. [49] and Koops and Leenes [25] on the other hand

discuss challenges that PbD's implementation into the legislative measures will bring. While both articles were written when the GDPR was only a draft version, the core of their criticisms is still relevant today as the GDPR has been in effect for relatively little time with not much time for legal precedents yet. Tsormpatzoudi et al. [49] discuss challenges arising from the wording in the GDPR, legal compliance in implementation, difficulties of understanding between principles and the role of the data protection officer. As a further future challenge, they introduce the involvement of stakeholders that are not from the organisations of the data processors that implement the measures introduced in the GDPR and how those stakeholders may also need to be educated about PbD or GDPR's privacy by design and by default [49].

Koops and Leenes [25] argue that PbD should not be interpreted so that technologies or coding is the only acceptable solution for complying with the regulation. Instead, communication strategies should be thought of, and mindsets of the designers and developers influenced [25]. They argue that techno-oriented implementation holds too many problems in it such as contradictions with the rest of the regulation and the difficulty of defining the scope of data protection requirements [25]. System developers should not try to integrate as many data protection measures as they can, but instead, organisational measures would be more fitting [25]. The GDPR has included minimisation in the final version of the regulation, so Koops' and Leenes' aims came at least partially true. Their argument about the interpretation of PbD is reasonable since usually there are not any universal technical solutions that could guarantee compliance. Fortunately, the GDPR's final version tries to be technology neutral and emphasises appropriate solutions and measures depending on the situation. Koops and Leenes do not entirely rule out technology related solutions, but they want to emphasise that technological solutions are not the only way of complying with the regulation.

It seems that the openness of the GDPR and in part PbD has two consequences. On the other hand, it ensures that the regulation is not too specific so that it does not rule out legal cases where the regulation should be applicable. However, on the other hand, it may confuse those that need to follow the regulation and make adhering to the regulation unnecessarily tricky and defeating the aim and purpose of the regulation. As Liebwald [26] discusses, legal language has specific challenges that it needs to deal with. These challenges are the need to build general norms using abstract language and the distance that exists between the general ruling and a legal decision taken in an individual legal case. As language itself is imprecise, there can never exist a maximum precision in the legislative text [26]. Also, there is an added vagueness in legal text that sometimes the courts must interpret and possibly substitute for the legislator [26]. Liebwald lists several reasons why vagueness might be added into legislation: covering future circumstances that are not entirely predictable, covering the typical cases, leaving room for more specific rules and interpretation, or that genuine political willingness or consent are lacking.

Vagueness in law can be considered a good thing since it moves influence from the legislator to the courts. Legislators can be perceived to be fickle and more affected by concepts such as party politics. Also, it can be perceived that there is no reason to question the independence of a judge. Still, if there exists too much vagueness in the legal system, the separation of power between different government entities might become blurred, and the laws lose their verifiability and predictability. [26]

In the GDPR's case, many of the reasons for vagueness presented by Liebwald [26] seem to apply. As a Union regulation, it tries to leave room for the national more specific laws while trying to consider the technological advancement of the future. The more specific ways of complying with the regulation in different circumstances might be better decided in national courts with the GDPR being framework in which the decision is made.

It is not an easy task to implement something like PbD into widely used legislation, and careful thought must be put into how it works in practice as PbD, too, can cause unintended damaging consequences if worded wrongly. Then again, such a paradigm shift in the way that applications and system are designed and how people's personal data is used is bound to cause issues. An appropriate question would be that is individual's personal information so invaluable, that inconveniences to the developers and designers weigh more heavily? Of course, real-world issues and facts demand that concept solutions need to be thought upon and modified when adopting concepts such as PbD into laws.

4. INFORMATION SECURITY

Security is a term that people use in several ways in everyday language. It relates to a multitude of concepts, and one of them is information. Information security along with data protection is at the heart of the GDPR, so information security and the related concepts are analysed in this chapter.

4.1 A definition of information security

The GDPR defines network and information security as "the ability of a network or an information system to resist, at a given level of confidence, accidental events or unlawful or malicious actions that compromise the availability, authenticity, integrity and confidentiality of stored or transmitted personal data" [13]. Three of the concepts mentioned in the GDPR, confidentiality, integrity, and availability form a so-called CIA definition that is argued to be the most used information security definition in the literature [29]. Figure 1. shows the relationships between confidentiality, integrity and availability and how these three concepts form security.

Confidentiality tries to make sure that only those with the proper rights and privileges have access to protected data [36, p. 10][46]. Access includes but is not limited to viewing, reading and knowing [36, p.10]. Confidentiality is then breached when someone who is not allowed to can access the information. An example of a confidentiality breach is when some malicious individual infiltrates a computer system and steals sensitive data from it.

Information has integrity when an authorised party can only access or modify an asset in authorised ways [46]. The modification includes writing, changing (status), deleting and creating data [36, p. 10]. Integrity as a protective measure is not only limited to preventing unauthorised modifications by a user but also concerns itself with situations where data changes due to an error or a failure. Database corruption and information loss while it is transmitted from a location to another are also examples of integrity issues.

Availability means that authorised users can access information without interference and receive it as it was supposed to be received [36, p. 10]. The prevention of access should not occur in case of legitimate access. Pfleeger and Pfleeger [36, p. 12] list characteristics of available information as the timely response to a request, the requesters are equal, the system is fault tolerant so that information is not lost in case of a failure, the system can be used as it was intended, and that concurrency is controlled. The ISO/IEC 25010 standard's [46] Software product quality model does not place availability into security characteristics but to reliability characteristics. Nevertheless, the definition in the standard for availability is like Pfleeger and Pfleeger's definition.

Sometimes more properties are added to the three laid out ones of the CIA definition [29], as the GDPR does by including authenticity [13]. ISO/IEC 25010 Software product quality model also includes authenticity, as well as non-repudiation and accountability to measured security sub-characteristics [46]. Authenticity can be defined as the process where a user's identity is verified to be the one that is claimed before they are granted access to services [46]. A general example of enforcing authenticity is asking a user to input their password before letting them log on to a service.

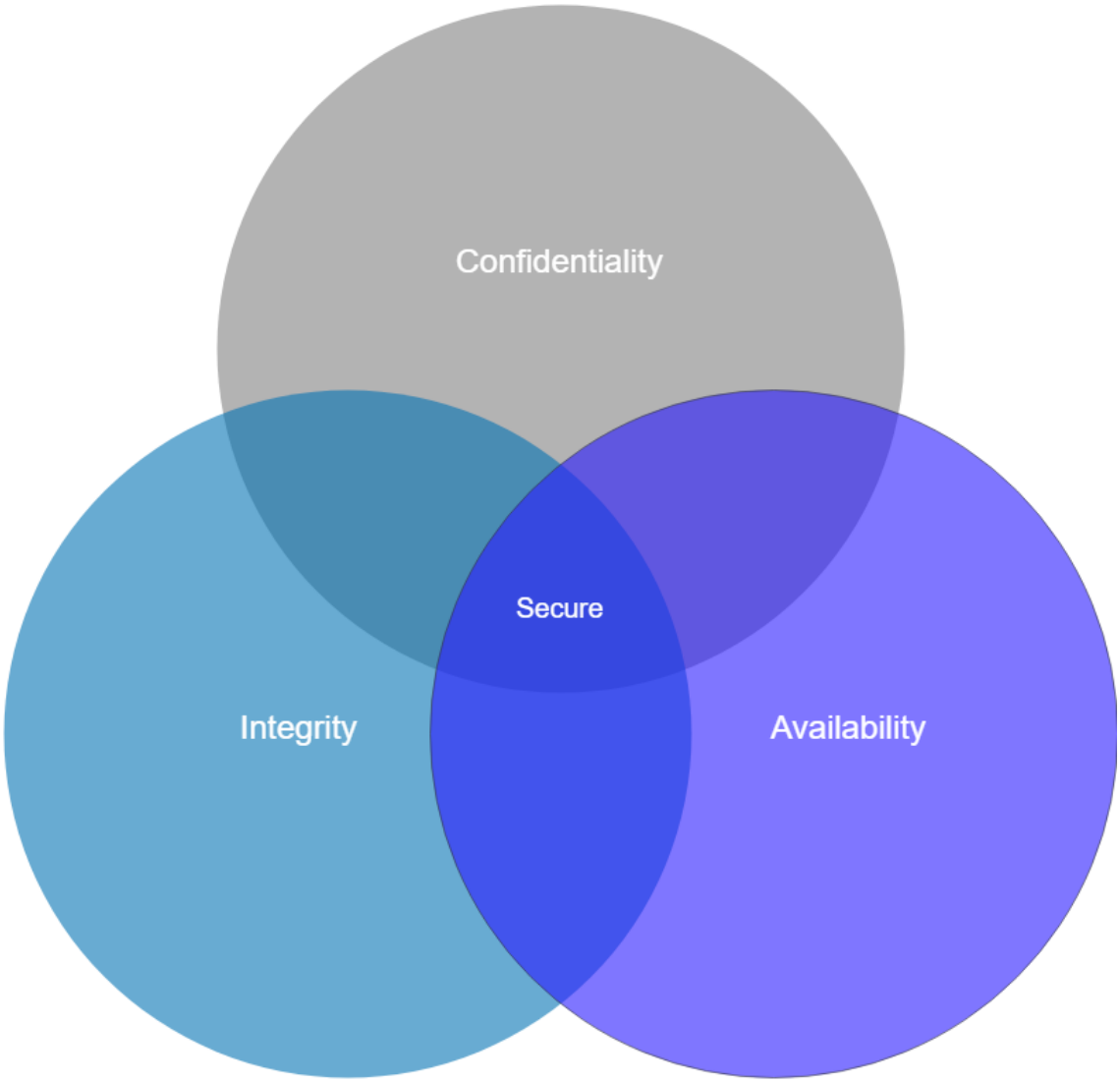

Figure 1. *Relationships of CIA concepts [36, p. 11]*

Authenticity adds to the CIA definition a very concrete measure of ascertaining that the entity accessing a piece of information is whom they say they are. As confirming a data

subject's identity is one of the GDPR's requirements, it is no wonder that authenticity is included in the wording of the regulation. Accountability and non-repudiation share similar goals, as non-repudiation is defined as the ability to prove that an event has taken place without the possibility to repudiate it later and accountability is defined as ensuring that entity's actions can be traced uniquely to the entity [46].

The exact definitions of CIA concepts may differ based on the source [29]. Lundgren and Möller present that the CIA definition is itself too narrow and while the definition is a fitting way to analyse security, information security should not be defined through the CIA definition. Sometimes the three concepts also contradict each other [29][36, p. 10]. It seems that the definition of information security, like the definition of privacy, is hard to pin down. The GDPR itself does use a variant of the CIA definition, so it is an appropriate definition to use here with added attributes. The popularity of the CIA definition is also its merit.

4.2 Data at rest, in motion and in use

Data that is protected can be in an inactive state in a file system, transmitted from one place to another or currently in use in some context. Respectively the terms data at rest, in motion and in use are used to describe these states. That is why different protection measures need to be applied so that the data's confidentiality, integrity and availability, among other properties, is guaranteed. While data is in motion or a transmission state, confidentiality means that an unauthorised person cannot read the data. Integrity means in this case that data cannot be modified or falsified by an unauthorised user [54, p. 2]. Availability then means that the transmitted data is available to those that are authorised, and they receive it as it should be. Authenticity is used to define that legitimate access.

While data is at rest or in a storage state confidentiality means that no authorised user can access it through network and integrity means that the data stored cannot be modified or falsified by an unauthorised user through a network [54, p. 2]. Physically accessing the stored data could also be added here even though someone physically accessing the space the storage device is in is likely lower than accessing the data through a network from anywhere in the world. Availability and authenticity mean virtually the same here as in motion.

Data in use state is defined as data that is in device memory, so the data has been recently or is currently manipulated [45]. While the data is usually loaded to memory through legitimate actions, protections should be still placed so that CIA and authenticity are held for data in memory too. As an example of the need to protect data in use, Stirparo et al. [45] analysed data in use leakages in the memory of Android smartphones, and they found that many applications leave sensitive data into the device memory and do not appropriately protect it. The result shows that along with protecting data in motion and at rest, attention should also be paid to secure data in the memory of the devices.

As can be seen, the CIA definition's attributes can theoretically be guaranteed similarly even though the data is in a different state. The exact measures of doing so change and several solutions for this have been developed over the years. For data at rest, the primary method used is encryption, especially when thinking about the potential theft of the device where the data is located [1, p. 75]. Physical security also relates to data at rest, so making it difficult to physically access the areas, where the data storage devices are, is essential. [1, p. 76]

For data in motion, there is a need to protect the data itself and the connection through which the data travels. For the data itself, there exist several secure versions of transmitting protocols, like SSL/TLS (Secure Socket Layer/Transport-Level Security) which can be used appropriately to ensure that the data is transmitted securely. As for the connection, a virtual private network (VPN) connection can be constructed so that the whole network traffic is encrypted. [1, p. 77]

As for the data in use part, the measures are more limited since the data is accessed by those who have legitimate access to it. [1, p. 78] Buffer overflows are a typical example of trying to exploit data that is stored in memory. In a buffer overflow, the software accesses a part of memory that otherwise not reserved for it. Buffer overflow is achieved by using an array reference to read or write to a location before or after the array. Through the buffer overflow, sensitive data such as old passwords left in memory after processing could be accessed, violating confidentiality. Integrity and availability can also be violated if data is corrupted or changed. These overflows can be prevented several measures, including programming language choices and verifying that accesses are within bounds in the program code. [3]

Access control is needed to help ensure the confidentiality, integrity, availability and authenticity. In information security, access control is a fundamental part of ensuring that objects are only accessed by those who should have access to them [36, p. 109]. Even though access control is fundamental, it is hard to implement correctly and extensively. The Open Web Application Security Project (OWASP) [34] mention broken access control in their 2017 Top 10 Application Security Risks listing as one of the most common risks applications face. The GDPR does not explicitly mention access control, but it is safe to say that it is an integral part of a software system especially since the risks of exploiting a poor implementation are high. Related to access control is the concept of least privilege. As defined by Saltzer and Schroeder [39], every user, as well as a program, should only have the least amount of privileges to complete a task. Having a few privileges limits the potential damage caused due to an error, accident or a deliberate attempt to misuse a system.

As the processor of the data needs to be able to prove that they are complying with the GDPR and that they have acted with the compliance of the regulation, it is useful to document activities in a software system. Log keeping and the audit trail can help with that

and logging is mentioned as a responsibility of those entities that process or control the data [13]. An action that has an impact on security can be minor like an individual accessing a file or major, like a change in an access control change affecting the whole database [36, p. 272]. Accountability of actions is enforced by logging these security-related events into a log that lists the event and who caused the event. This logging procedure forms an audit log which must be kept secure from unauthorised access. [36, p. 272]

The problem of audit logs is that they can grow too large if every instance of every event is logged. In addition to the issue of volume, the analysis of the log would become too cumbersome if the log is too big. That is why the events that require logging should be carefully decided. [36, p. 272] Regulatory measures, for example, can dictate what should be stored and what not. Audit log and can also be reduced, so that the log itself contains only the major events and more insignificant logging data is stored elsewhere [36, p. 273].

4.3 Risk analysis

No software can be completely secure. Attempting to combat every single possible threat whether it is an error, fault or adversary would be too resource consuming to try. That is why different threats and possibilities need to be assessed somehow and then try to protect the software from these perceived threats that are relevant.

Although there are several different definitions for risk, an information security-oriented definition is suitable in this case. Wheeler [56, p. 23] defines risk from an information security standpoint as “the probable frequency and probable magnitude of future loss of confidentiality, integrity, availability, or accountability”. A risk has both the probability of it materialising and the effect that it causes when it materialises. The goal of risk management is to maximise the organisation’s effectiveness while at the same time minimising the chance of adverse outcomes or incidents [56, p. 24]. The goal is not to erase every risk, but to prioritise the most important ones systematically so that the critical risks will not go unnoticed.

The general workflow of risk management is shown in Figure 2. Risk analysis and management is a cyclical process, and well-established risk management frameworks use this type of lifecycle approach [56, p. 46]. The risk assessment stage of this workflow contains risk analysis where the risk is measured by its likelihood and severity [56, p. 47]. As can be seen from Figure 2, several different people and roles take part in the process. Also, the responsibilities should be shared since the security function is merely helping and guiding, while the business owner is the one who owns the risk [56, p. 47].

The risks that were identified and analysed in step 2 will be evaluated in step 3, where the newly analysed risks are also compared against possible previous ones to form prioritisation between them [56, p. 47]. The decisions should be documented as seen in step 4. In

step 5 the mitigations measures are decided. Not all risks can be eliminated, so sometimes exceptions must be made [56, p. 47].

After mitigation, the developed measures must be validated against the real world to ensure that the reduction in risk is achieved [56, p. 47]. Sometimes the theoretically sound mitigation means do not work when they are implemented or end up increasing another risk while mitigating another. Wheeler [56, p. 48] presents an example of this problem where the increase of logging level on servers to provide more accurate information about potential unauthorised activity might start to consume too many resources and slow down the system.

The last stage is the monitoring and audit stage the resources and risks related to it are monitored. If there are any significant changes regarding the risks or an agreed amount of time has passed the risk management process is started again from the profiling stage. After the monitoring and audit stage, the next cycle of risk management can begin when needed. [56, p. 48] This process is continuous since new risks present themselves, and the magnitudes of old risks can increase or decrease as time goes on.

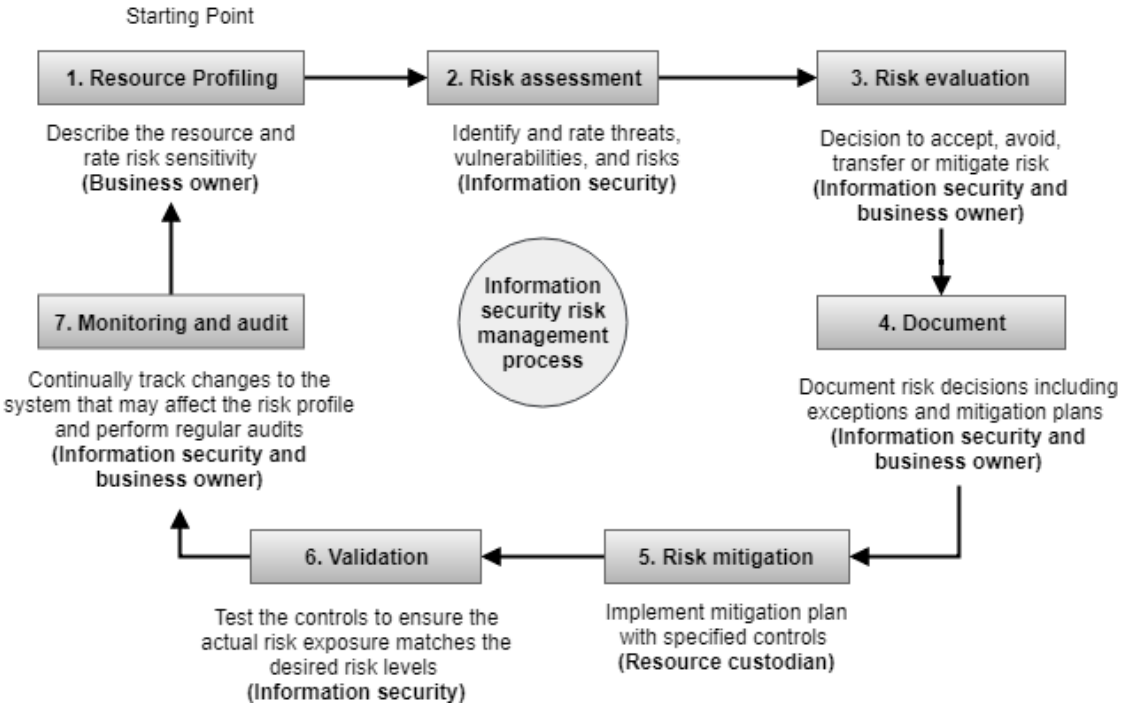

Figure 2. *Information security risk management workflow [56, p. 46]*

The GDPR discusses risk in several articles and sections. In general, the regulation's approach is risk-based where the suitability of different data protection and information security measures are depended on the perceived risk. The likelihood and severity of risks for the rights and freedoms of natural persons need to be considered, and the evaluation process should be updated and reviewed when necessary [13]. As such, the GDPR talks about risks and risk-assessment process similarly to previously existing literature. The

GDPR mentions situations where the risk can be almost automatically considered to be high, such as when the processing is done on a large scale where many natural persons would be affected or when the data concerns children [13].

The regulation also discusses data protection impact assessments, that controller shall carry out before processing the data if the risk for the rights and freedoms of natural persons is considered high. The controller can ask assistance for this from a supervisory authority. In short, the impact assessment should contain why, what and how information is processed, what is the result of risk assessment and what are the risk mitigation means that will be applied to lessen the impact of the risks. [13] As such, the data protection impact assessment seems to be a broadened risk-assessment where the controller needs to specify and think about why they process pieces of specific information. Since risk management is and should be a part of organisations way of operating, the GDPR does not add large amounts of new required tasks for the controller. It is clear that the makers of the GDPR want the controllers and processors to stop and think about why they are processing data and how the data is used. The need for data minimisation presents itself clearly when the controller cannot adequately explain why a data point is needed and the individual's risk of some specific data about them being misused or unnecessarily processed is mitigated.

The theme of risk is indeed a central topic in the regulation, as it also is in the information security space. The regulation passes the burden of defining the appropriate measures to the data processors and defines the framework on how that definition should be done. The GDPR attempts to pressure data processors to get their risk assessment routines in order.

5. GENERAL DATA PROTECTION REGULATION

In this chapter, the focus points of the General Data Protection Regulation are laid out. The European General Data Protection Regulation was finalised in May 2016 and the transition period ended on May 25th in 2018. The basic principles of the regulation are laid out in article 5 of the regulation as lawfulness, fairness and transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality and accountability. Several key terms are introduced in the regulation. The most relevant of these concerning this thesis are:

- data subject: A natural person
- processing: Any operation that is performed on personal data
- data controller: natural or a legal person, public authority or other instance that alone or with others specifies the reasons and means of processing personal data
- data processor: An instance that on behalf of the data controller, processes the personal data [13]

A risk-based manner of approach has been adopted for assessing if the implemented security measures are enough regarding the nature or the amount of the data being stored and processed [13]. Through a risk analysis a data controller can find out the appropriate technical and organisational requirements needed in that specific case. Not all security measures apply in every situation which is valid for information security in general. The data controller is also responsible for informing their supervisory authority of a data breach without undue delay and possibly also the data subjects that were affected by the breach. The data controller has a reverse proving requirement: the controller needs to be able to document and, when needed, present how they handle the information they process. [13]

The GDPR defines two distinct ways of information gathering from individuals. Either the gathering is legislation based which means the controller and processor has a requirement in a member state law or union law to gather information about individuals or the gathering is based on a consent asked from the data subject. The consent must be gained through activity by the data subject, so the subject must opt-in rather than opt-out for the processing of their information. Silence or inactivity are therefore not acceptable means of getting the consent form a data subject. The data subject has a right to withdraw the given consent. The data subject has also gained several rights that they can exercise. In general, these rights must not conflict with the rights of other data subjects or with other member state laws. [13]

While the regulation overrides the previous EU directive and in turn the national laws that have been based on that directive, it is not all-encompassing. As member states have had

their laws considering for example patient data, the GDPR leaves room for further limitations [13]. These limitations must not prevent the free flow of data where applicable, but data concerning health is an area where member state law or future union laws can introduce further restrictions [13]. The possibility of member states to further specify and expand the basis which the GDPR has created means that while the regulation may, in the beginning, overwrite some previous national laws, the member states can bring their previous regulatory measures back if they do not conflict with the GDPR.

Although the GDPR is an EU regulation, the scope of the regulation has been broadened to all controllers and processors even if they are not established in the EU if the processing activities relate to offering goods or services to data subjects inside the EU. The same applies to monitoring the behaviour of data subjects when that behaviour takes place inside the EU. [13] That way the regulation affects large parts of the world and has several implications for business practices even for companies based outside of the European Union. In addition to service providing activities, the companies that analyse data that has come from EU citizens are also subject to the GDPR.

The regulation defines supervisory authority as an independent public authority established by a member state. There also exists a concept of the supervisory authority concerned, meaning the authority which is concerned by specific processing of personal data because the controller or processor is established on the territory of that authority, data subjects in that territory are substantially affected by the processing, or a complaint has been lodged with the authority. In the regulation, Article 51 details the specifics of the supervisory authority. Each member state should have at least one supervisory authority to monitor the application of the GDPR. [13]

The GDPR details several kinds of consequences of not following the regulation. Every data subject has the right to lodge a complaint to a supervisory authority and even to take judicial measures against that authority if they do not handle their complaint in due time. Same applies to the controller or the processor. Data subjects can also receive compensation from the controller or the processor if they have suffered material or non-material damages as the results of an infringement of the GDPR. [13]

Also, the supervisory authorities can impose administrative fines on data controllers or processors. The regulation lists several factors that the authority should consider before deciding on the fine, such as the nature of the infringement or the actions taken by the controller or processor to mitigate damages. As to the size of the fines, the regulation mentions three categories of fine sizes. The choice between them depends on which article of the regulation has been infringed. The first group of provisions is the lesser one of the three and includes the infringement of the obligations of the controller or the processor, certification body or the monitoring body. The size of the fines is up to 10 000 000 EUR or up to 2% of the total worldwide annual turnover of the previous financial year, whichever is larger. [13]

In two of the three categories, the fines are up to 20 000 000 EUR or 4% of the turnover depending on which is larger. These groupings consider more severe infringements that violate, for example, the basic principles of the regulation and the data subjects' rights or if the controller or processor is not complying with an order from the supervisory authority. The regulation also orders the member states to abide by these measures and without delay inform the Commission of the laws and changes to said laws that enforce these penalties in the member states. [13]

5.1 Data subject's rights

Under the GDPR the data subject has several rights which apply in most scenarios with exceptions regarding individual circumstances. These rights are detailed in the Chapter III articles 12-21. The regulation emphasises the openness of information processing and data collector must provide information about how the data subject's information is processed in a compact, transparent, quickly understood and widely used manner. The GDPR also sets a deadline for fulfilling the data subject's request: without undue delays and at a maximum one month from the request. [13]

The data subject has a right to know if their data is processed and if it is, then a right to access their data that is processed and stored. Article 15 outlines this right. Also, the following information has to be sent to the data subject: reason for processing data, which categories the data falls into, recipients of the data if it has been or will be disclosed to another party, the period for which the data is being stored if possible, right to request erasure of the data, right to lodge a complaint, source of the information if it was not collected directly from the subject and the existence of automated decision making. Right to obtain the copy of personal information must not conflict with the rights and freedoms of other data subjects. [13]

Article 16 details the right of the data subject to request the correction of their inaccurate personal data and to have possible incomplete data completed. Article 17 introduces the right of erasure, meaning that the data subject has the right to ask for their data to be removed entirely from the data controller's registry without undue delay. The data controller is obligated to comply with the request if one of the following grounds is applicable:

- the data is no longer necessary for the purposes it was collected,
- the data subject withdraws their consent for the processing of the data, and there are no legal grounds for processing the data,
- the data subject objects to processing explained in Article 21(1) or Article 21(2), and there are no legitimate grounds to process the data,
- data has been processed unlawfully,
- the data is going to be erased due to legal obligation from a member state or the union, or

- the reason why personal data was collected were information society services referred to in Article 8(1) [13]

If the erasure is applicable, then the controller must take the necessary steps to inform other controllers processing the data that its deletion has been requested. The right of erasure is not applicable in following special situations, where the processing is necessary:

- exercising the right of freedom of expression and information
- there is a legal obligation from the member state law or from the Union to process the data
- there is a public interest to process the data
- the controller exercises their official authority
- reasons of interest in the public health area following points h and i of Article 9(2) and Article 9(3)
- the data is used for archiving purposes in areas of public interest, scientific research purposes, historical research purposes or statistical purposes and the right of erasure severely impairs achieving the objectives of the processing
- the data is used concerning legal claims [13]

Article 18 details the right to a restriction of processing. If the correctness of the data is questioned, processing is deemed unlawful, or the controller no longer needs the data for the purposes it was collected for, then the data subject has the right to ask for their data processing to be restricted. The controller can still store this restricted data, but all processing is ceased. There are exceptions to this. For example, the processor can analyse the data if it is needed for a legal claim. [13]

The data subject has a right to data portability which is presented in article 20. Data portability means that the data subject has the right to obtain the data a controller has on them in a commonly used machine-readable format and to send that data to another controller. If possible, the controller should send the data directly to that another controller without having to send it to the subject. The collection of data must be based on consent and processing must be automated for this right to be applicable. If the processing is carried out in public interest or controller exercises their official authority, then this right does not apply. [13]

The last right presented by the GDPR is the right to object in Article 21. The data subject can at any time object to the processing of their data, including profiling. Without a compelling reason or without an existing legal requirement, the controller must stop processing the subject's information. In the regulation, this right is mainly directed towards direct marketing. In case the personal data is processed for scientific, historical or statistical reasons, the right to object is still relevant unless these tasks are carried out in the name of public interest. [13]

5.2 The GDPR and information security

In addition to privacy improvements, the GDPR also guides data controllers on information security. These guidelines are laid out in more detail in part 83 of the introduction chapter of the legislation and article 32. The GDPR requires the controller and the processor to ensure through risk analysis that their information security measures are in order. The risk concerning personal information is higher when the data processor is processing large amounts of data or if the processing involves data that merits specific protection. Such data is particularly sensitive, and the regulation lists examples such as ethnicity, religion and data concerning children. Regulations related to information security are non-specific and technology independent on purpose since the protection should be technology independent and applicable in many different situations. [13]

The data controller must ensure that the subjects' data is protected during each step from the point the data is gathered through the time the data is in storage or being processed and finally ensure that the data is safely deleted after it is no longer being used. [13] These requirements effectively seem to mean data protection during transit, at rest and in use.

While the GDPR remains general in the information security specifications, it still suggests ways to increase the level of security concerning the results of risk analysis. These measures include encryption, pseudonymisation of the data, enforcing the attributes presented by the CIA definition, the resilience of the systems, keeping necessary backups, testing and evaluating the software and organisational measures [13]. The appropriate level of security is formed based on the risks presented by the processing of the data [13]. At a glance, all these measures seem logical and concur with many of the best practices presented in the literature. Legislative measures can compel developers and designers to take these measures into account better when designing software and that way help in overall preparedness. While some more direct measures aimed at improving information security can exist on a national level, for example in official guidelines, a more general approach is fitting when designing regulation for the whole continent.

6. NATIONAL PRIVACY LAWS

To get a better view of the effects of the GDPR, a closer examination of national privacy laws is taken. Three example countries and their legislations related to individual privacy and health information were chosen. Two of them, Finland and France, are EU member states, and the third is The United States of America. The GDPR does not directly affect the USA but comparing two European legislations to legislations outside Europe shows the differences and potential markets for different software products. The comparison also highlights the differences between the GDPR and previously existing legislations. The analysis of the laws of Finland and France pertained to the legislations made before the GDPR. The regulation has changed or will change these laws so that they are compliant with the regulation.

6.1 The United States of America

The USA differs in its privacy-related regulations from many other industrialised nations. Solove and Hartzog [43] describe the current laws as "a hodgepodge of various constitutional protections, federal and state statutes, torts, regulatory rules, and treaties". Whereas in other countries privacy laws are more all-encompassing, in the US different laws regulate different industries [43]. Esteve [12] points out that the fragmented nature of the legislative framework of the US makes the legislation harder for European scholars to understand.

The sectoral approach in the US leaves gaps in the overall regulation notably on the federal level [43]. These gaps are still regulated though by privacy policies that companies have, and the Federal Trade Commission (FTC) enforces these policies. The FTC can act on perceived breaching of a promise made in these policies or if some practice is deemed unfair or deceptive [43]. These measures are declared in Section 5 of the Federal Trade Commission Act [17]. The FTC can sanction companies from wrongdoings but does so rarely, and many cases are settled outside of the courts [43]. The fines given by the FTC are perceived to be small. Solove and Hartzog [43] also note that FTC can influence companies through fear since their auditing process is long and extensive.

As Solove and Hartzog [43] point out, the enforcement of these regulations by the FTC has given it a sprawling jurisdiction, and currently, the FTC has more territory regarding privacy than any other agency. The largeness of its jurisdiction makes FTC the primary source of regulation in several instances where companies are not in the domain of other specific privacy laws. As for the reasons this expansion has happened, Solove and Hartzog give two reasons: FTC's jurisdiction has broadened, and FTC's enforcement frame-

work has been a good fit for self-regulatory attitudes of policy-makers. [43] Self-regulation seems to be prevalent in the US with companies enacting their privacy policies and FTC making sure that companies abide by their privacy policies [12].

Solove [42] discusses the problems that the idea of privacy self-management causes in the US. Privacy self-management is an idea that an individual should self-manage what information is available of them. Solove [42] points out that privacy self-management is challenging in practice because of several reasons:

- the individual might not be informed enough to able to make those decisions,
- the individual's behaviour is affected mainly by how the question of privacy is framed and based on the background knowledge a person has,
- the individual has a problem with the sheer amount of entities that collect personal data,
- the individual cannot know how even smaller pieces of data can be aggregated in the future,
- the individual is unable to accurately assess the harm that sharing information might bring alongside it.

Privacy self-management also relates to the concept of consent. Solove [42] argues that consent is an easy concept to hide behind since consent can legitimise almost any kind of information collection. Esteve reminds that even though a choice is given to an individual on how they can use their personal information and that is enforced by the FTC; there is no requirement from the US law and no detailed rules limiting the data collection that companies can do [12].

Privacy self-management as an idea seems to be the antithesis of what the GDPR is. Solove's arguments about the concept of consent being problematic are valid, since if anything can be legitimised, then every practice could be justified by asking for a person's consent. That person might not know all they should and might even be deceived by framing the consent form or a privacy policy so that it is complicated for an individual to know what they are consenting to. Of course, the US is not a singular place concerning legislation, but a mix of individual responsibility and federal laws, so many examples cannot be strictly generalised. It does seem however that the splintered kind of privacy legislation can cause unnecessary confusion and reasonable doubts.

As different industries have their privacy laws, the medical field has a dedicated law. The Health Insurance Portability and Accountability Act (HIPAA) of 1996 required the US Department of Human Services (HHS) to adopt national standards for electronic health care transactions and security [50]. HIPAA consists of two parts: portability, meaning that an individual should be able to keep their health insurance if they are changing jobs and accountability to ensure the confidentiality and security of patients' information [7].

The latter of these parts is more relevant to the subject matter of this thesis. Several provisions were added to the act to help ensure federal protections for health information that could be individually identifiable. These provisions are the privacy rule, security rule, enforcement rule and final omnibus rule [50]. Solove and Hartzog [43] note that HIPAA does not cover all medical data and state laws can cover medical data more thoroughly.

The privacy rule is meant to ensure that individuals' personally identifiable health information is protected accordingly. Anonymous data is left out of the scope of the privacy rule. The rule covers health plans, health care providers and health care clearinghouses. The privacy rule attempts to limit the circumstances where the identifiable health data is used and who can use or see the data. For example, the patient's data should only be shown to the patient or a government official. The privacy rule details individual rights, such as that a patient has the right to obtain and see their protected health information and to restrict the use or disclosure of their patient data. Although in the case of restriction, the privacy rule mentions that the entity that processes the data is not obligated to restrict the processing. This most likely means the same as in the GDPR, that there are instances where the restriction right is not applicable. Individuals also have a right to amend incorrect or incomplete health information. The entity is also required to implement the necessary safeguards so that the privacy rule can be followed. These measures include assigning a privacy official and writing a privacy policy as well as training personnel. [51]

The security rule makes the privacy rule more concrete by addressing the technical and non-technical safeguards to secure individuals' electronic protected health information (e-PHI) [52]. The general rules of the security rule are:

1. Ensure the confidentiality, integrity and availability of all e-PHI that the entity creates, receives, maintains or transmits
2. Identify and protect against foreseeable threats to the security or integrity of the information.
3. Protect against foreseeable impermissible uses or disclosures
4. Ensure workforce compliance [52]

The security rule defines the CIA attributes as: "e-PHI is not available or disclosed to unauthorised persons" (confidentiality), "e-PHI is not altered or destroyed in an unauthorised manner" (integrity) and "e-PHI is accessible and usable on demand by an authorised person" (availability). The security rule is meant to be flexible, as was the privacy rule. Risk analysis and management are suggested to find out which protection measures are the most useful. In addition to administrative safeguards, the security rule requires physical and technical safeguards. Physical safeguards include facility access and control and workstation and device security. Technical safeguards consist of access control, audit control, integrity control and transmission security. [52]

The enforcement rule of HIPAA contains compliance and investigative provisions, procedures for hearings organised when there is a suspicion of wrongdoing and imposition

of civil money penalties for violations. In 2013 an omnibus rule was added to HIPAA because of another act, Health Information Technology for Economic and Clinical Health (HITECH), mandated changes. These changes specified more accurate measures, such as business associates of the entities under the influence of HIPAA also became liable to a certain extent. Also, a breach notification rule was defined, so that the entities would have to report data breaches to the Secretary of HHS within 60 days. If the breach affects over 500 individuals, then the entity must notify the media. HHS's Office for Civil Rights is the instance responsible for enforcing the rules of the HIPAA. Violations of the HIPAA may result in civil money penalties or a criminal investigation if the violation is severe enough. [53]

HIPAA's privacy and security measures are not far from the GDPR. HIPAA seems to be more specific than the GDPR in some instances, like providing concrete examples of safeguards that could prevent violations. Then again HIPAA is an act of one nation, albeit a federal one concerning all the states in the US whereas the GDPR is almost a continent-wide regulation and keeping it as general as possible might be deemed better. Both legislations are formed as all-encompassing, so they share similarities. What is surprising is the mention that an individual under HIPAA does have the right to restrict processing, but the entity is not obligated to comply with the request. Of course, the same applies to the GDPR since there are situations where the restriction is not possible, and the data controller must tell the individual why.

Regarding the chosen examples for this thesis, the USA is unique. It is apparent from the nature of the US privacy law framework why an outsider might view the legislation as lacking or loose. Then again as Esteve notes, companies like Google and Facebook can be the target of fines or other legislative measures in both the EU and the US [12]. The FTC does have influence the US. It also seems apparent that various legal protections bring with them the issue of according to which law should a perceived violation of privacy be judged? Solove and Hartzog [43] mention the attempt in the past to treat companies' privacy policies as contracts so that contract law could be applied to them. That has not been successful, and privacy policies have not been thought of as contracts in court cases so that argument has not stuck [43].

6.2 Finland

Before the GDPR, Finland as an EU member state was affected by the previous EU directive 95/46/EC. Finland's Personal Data Act (Henkilötietolaki) 523/1999 [19] was the primary legislature considering data protection before the GDPR and Act on the Electronic Processing of Client Data in Healthcare and Social Welfare (Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä) 159/2007 [18] included measures to protect data that relates to patients.

Terms familiar from the GDPR are also included in the Personal Data Act, such as personal data, processing of personal data, controller and data subject and they mean essentially the same here. The processing of personal data must be planned before the data is collected and the data must be used according to the original purpose. The Personal Data Act details the general prerequisites for processing personal data as follows:

1. the data subject has given an unambiguous consent for the processing,
2. the data subject has given an assignment, the data subject is a party of a contract or processing is necessary to take steps at the request of the data subject before entering a contract,
3. processing is necessary to protect data subject's vital interests,
4. processing is based on the provisions of another Act or is necessary for compliance with an obligation that is directed at the controller,
5. there is a connection requirement, meaning that there is a relevant connection between the data subject and controller. For example, the data subject is in service of the controller,
6. the data relates to clients or employees of a group of companies or a comparable organisation,
7. processing is necessary for payment traffic, computing or comparable task,
8. circumstance concerns generally available information on the status, duties or performance of a person in a public corporation or business. Data is processed to safeguard the controller or a third party, or
9. the Data Protection Board (Tietosuojalautakunta) has issued a permit for the processing. [19]

Only accurate and necessary personal data should be processed. Processing of sensitive data, such as race or sexual preferences, is prohibited unless there is a good reason for it, such as the processing is based on provisions of an act requiring it. The authorities of data protection are the Data Protection Ombudsman (Tietosuojavaltuutettu) and the Data Protection Board. The ombudsman provides guidance and direction, supervises the processing and makes decisions based on the Personal Data Act. The board deals with questions that relate to the processing of personal data. Both the ombudsman and the board have the right to access personal data that is processed, and information related to the legality of the processing. [19]

The Personal Data Act defines several data subject rights. Section 24 "Information on the processing of data" demands that the controller must provide information about itself and the purpose of processing the personal data to the data subject. Section 25 adds to information on the processing of data by declaring that if the personal data has been obtained for direct marketing, the data subject has the right to know the data controller, controller's address and the name of the person register they used. The right of access details that data subjects have a right to access data saved of them or to a notice that no data is saved. The right of access is not applicable in all situations, for example, in matters concerning national security or if the data is used exclusively for a historical or scientific study. The right of access entails that the data subject must prove their own identity when requesting

access and that the controller must provide the information without undue delay. The data controller must provide a reason for declining the request. [19]

If the right of access request relates to data about the data subject in the files of healthcare authorities, institutions, physicians, dentists or other health care professionals related to the state of the data subject's illness or health, the data subject must request a physician or other healthcare professional. The professional will then act on behalf of the data subject and obtain the data for the data subject to view. The data subject also has the right to rectify erroneous, unnecessary, obsolete or incomplete data without undue delay. The data controller can decline if a compelling reason exists. The last right of the data subject is to prohibit processing of their data for direct marketing, distance selling, market research, opinion polls, public registers or genealogical research. Also, automated decision making is only permitted if an act provides the decision making or the decision is made due to an agreement. [19]

The Personal Data Act additionally includes information security requirements to improve data security and storage of personal data. For data security, the controller must implement technical and organisational measures for securing personal data from unauthorised access, accidental or unlawful destruction, manipulation, disclosure, transfer or unlawful processing in general. These measures should consider the techniques available, costs, quality and quantity of the data. Those who have gained access to and knowledge of characteristic of personal data shall not disclose the data. This non-disclosure is called the secrecy obligation. If personal data is no longer necessary, then it must be destroyed unless specific provisions are preventing that. Personal data may be transferred for archival or to be used by a higher education facility if the National Archives grant permission. [19]

As for the penalties detailed in the act, the registry keeper is obligated to compensate for the damages caused by the possible processing of personal data in violation of the Personal Data Act. The penalty for a personal data offence, for breaking into a personal data file and for violating a secrecy obligation is detailed in the Penal Code (Rikoslaki) (39/1889). The Personal Data Act declares that a person shall be fined for a personal data violation if they are found guilty of gross negligence or intentionally breaking the provisions of the Act. Then again if a more severe penalty is detailed elsewhere, then the person will be sentenced according to that. [19]

Because of the national scale of the Personal Data Act, it describes more detailed measures than the GDPR does. Both the Finnish Act and the regulation mention the need for a consent from the data subject and that personal data should be processed for only those ends that the data was originally collected. Both require the data to be accurate, and that sensitive data should not be processed without a proper reason. The data subject rights also share similarities like the right to request information and correct erroneous information.

There are differences too. Overall the GDPR's data subject rights are much more detailed and, for example, the right of erasure is more explicit. The Finnish Personal Data Act requires that data be erased if it is no longer needed but does not explicitly state that an individual data subject has the right to request the erasure of their data. Also, the act does not define sanctions for breaking it but relegates that duty to other acts and codes depending on the charge. The Personal Data Act is more detailed about the practicalities of following the Act correctly which is explained by the national scope of the law. Overall though the Personal Data Act does not differ that much from the GDPR apart from more concrete measures presented in the regulation in some parts.

Act on the Electronic Processing of Client Data in Healthcare includes measures on healthcare data processing. The act also produces a unified electronic handling and filing system for patient data to produce health care services securely and to give citizens access to their data, called Kanta. [18]

The handling of patient data must be secured. The data's availability and usability must be guaranteed, as well as the integrity. The service providers must keep logging information and registry about the usage of their services and when the data is handed over to another party. This other party must be another healthcare service provider, and the data must be used to guarantee the treatment to a patient. The patient can also prohibit this data exchange. [18]

The act uses the term customer for the data subjects of these health care services. The customer has the right to be informed about the health care services, how they are used and of the legal rights the patient has. They also have the right to be informed of how their information is passed to other parties. The customer has a right to view information that has been stored of them as was detailed in the Personal Data Act. In addition to viewing the stored information they have the right to view log information about who has used that information and whether the information has been extradited to another party and why it has been used or extradited. There are exceptions though and over two-year-old log information is not required to be handed over. Log information is also not required to be shown if it is viewed that the information in the logs would endanger the patient or someone else's rights. [18]

The act details information security measures. The system fulfils the requirements when it is designed and produced so that it fulfils the laws regarding privacy and information security and works according to them. The act divides information systems to two categories A and B according to their connection to the Kanta service. If the service is part of Kanta service or connects to it through a proxy service, it is category A. All other systems belong to category B. Category A service must prove that the system fulfils all privacy and information security requirements given by approved joint testing and with a certificate given by information security evaluation facility. Category B software must prove that their system fulfils the requirements talked earlier. The service provider must develop

a self-regulatory plan regarding the privacy and information security of the service. The plan must detail how the software is administrated, used and so forth. [18]

Ministry of Social Affairs and Health's decree 298/2009 defines separately about patient files, how they should be used and what is saved into them. The decree details that the logging files defined in the healthcare act must be preserved intact for 12 years. Along these instructions, the decree's appendix includes orders on the preservation of patient files. For example, the patient's necessary information and files containing essential health care information must be preserved for 12 years after the patient has died or 120 years from the patient's birthday, if the time of death is not known. [20]

The Finnish healthcare act is much more general in appearance than USA's HIPAA. Then again in Finland, the Personal data Act contains most of the more practical legislature and the health care act the needed additional details. The Kanta service details take a large part of the act, and not much additional detail is given on how the service providers can make sure if the services are sound concerning information security and privacy. The personal data act is much more useful in that regard. The decree 298/2009 offers more regarding how long files should be preserved.

6.3 France

In France, the authority responsible for data protection is Commission Nationale de l' Informatique et des Libertés (CNIL) [9]. The critical regulation regarding data protection is Act No 78-17 of 6 January 1978 on Information Technology, Data Files and Civil Liberties [9]. There have been several amends to the act over the years, and the directive that preceded the GDPR has affected the act as well. The act applies to the automatic and non-automatic processing of personal data unless the processing is done for purely private reasons.

The data protection act defines personal data as any information related to a natural person who can be identified (indirectly or directly) by reference to an identification number or other specific factors. The personal data is lawfully processed if it has been obtained lawfully or through an explicit consent. The data must also be accurate and relevant to the processing and not retained longer than there is a need for retaining. The act does mention that the data can be retained longer if it is used for historical, statistical and scientific purposes, an explicit agreement has been given, CNIL has authorised it, or the data is used for medical research or public interest. The act differentiates between the data controller and data processor like the GDPR. Specific categories, like race or religion, of personal data, cannot be processed without reasons such as protection of human life, explicit consent is given or legal claim. There are fascinating details related to specific categories. For example, these categories can be processed when the data subject has made the information public or processing is carried out by a non-profit association. [9]

The act also details the duties of CNIL. The authority shall inform data subjects and data controllers of their rights and duties and ensure that the processing of personal data is carried out according to the act. If personal data is processed automatically, the CNIL must be notified, though not in all instances. An example of this is state security and criminal offences. The CNIL has specified standard formats for the notifications which the data processors must send to CNIL. [9]

The act describes the obligations for data controller and the rights of individuals. The data controller must inform the data subject from whom the data is obtained. This information contains parts such as the purpose of processing the data, the controller's identity and if the data is going to be transferred to a non-member state of the EU. The requirement of informing the data subject does not apply, for example, in state security-related data collecting. In parts, the act is comprehensive and for example, describes how the certificates for electronic signatures must be obtained directly from the data subject. [9]

Regarding the nature of the data and risks of processing the act compels the data controller to take useful precautions and prevent the alteration, damage and unauthorised access to the data. The CNIL must be immediately notified of data breaches and the affected individual. The notification to the affected individuals is not mandatory if CNIL has determined that the data that was breached have been made undecipherable. [9]

Regarding the rights of individuals, the act describes several. A natural person is entitled to object to the processing of their data. Right to object does not apply if there are legal obligations for data processing. The data subject also has the right to be informed about the processing concerning them if their identity has been proven, the right of access. This information includes:

- confirmation that their data or part of it is indeed processed
- purpose of processing, categories of personal data processed and to whom the data is disclosed to
- whether the data is transferred to a non-EU member state
- the personal data itself and its origin
- information on the logic of automated processing, although this must not violate any intellectual property rights. [9]

Data controllers can ask for a sum of money for delivering the data subject their data and can decline unreasonable requests. Data subjects can also request their data to be rectified, completed, updated, blocked or deleted if their data is inaccurate, incomplete, equivocal, expired or if the collection, usage, disclosing or retention is prohibited. This request is called the right of rectification. The burden of proof is on the data controller to justify that a necessary operation has been carried and if the case is disputed. The act also mentions explicitly personal medical data with the right of access. The data may be disclosed to the data subject directly, or through a doctor the data subject designates. [9]

If a data controller does not process the data according to the act, CNIL has several means of sanctioning the controller. CNIL can serve a notice to the data controller urging them to change their data processing activities. If the data controller does not comply with this request in due time, CNIL may pronounce financial sanctions or an injunction to cease the processing. If the processing activities are deemed to lead to violations of rights and liberties of data subjects, CNIL can interrupt the processing or in more critical cases use any security measure necessary for the protection of said liberties and rights. Regarding the financial sanctions, the first sanction due to a breach cannot be over €150 000 in size. If a second one happens during the next five years, the maximum is €300 000 or 5% of the previous annual revenue, although that cannot exceed €300 000 either. There are also criminal provisions mentioned. These provisions are used, for example, when the data controller is not providing the correct information about stored information records. [9]

The act mentions explicitly medical research related data processing in chapter nine. CNIL gives authorisation to the use of personal medical data for research purposes. Data which allows for the identification of individuals must be codified, although it is not mandatory if the nature of the research demands identification. Data subjects have the right to object to this, and in case of identifying biological samples, express consent is required. Chapter ten then goes on to detail measures related to the processing of medical data for the evaluation or analysis of care and prevention. This kind of processing is permitted by the CNIL. This kind of personal data cannot be used so that a subject can be identified although some data points can exist, but not attributes like the social security number. The Public Health Code details more concerning personal medical data. [9]

As the United States of America and Finland, France has their more healthcare related privacy legislation in a separate law called the Public Health Code. Some of the articles of the Public Health Code are mentioned in the Act No 78-17 of 6 January 1978 on Information Technology, Data Files and Civil Liberties [9]. Since the French Public Health Code is remarkably more protracted than the health-related legislations of USA and Finland, it is not explained further in this thesis. Luckily the Act No 78-17 contains mentions of health-related rules, and since the directive preceding the GDPR naturally affected France as an EU member state, no vast gaps are formed to the discussion. France is still a good example to include in this comparison since the Act No 78-17 is detailed and provides an excellent counterpoint to the Finnish Personal Data Act.

In the French privacy legislation, the level of detail seems to be higher than in the Finnish Personal Data Act or the GDPR. It is interesting to see how the Finnish and French data protection acts differ from each other as they were both affected by the 95/46/EC directive. The unifying aim of the GDPR comes apparent as more differences in member state legislations are bound to exist. The French act contains data breach notifications requirements and detailed sanctions whereas the Finnish act is missing both. The similarities are still the majority, and the individual rights presented by the acts are for the most

part identical. HIPAA resembles the Act No 78-17 more than the Finnish Personal Data Act, as both HIPAA and Act No 78-17 are quite concrete.

6.4 Summary of the discussed laws

The United States of America differs from France and Finland partly due to the differentiation of federal and state laws. France and Finland are both members of the EU and having implemented the directive 95/46/EC; they still have some variety between their data protection related legislations. If the HIPAA is looked at like the acts of Finland and France, they are surprisingly similar considering how the law in the USA can be perceived to be different.

The comparison points of the different legislations are collected to Table 2. As can be seen from the table, when the requirements extracted from the legislations are put side by side and condensed, the differences are not significant. It may seem that the GDPR does not add to the overall field of privacy legislations that many new aspects while looking at Table 2. The table emphasises that the GDPR's influence is not necessarily the number of details or unique attributes. Instead, the aim seems to be more a more profound foundational change to how the information is handled and how individuals can affect that handling.

Table 2. *Summary of the discussed legislations*

Legislative attribute	The United States of America	Finland	France	The GDPR
Special categories of personal data were presented.	Not detailed in analysed legislative texts.	Several categories, such as: ethnicity, religious affiliation, trade union membership, health, sexual preferences.	Several categories, such as Race, ethnicity, political-, religious-, or philosophical opinions, health, sexual life	Several categories, such as: genetic data, health and sex life. Children's personal data is also a special category.
Separate laws/acts for electronic medical data processing and privacy.	Federal level privacy act does not exist, but FTC's Federal Trade Commission Act is applicable. For medical data: HIPAA.	Personal Data Act and Act on the Electronic Processing of Client Data in Healthcare	Act No 78-17 of 6 January 1978 on Information Technology, Data Files and Civil Liberties and The Public Health Code	Primarily contains privacy related articles, but data concerning health is also detailed.

The individual has a right to access or view personal data that has been stored of them.	HIPAA details that individuals have a right to access and review their protected health information. Not detailed in the Federal Trade Commission Act.	Detailed in the Personal Data Act. Also includes other information, such as the source and usage of the data.	Detailed in the Act No 78-17. Also includes other information, such as recipients and categories of the data.	Detailed in the GDPR. Also includes other information, such as purposes of processing, categories of data, storage period.
The individual has a right to restrict the processing of their personal data.	Under HIPAA, restrict use or disclosure of health information. Not detailed in the FTC act.	Not detailed in the analysed legislative texts.	Not detailed in the analysed legislative texts.	Processing must be restricted upon request where applicable.
The individual has a right to ask for their personal data to be removed.	Not detailed in analysed legislative texts.	Personal data should be deleted if it is no longer needed or it is unnecessary.	Individual may ask the data controller to delete their personal data.	A right of erasure is included in the regulation with exceptions.
The individual has a right to ask for correction or completion of their personal data.	Under HIPAA, incomplete or incorrect personal data can be amended. Not mentioned in the FTC act.	Erroneous, unnecessary, incomplete or obsolete personal data can be rectified.	Incomplete or incorrect personal data can be completed and corrected.	Rectification of incorrect or incomplete personal data is detailed.
The individual has a right for data portability from one data controller to another.	Not detailed in analysed legislative texts.	Not detailed in analysed legislative texts.	Not detailed in analysed legislative texts.	When processing is automatic and based on consent.
The individual has a right to object to the processing of their personal data.	Not detailed in the analysed legislative texts.	Yes, although constrained to specific cases.	Can object to the processing of any data related to the data subject.	At any time, the right to object to the processing of personal data is in effect.
Special measures are detailed against automatic processing.	Not detailed in the analysed legislative texts.	Decision making based on purely automatic processing is permitted only in certain situations.	Several mentions of the restriction of automatic processing and decision making.	Right to not be a subject for purely automated processing, including profiling.

Sanctions are detailed for not following the requirements presented in the legislation.	FTC can both fine companies and subject them to audits. HIPAA details civil money penalties and criminal sanctions.	Fines and other measures, although these are not discussed in detail in the analysed acts.	Fines between 150 000€ - 300 000€. Also, criminal provisions are possible.	Different sanction categories. The maximum is 20 000 000€ or 4% annual turnover.
A supervisory authority is detailed.	Yes, FTC for privacy and HHS for HIPAA.	Data protection board and the data protection ombudsman.	The CNIL	Member states provide public supervisory authorities who work with the Commission.
Data breach notification requirement.	HIPAA has a Breach Notification Rule. Individuals and HHS secretary must be notified after a data breach and the media if the affected amount of people is large.	Not detailed in the analysed legislative texts.	Notify CNIL immediately, individuals if their personal data has been breached unless CNIL orders otherwise.	Within 72 hours notify the supervisory authority, data subject if breach results in high risk to subjects.
Means of gathering personal data is discussed.	No mention of how personal data is gathered in HIPAA or the FTC act.	Data is gathered and processed through an explicit consent or a legal requirement.	Data is gathered and processed through an explicit consent or a legal requirement.	Data is gathered and processed through an explicit consent or a legal requirement.

France and Finland do not share the idea of self-management when related to privacy as much as the US does. The GDPR emphasises even to a greater extend the vaster level of unified data protection than the directive that preceded it, so the trend does not seem to be going towards the model of the US.

It will be interesting to see if the GDPR changes anything regarding the legislation in the US and how. One effect might be the self-regulation of companies working both in the EU and in the US since those companies must abide by the GDPR. As having multiple data processing and protection conventions overlapping inside the same company might be deemed redundant, the companies could make the requirements of the GDPR the norm. Then again, the GDPR might be deemed too strict. While privacy laws of the US may be

deemed lacking, the GDPR might be a step into the other end. Esteve [12] notes that a balance between strict and lenient privacy law should be reached.

One glaring difference between The US and the European legislations analysed was that no inclusion of data gathering means could be found considering the US. Finland, France and the GDPR all mention strictly about how the information collection must be based on an explicit consent or to a legal requirement. It is not surprising that HIPAA did not contain a mention of data collecting since most likely it is, for the most part, collected because of a legal requirement regarding the treating of a patient. Solove and Hartzog [43] explain that there does not exist a federal law that directly protects the data collection or use by entities like Amazon or Facebook. As was mentioned before, the companies' privacy policies have a lot of say and FTC then supervises that these policies are respected. The legislative attribute considering separated acts for medical data is marked "to some extent" in the case of the US because of this missing federal level law for privacy. There exists a federal level law for processing health related data, but not a single law for privacy.

France and Finland do not have any mentions of specific technical measures or mentions of risk analysis in their privacy laws. The GDPR and HIPAA, even with their grander scale are more specific in this case than analysed member state acts. The GDPR might have included these measures to make the regulation more detailed and to be in line with larger best practices regarding information security and data protection.

The Finnish data protection related acts were the only ones that did not mention any specifics about how data breaches should be handled. Personal Data Act does discuss sanctioning the data controllers that do not follow the rules of the act and data breaches might be interpreted as belonging to that section of the act. The notification conventions related to data breaches are likely detailed elsewhere, but still, the Finnish data act differs in this case from the others.

The magnitude of the fines that can be issued due to a data breach or other violation seems to have risen with the GDPR. While the fines the FTC has given in the US are perceived as small [43], and the fines of France are also relatively small, the GDPR's sanctions are at least on paper many times what the previous maximums have been. What remains to be seen is how these fines are issued and whether cases are settled outside of the courts like what has happened in the US.

Every legislation except the ones from the US contains a notion that some specific pieces of personal data, like religion and ethnicity, are more sensitive than other data points. What also seems to be missing from the HIPAA and FTC's guidelines that are contained in other analysed legislations is automatic processing. Even before the GDPR, France and Finland specified measures against purely automatic processing and the GDPR continues

this path. HIPAA not mentioning automatic processing is not peculiar since such processing could be more common in consumer services and there might be state level measures that restrict automatic processing.

Regarding the data subject rights, the right of access to personal data and the right of rectification were in some name and form mentioned in every legislation. The right of restriction of processing was not detailed only in the Finnish Personal Data Act, although the right to object could be considered to cover some parts of the restriction. The right to object is then missing from the US federal law level. The right of erasure is missing from HIPAA which is understandable considering the application area of it, but it is apparent that the US does not have a federal level right for the individuals to erase their data from different data processors or data controllers. France and the GDPR have the most widely affecting applications of this right since in the Finnish perspective the data is only supposed to be removed when it is no longer needed. An entirely new right the GDPR has is the right for data portability. HIPAA and

The differences in the level of detail between the legislations are apparent. There exist strong reasons for this considering the different scopes of the analysed legislations, but excluding the GDPR's scale, the Finnish acts are noticeably more ambiguous than the others. The Finnish health care-oriented act spends quite a lot of space in defining the Kanta service and leaves most of the data protection side to the Personal Data Act. In France, this kind of divide seems to be also true. In the United States, HIPAA is the de facto act for the health care industry, and otherwise, all-encompassing data protection act is missing. The FTC operates as the supervisory authority in these in-between cases and in that way reminds the French CNIL and the Finnish data protection board.

The state laws of the USA were not analysed here which leaves gaps considering the comparison of legislations done in this chapter. The GDPR is a widescale regulation affecting EU member states, so the analysed levels are similar. Analysing the legislation of each of the US state would nevertheless be out of the scope of this thesis.

Based on this analysis, the GDPR can be determined to be the next step by the authorities to try to ensure an individual's privacy. While the regulation does not seem to try and reinvent the wheel, it details as its basis a new way of thinking about data protection and how the companies must consider the individual while they are processing data. Whether there are glaring flaws or omissions in the regulation will be seen in the next few years, and the current contents might be refined as real life incidents, and cases arise.

7. MEDICAL SOFTWARE ANALYSIS

The GDPR presents several new requirements for software applications as well as refining old ones. Since the regulation at the time of writing this thesis is new, some of these requirements and their impacts are not unequivocal. To make the changes presented by the GDPR more concrete, the requirements of the regulation are analysed concerning a medical software application. A medical software application is an excellent example of a software category that contains sensitive data and has been more heavily regulated even before the GDPR [22].

The software application in question is a web application that is used for risk calculation related to screening taking place during pregnancy. The application contains patients' personal information and diagnosis data. The data the application primarily holds is classified by the GDPR as data concerning health in part 35 of the introductory section and is one special category of information gathering [13]. That is why the risks of processing such data can be considered high.

In Finland, the software is classified as category B since it does not connect to the Kanta service. The software is used worldwide, so it has already had to comply with the legislations of several different nations. Patients themselves do not use the software as it is meant for medical professionals to use. Also, system administrators administer the system, so they have elevated privileges.

The application's architecture consists of a web server that client computers connect to and of a relational database for storing data to which the web-server is connected. The relational database used is Microsoft's SQL Server and the web application framework used is .NET. The application supports Google Chrome, Mozilla Firefox and Microsoft Internet Explorer browsers. There are several views in the application, where a user can, for example, search for patients based on several different parameters, view and fill in patient's personal and diagnosis data and see the workflow status of patients. These all depend on the given user group rights.

The application is deployed into an internal network of a site that is, for example, a hospital with limited access to other networks via a demilitarised zone (DMZ). In network security a DMZ is denoted as a subnetwork in the internal network that is between two firewalls, the external one protecting DMZ from external networks and the internal firewall protecting the internal network from DMZ itself [54, p. 297].

The users are assigned to user groups of which there are several. Groups can also be added to the application. The groups restrict the actions a user can perform in the application,

and a user can belong to multiple groups. A user can also be removed from groups and the account shut off entirely.

The application requires the user to log in with their user account. The password expires after 30 days, and the new password must be different from 10 previous passwords. Every password is encrypted, so they are not stored as plaintext. The application automatically signs the user out if they have not used the application during the last ten minutes to prevent unauthorised access and authentication issues.

Two-step verification with a one-time token code can be used when logging in to add a layer of security into the application especially when a user logs in from outside of the internal network. The two-step verification is required when a user logs in to the application from a public network with remote access. The application restricts the rights of the user when they log in remotely. When a user is part of a remote user user-group, they may only view, and query a limited amount of information related to their assigned patients.

The descriptions of the software are kept general because of the sensitivity of privacy and information security related subjects. That is why no exact details are given so the software application could not be singled out. However, the relevant and essential details related to this thesis' analysis of the application are given so that an understanding of the application and its essential functions can be formed.

The GDPR is a recent development and came to effect during the writing of this thesis. Regulations are contemplated in the national and EU courts where the regulation's articles are thought upon on a case by case basis, forming more strong precedents regarding what practices of personal data processing are deemed suitable and lawful. The practicalities of the regulation are not entirely known since the GDPR is so new. Previous literature has also been mostly speculative and scarce. Lack of previous research is understandable since the regulation is new and has been in a transitional period for two years.

The data controllers and processors must abide by the regulation if they intend to operate on the EU soil collecting the data of EU citizens even when some uncertainties exist. That is why it is beneficial to try and apply the regulation into a concrete example where the regulation must be considered when the application in question is developed further. This way a working example can be produced on how a software application can adhere to the GDPR which can then be hopefully applied to other applications as well.

Privacy and security laws before the GDPR have been stricter for medical software applications because of the nature of the data they contain. While the GDPR brings new demands with it, such software has already had to comply with several requirements especially if the software is used internationally. By analysing a medical software application from the GDPR's viewpoint, the nature of the regulation's changes is visible, and the

regulation's new calls for the protection of sensitive data are weighed against previously existing rules and features of the software application.

The knowledge of the features of the analysed medical software was gained through reading the documents and the manual, further development and usage of the software. First, the brief overview of the software application is given, and the relevant requirements of the GDPR are discussed as some of the regulation's measures are more relevant than others. Then the application's present state is described and compared to the requirements of the GDPR.

The comparison between the software's current state and the requirements the software faces, in this case from the GDPR, is done with the help of ISO/IEC 25010 standard's Product Quality Model [46]. After the evaluation of the present state, the required changes are listed. There are also some changes that would be beneficial but not compulsory to add to the overall security and privacy level of the application. These change proposals are then evaluated since they are bound to affect the functionality and the overall information security and data protection level of the application.

7.1 Method of analysis

The ISO/IEC 25010 standard defines two models: a quality in use model and product quality model, of which the latter is more appropriate in this case. The models provide a set of quality characteristics which can be evaluated with the help of needed requirements to measure the overall completeness of the software product. [46]

The standard is part of the 25000 families of standards that form the System and software Quality Requirements and Evaluation (SQuaRE) for system and software engineering. The standard 2501n is the Quality Model Division that provides detailed quality models for computer systems. [46]

The product quality model divides quality properties into eight characteristics which are further divided into sub-characteristics [46]. These are visualised in Figure 3. Not all these characteristics are relevant for this thesis, so the application is not analysed regarding all of them. The most relevant main characteristics regarding this thesis are functional suitability and security. As explained in the standard the characteristics can be used as a checklist, providing a basis for estimating the effort and activities that are needed in development [46].

The application is analysed concerning the functional suitability sub-characteristics. The sub-characteristics are functional completeness, functional correctness and functional appropriateness. The standard defines these characteristics as:

- functional completeness: how well the set of functions covers all the specified tasks and user objectives

- functional correctness: how well the system provides the correct results with the needed degree of precision
- functional appropriateness: how well the functions facilitate the accomplishment of specified tasks and objectives [46]

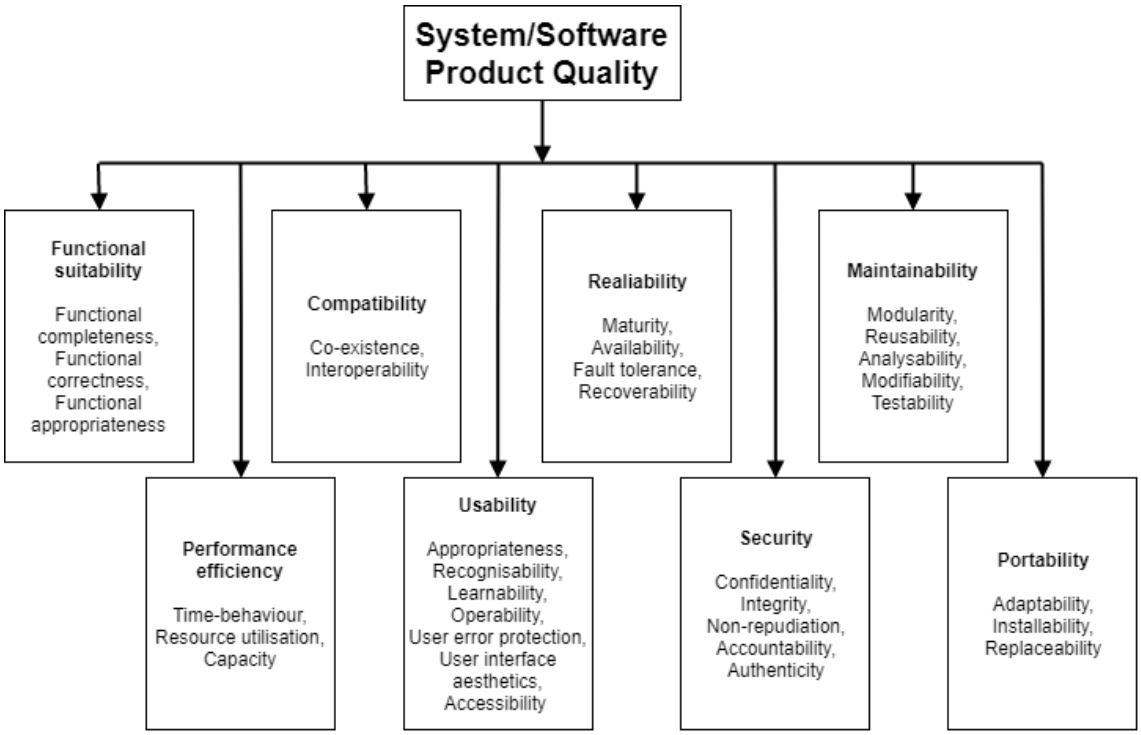

Figure 3. *Software Product Quality Model [46]*

The ISO/IEC 25010 standard does not define metrics for the analysis, so such metrics are formed. A numerical score is given from a range specified in Table 3. for every detailed and relevant requirement of the GDPR concerning the sub-characteristics of functional suitability. The scoring is done to quantify the current state of the application. The table demonstrates an example concerning the right of data subjects to gain access to their personal data, to clarify the meaning of the grades.

Grade one means that the sub-characteristic in the application's current state is not compliant with the requirement. The partial compliance represented by grade two means that while the application does contain some functionality or features to support the requirement, the software is mostly not compliant, and a fair amount of work is needed to make the application compliant. Grade three means that while the application is mostly compliant with appropriate functionality and features, some work must be done to achieve compliance. Grade four means that the software is fully compliant already with a suitable functionality to fulfil the specified requirement.

After the sub-characteristics have been rated, they must be aggregated to find out the overall state of the requirement and overall state of the application. Xu et al. [57] mention

that the arithmetic mean is not entirely suitable when analysing information security since it does not consider the principle of the weakest link. Merely using the minimum function would not fit either as every system with one serious flaw would be rated the same, so the analysis would not be distinctive. Xu et al. propose using the power mean, also called the generalised mean, with parameter value -2, as the function then emphasises lower values without completely disregarding the higher, better values.

Table 3. *The grading scale for the quality sub-characteristics*

Explanation	Grade	Example of completeness	Example of correctness	Example of appropriateness
Not fulfilled	1	The application's functionality does not cover the right of access.	The collected data does not contain any personal data of the data subject.	The application does not contain functionality that suits collecting the data subject personal data for the right of access.
Partially fulfilled	2	The application's functionality addresses some parts of the right of access, like an ability to generate reports in a widely used format.	The personal data collected contains some pieces of data that is needed. For example, relevant audit trail data but not any personal information.	There is a functionality to collect personal data, but it also gathers other data not suited for the data subject.
Mostly fulfilled	3	The application's functionality is mostly complete, but for example, the data is collected to a difficult to use format.	The personal data collected contains nearly every piece of data that is needed. For example, all personal data and most of the audit data.	There is a functionality that gives data subjects access to their personal data, but it is meant for another usage.
Fulfilled	4	The functionality covers all the steps required to gather and show data subject's personal data in a commonly used form.	The personal data collected contains every piece of data that is required by the right of access.	There is a dedicated functionality to give a data subject access to their data, and the functionality is designed for this task.

The approach of Xu et al. [57] fit the purposes of this analysis as well. The GDPR emphasises the overall security and data protection level of an application, so even a small number of serious flaws is a bad thing. Breaking any of the given user rights, for example, lands the data processor to the highest sanction category. So, it is not enough that most of

the rights are adequately implemented. That is why the smaller values should be emphasised.

$$\boldsymbol{M}_p(a_1, a_2, \ldots, a_n) \equiv \left(\frac{1}{n} \sum_{k=1}^{n} a_k^p\right)^{\frac{1}{p}} \quad (1)$$

The generalised mean is presented the Equation 1. All $a_k \geq 0$. By changing the value of parameter p, the equation produces different means. For example, p=1 calculates the arithmetic mean. [5] In this analysis, the value p = -2 is used to emphasise the lower values accordingly.

7.2 The GDPR's requirements

The company that has created the application typically does not process the data themselves and does not receive it. The institutions where the application is used process the data. The data is collected both through a consent form that a patient fills out when they check in to an institution and through legal requirements. Since parts of the data gathering are carried out due to an explicit consent was given by a data subject, the rights of the subjects needed to be analysed and considered [13]. This analysis does not consider how the consent is collected since it is done on sites where this application is deployed. Since the company making the software does not receive the data, their data collection policies are not analysed. What they can do is to make sure that the application complies with the GDPR, so that their customers can enforce their compliance. Since it is also the aim of this thesis to analyse the GDPR and the requirements it brings, only the software application is considered.

Concerning all the rights, although some more than others, one significant question is how to get the patient's request to the users of the application and then how to deliver the answer back to the patient. These questions will not be answered in this thesis since these are also the responsibility of the sites where this application is deployed. The Finnish Personal Data Act that preceded the GDPR stated that a data subject needs to contact a medical professional if they want to access their information contained in a healthcare related system and provide a valid identification to be able to see their data [19]. Even though the GDPR does not explicitly mention such qualification, it seems reasonable to conclude that confirming a data subject's identity and answering their request would work in the same way in the future.

As information is collected through consent from a patient, the individuals' rights need to be analysed and considered. The patient has the right to request access to the information that has been collected from them [13]. That data needs to be obtained transparently from the application and in a widely used format [13]. As the data needs to be in a clear format, it means that some thought must be put into the presentation of the information and that the format, such as the file type, should not be an obscure one, but instead, something

data subjects can access. Also, if the personal data is deemed to be inaccurate or insufficient, the patient has the right to request a correction to the data [13]. These two rights have already been included in the Finnish Personal Data Act [19] so, they do not bring any new major requirements as the application is approved to be used in Finland. The application should have a reporting functionality that gathers patient data into a widely used format that can be then sent to the patient that requested their data. The right of rectification details that the patient has a right to rectify inaccurate personal data belonging to them.

The right of erasure is not so clear-cut as the right to obtain data and the right to rectify data. While through consent the patient does have the right to request erasure of their data, other legislative measures may set restrictions to the usage of the erasure right. The GDPR specifies several exceptions to the right to erase data, and many of them are applicable here. In this case, public authority takes precedence since the medical organisation is exercising their public authority by treating a patient according to their personal information. Also, the data is of interest to public health and for archival purposes. Member state law is also in the way in Finland for example, where the Ministry of Social Affairs and Health's decree 298/2009 specifies the extent of information preservation for different types of patient information [20].

Then again it could be argued that in an individual case it would be beneficial to be able to remove someone's information from the application. For example, in the case of an erroneous entry made into the application that serves no purpose for the patient or the registry keeper, the entry should be able to be deleted due to the GDPR's data minimisation requirements. Another case would be that someone's data has been exported out of the application for archival and that data is not needed in that application anymore. Such a feature should be implemented so that only more privileged users can do it, only in a specific circumstance and not by accident. Still, while the functionality for erasing data would be beneficial, the wording of the GDPR does not require the right of erasure in the case of this application.

Right to the restriction of processing is more straightforward than the right of erasure. If a patient wants the processing of their data to be restricted, then their patient information should be put aside and not processed until whatever the basis for the restriction was, is cleared. The restriction should not interfere with the obligations of the data collector. It can be inferred that it should be possible to restrict the processing of data if that is requested and then be able to successfully and efficiently allow the usage if the reason for the restriction is no longer valid or if the data is needed again. A situation where the data might be needed again is, for example, a health-related reason where the personal data is needed for an accurate diagnosis or information on the state of the patient.

Right to data portability can be ruled out in this case since one of the requirements along with the given consent is that the processing is automated. In the application, processing

is not automatic and is always initiated by a human user. That is why this right is not relevant or applicable to this application. Right to object is quite like restriction and erasure since it prevents processing of data subject's data. It, however, is not deleted or put aside, instead, the processing is just stopped. This right should be considered in the case of the analysed application although if such a request comes from a data subject, the collector can provide compelling reasons why the right to object is not applicable such as the patient needs to be treated according to law and the processing of the data is necessary for the treatment.

In addition to the data subject's rights, the information security measures and guidelines in the regulation need to be considered. The data processed in the application likely has a high-risk nature which needs to be considered when analysing the application. The personal data must be adequately protected when it is stored in the database or elsewhere, transmitted from a part of the application to another and when the data is in use. Encryption is one answer since with the necessary and appropriate level of encryption it is a challenging task for an adversary to figure out what the data contains. The case for encryption is especially real for the transmitted data as it is virtually impossible for an eavesdropper to decipher encrypted data moving in a network if the used encryption is strong enough.

Anonymisation of the data is not applicable in the case of the analysed application and anonymising the data would have taken out of the GDPR's scope. Pseudonymisation is more applicable as it is not necessary to use a patient's name, social security number or similar for identifying them in every part of the application and the database. By using pseudonyms even though a part of the database would be compromised the real identities of the patients might not be revealed as the patients would be identified by their pseudonyms across the database excluding a few tables that link the pseudonym to the actual data subject.

Enforcing confidentiality, integrity and availability as well as authenticity is paramount in the application. Previously mentioned encryption is one way, but proper access control and other industry best practices should also be followed. Measures such as organisational ones, backups and resilience while necessary are not in focus here as sites who operate this application most likely implement these practices. The application should be able to support these measures, and the company that is developing the application can issue guidelines on how the product should be used.

The requirements derived from the GDPR are as follows:

1. Right to access data
2. Right to rectification
3. Right to a restriction of processing
4. Right to object
5. Data protected in use

6. Data protected in transit
7. Data protected at rest

The data subject rights are straightforward when forming the requirements, but the information security requirements are more general in nature. Because the regulation requires complete data lifecycle protection, the requirements were distilled into three requirements: data protected at rest, in transit and in use. These cover the use cases of the personal data and current information security and data protection measures for each can be evaluated.

7.3 Present state of the application

Currently, the application has extensive reporting functionality. The functionality is extensive in part because the application generates medical reports of the patients, so most of the reports are medical. The purely medical report is ill-fitted for sending to the data subjects because it may contain incomplete diagnosis data which could be misinterpreted by the data subject. That is why the developers of the application do not want to show the existing medical report to patients.

In addition to the medical report, the application does have the ability to generate an audit log report on a patient. The application collects data about what actions each user has made on which parts of the application and an extensive audit trail is formed from the tracked audit data. The audit log report collects information regarding one patient about who accessed or modified the patient's information, what the action was and when was it made. The report is highly detailed and maybe too detailed for a non-expert user of the application to get useful information out of it. The Finnish Act on the Electronic Processing of Client Data in Healthcare also details that potentially harmful log information should not be shown to the patients [18]. The application does have several reports, but a properly fitting report template for the GDPR's data subject's rights does not exist. The data subject's right to request information stored of them cannot be fulfilled currently.

Regarding functional suitability sub-characteristics, functional completeness is rated three, since while there exists a functionality to generate reports in a widely used format, none of the report templates contains all the information needed. Amending the situation should not demand much work however since all that is needed is a new report template. Functional correctness is rated two since currently, the functionality does not provide principally correct results regarding the GDPR requirements. The information on each separate report is not correct as the required precision is that the report should have patient personal data and condensed audit data concerning their files. The functional appropriateness is rated four since the application has perfectly appropriate functionality for this requirement.

The right to rectification can be currently fulfilled within the application. This right has more to do with providing the ability for the patient to make the request and that the changes are then made to the personal data of that patient. Since the appropriate user can make changes to the personal data of a patient and these changes leave logs, the functional sub-characteristics of this requirements can all be rated four.

Currently, there is no way to delete a patient from the application or its database. Not being able to delete patient information is understandable since legislation can require the storage of patient data up to several years and because there might not be a need to make more room to a database if an entry is not needed or erroneous. Deletion functionality is also risky since it is possible to erase data that is required by law to store. Although the GDPR does present several exceptions as to when the right of erasure is not applicable, there might be situations, where this functionality is needed as was discussed in the previous chapter. If done right, the functionality will not interfere with the routine use of the application, and there is no harm in preparing for a sudden need rather than leave the right out and suddenly need it. Since the GDPR does not require this right to be fulfilled in this application's case, the right is left out of the analysis

Right to the restriction of processing can be handled with a feature called on-hold that is already a feature of the application. This feature effectively excludes a specific patient from the operations that the users of the application can perform until the patient is taken out from hold. The user that puts a patient on-hold needs to give a reason for the action and that given reason is also saved so that other users know why the patient is excluded. Putting a patient on-hold is immediate and not tasking to undo. As such the software already has the means to comply with this data subject right. Then again, the feature is not originally designed to be used for this purpose which might present some complications.

Regarding the sub-characteristics, functional completeness is graded two, since while the on-hold function itself is ready, it is not initially designed for this use. As that function is used for other tasks as well, there needs to be some way of segregating between patients that are on hold because they have exercised their right or because of another reason. The user can specify the reason for putting the patient on-hold, put a clear representation should be ensured. For correctness, the application is graded two. Appropriateness is graded three since the functionality is appropriate although not initially designed for this use, so some modifications are needed.

It is possible to export patient's data from the application into a machine-readable format, for example in a comma-separated values (CSV) format, to move the information from the system to another. The export feature could be used for archival purposes too if the archive is another dedicated system. Although the GDPR's wording does not strictly require the compliance to the right to data portability in this application's case, the right could be fulfilled if needed. Part of the information gathering is based on consent, so in

that sense, this right could be applicable. The data would probably not be sent to the data subject, but directly to another system. As the GDPR does not explicitly require this right in the case of the analysed application, it is left out of the analysis.

Right to object based on the regulation's specification could also be handled with the on-hold-feature. The necessity of complying with the right to object in this case is unclear. Even if complying is necessary, the data processor can provide a compelling reason why the processing needs to continue. In the application, the reason is health-related data. Still, if this right is analysed here, the on-hold feature would probably not be the best option. Since the regulation does not mention anything about reversing the right to object, then if the on-hold feature is used, the patient is permanently on-hold. This distinction would be required to show to the users, so they do not accidentally take the patient out of on-hold. Functional completeness is therefore rated two, correctness two and appropriateness two. These grades mirror right to restriction on some level, but the existing functionality is less appropriate in this case.

The access control of the application divides users into roles with the possibility of allocating each role with different rights and assign users to these roles, so not all functionality is available to all users. The least privilege concept can then be used correctly. Authentication is also robust with measures to authenticate the user and locking the application when it has not been used in a while. Log traces are also generated upon entries and for example when a user accesses a patient's file. This way unauthorised accesses still leave a trace behind.

Andress [1, p. 78] mentioned that data in use is difficult to protect since a user can do several different things with the data. Also, as was mentioned in Chapter 3, some devices can leak sensitive data from memory, and attacks such as buffer overflows can force leakages to happen. Securing data in use puts much responsibility to the user as they can break the confidentiality, integrity, availability and authenticity of the data. One mitigating factor the application has are the audit logs, so modifications at least do not go unnoticed. Because the responsibility rests on the user, one way the application can control this is the mentioned access control level. Not giving everyone the same rights enhances the CIA and authenticity attributes. Logically another way to increase data in use protection would be device full disk encryption since it would encrypt the data currently stored on the user's computer. Full disk encryption of the users' devices is outside the bounds of this approach since the sites would implement this. The risk for an attack against the users' computers physically in a closed and guarded environment is not as likely as some more mundane ways of breaking information security.

Buffer overflows or similar types of means to try and acquire information from the application during use is another vector but by themselves worthy of a wholly separate analysis since there are a substantial amount of ways to conduct this type of attack and how to

protect against it. OWASP [35] mentions that the .NET platform is not particularly affected by buffer overflows, so the framework of choice also mitigates the risk of overflows. When the protection level of data in use is evaluated, the primary means of evaluation is access control. Since it is robust the grade for all functional suitability characteristics is four.

Not being in a publicly accessible network might make the network more secure, but an internal network is not threat free. An organisation can quickly focus on external threats and forget about internal threats, even though the internal threats could be as common as external ones [24]. As Kemp [24] notes, it is far easier to sniff web traffic if that individual already has access to the infrastructure because they work in the said organisation. That is why the internal network and its potential risks need to be considered and not take the perceived security for granted. A DMZ creates a buffer area where external networks can connect to and makes unauthorised intrusion to a network harder. DMZ-type of protections are useful, but they are not by themselves enough in securing an internal network [54, p. 337]. DMZ is also not by itself enough when data protection is the aim since the DMZ does not consider itself with what information contains if it is deemed safe. Confidentiality can be broken even with DMZ in place. Limited remote operating functionalities are also present, and users can perform limited tasks while connecting from another network. Andress [1, p. 77] mentions VPN's as a possible solution too when a private connection is wanted.

Even considering the caveats, the internal network provides a high level of safety. The connection from the client computers to the web-servers can be made via Hyper-Text Transfer Protocol Secure (HTTPS) encrypting the connection between the web-server and the client computers. HTTPS supports the protection of the data in transit, at least between those two points. As the application works in an internal network, the security certificates are self-signed which is adequate in this case.

The connection between the web-server and the database server is not encrypted, so the data is sent in the clear. As the data is sent unprotected the data in transit protections are not wholly complete. That means that the confidentiality and integrity of the information cannot wholly be ensured. Authenticity is also questionable if the messages can be changed while transmitting. That is why the functional completeness grade is two since data is only partially delivered encrypted. Functional correctness is therefore also graded two. Functional appropriateness is graded three since as the only missing functionality is an encrypted connection between the database and the web server. This functionality is already included in the .NET framework and SQL Server [31], so the functionality itself exists but just not used currently.

The database uses pseudonymisation of data widely, and the patients are identified based on pseudonym in most instances. There is only one table that links the pseudonym to the actual person. The database access control is handled on a user credential basis although

the database itself is not encrypted on disk. As was discussed in Chapter 3 concerning data protection at rest, physical access limitation and encryption are fitting means of securing data that at rest. While accounting for physical barriers to users' computers is not possible in this thesis since that falls under the responsibility of the sites themselves, access control limits the ability to use the application. Since the application does not stay open during extended times of inactivity, it limits the unauthorised accesses.

Andress [1, p. 75] mentions that encrypting the device disk also protect data at rest. Encrypting the discs means that if someone would try and access the data directly from the disk, they would only see encrypted data. The availability of the data needs to be ensured, so the encryption must not interfere too much with the everyday use of the application. As for the grading, functional completeness is measured as two since there is working to be done if complete protection is desired. Same goes for the functional correctness. As for the appropriateness, data at rest protection is graded three since the protections in place are for the most part appropriate.

The results of the analysis of the application's current state are presented in Table 4. All in all, the state of the application is decent with a fair amount of protections already in place. The total grade is calculated with the generalised mean presented in Equation 1. and is presented for each requirement separately as well as the overall total grade for the whole application including all the requirements. The overall total grade for the application is 2,49.

Table 4. *Results regarding the current state of the application and the GDPR*

The GDPR's requirement	Functional completeness	Functional correctness	Functional appropriateness	Total grade
Right to access data	3	2	4	2,66
Right to rectification	4	4	4	4
Right to a restriction of processing	2	2	3	2,22
Right to object	2	2	2	2
Data protected in use	4	4	4	4
Data protected in transit	2	2	3	2,22
Data protected at rest	2	2	3	2,22
Total grade	2,38	2,26	3.01	2,49

As can be seen from the resulting score, the application is not entirely in line with the requirements. Nevertheless, the score can be called satisfactory since the generalised

mean drags the score downward rather than upward. For example, if the internal network were to be considered perfectly secure, the grade would be much closer to three. The grade describes the application's current state well and considers how easy or hard the modifications presumably are.

The right to request correction, data portability and data protected in use are the only requirements entirely fulfilled from the start, and there is no need for changes related to them. All the other functionality requires work though and ideas for the improvement. The functional appropriateness score of 3.01 is the highest of sub-characteristic grades which indicates that the application contains sufficiently proper functionality and that several changes can be made relatively quickly. The .NET framework affects this in part since it supports the easy addition of several functionalities.

Functional completeness and correctness did not fare as well as appropriateness. This difference shows that while the functionality in place or offered is appropriate, it is not complete and does not provide correct results in many instances. These gaps call for several improvements to be made. On the other hand, it might be better for the application that the results are the way they are. Since the functionality is mostly appropriate the only work that needs to be done is to fill the gaps and complete the functionality to be on the desired level.

The requirements grades are more in uniform, mostly because the right to rectification and data protection in use requirements are already fulfilled sufficiently. Three requirements have the same grading which again mirrors how the appropriateness is there but that the functionality is not wholly complete and correct. Right to access data is the highest graded requirement if the full graded ones are disregarded. As the reporting functionality is appropriate and suitable and most advanced regarding the requirements presented by the GDPR, the grade is fitting.

7.4 Proposed changes

Even though the present state of the application is good considering the GDPR’s requirements, there are evident changes that should be made. The required changes have mostly to do with the GDPR’s data subject rights, but some information security related functionality should be enhanced too.

As the application is missing the functionality to fulfil the GDPR’s right for the patient to access their data that is stored, a new report template needs to be generated. Fortunately, the previously generated reports have been PDF’s which fulfils the widely used format requirement, so the only task is to design the template. The template should include the patient’s personal information and additionally collated data about their diagnosis data. Also, aggregated audit trail data concerning that data subject should also be included. Even though the GDPR does not explicitly require the inclusion of audit data,

it would not be surprising if the requirement stays at least in the Finnish legislation, as it was there previously. The audit trail information should include who had accessed the data subject's information, why it was accessed and when it was accessed.

As the previous reports have been available in a drop-down menu in a toolbar located on view where the specific patient's record is open, the new report will be a new item on the drop-down. That way it does not majorly affect the functionality of the application, and the user right requirement is satisfied.

The right of erasure requires a whole new feature for the application. The user of the application should not just be able to look up the patient and delete the whole record. The option for deletion should only be visible to user groups that have higher privileges and should not be functional remotely. Another possibility is that only the user who has a patient-doctor relationship can see and interact with the button.

For the functionality, one possible solution is a button on the patient viewscreen, in the main toolbar. Upon pressing the button, a new dialogue screen is opened, and the user is asked to confirm their intention by writing into the dialogue some piece of information of the patient, such as their individualising ID, which is their pseudonym used throughout the system. By inserting the correct information and then closing the dialogue, the patient would be moved into a deleted state. The patients who are in a deleted state could be collated into the same section of a view for example, and they would be searchable. Then in a required timeframe, another user could review the patients to be deleted and permanently make the deletion decision. This way the deletion has a minute change of being accidental and even considers a malicious insider. Naturally, a backend functionality must be developed so that the specific rows are removed concerning the deleted patient in the database.

The problem is the audit logs. These logs likely contain some mention of the erased data subject's data since a trace must be formed if these are changed. So, if the data subject's data is to be erased, then these log entries should also be removed. However, the problem is that these log files are essential and needed at least in the eyes of the legislature preceding the GDPR [18]. The GDPR does overwrite these acts, but there is no reason to think that the member states such as Finland would not introduce these measures back into the legislation. The patient's pseudonym is a good candidate for the logs, but then the question remains that how can the pseudonym be linked back to the data subject when the linking information has been deleted? The requirement of storing the logs means that the information erasure can never be complete and questions the need for such functionality existing altogether.

As was argued before there is a strong reason why this kind of functionality is useful, but it must be noted that the erasure functionality in the analysed application's case cannot be wholly used. As the GDPR does not require this functionality for this application's

context, the decision is left to the company developing this application. The potential problems presented here must be one of the reasons for the right of erasure's several exceptions. Then again, this contradiction can reveal a larger question from within the regulation on how the right of erasure can co-exist with other rights in the regulation and how can the erasing be done in information security considering way.

The right of restriction and the right to object can be handled with the on-hold-feature. However, as the feature was not initially used for this purpose, some modifications should be planned. The feature made it possible to write a reason why the patient was put on-hold that is displayed alongside the record. The patients that have exercised their right can be marked so by writing the reason to a note related to the record that was put on hold. As was mentioned in the present state of the application section, the patients put on hold due to the GDPR should be differentiated somehow so that other users do not accidentally add them back into the use in the application.

In the right to object requirement's context, the on-hold functionality needs to accurately differentiate the patients also between those who have exercised the right of restriction and right to object. The right to object is more permanent than the right of restriction since there is no mention in the GDPR of allowing processing of the data after the right has been exercised. The right to object can be countered with compelling legitimate grounds, and since this right could be considered harmful to the patient if exercised, then such a legitimate reason could be formed, and this right disregarded. However, as the GDPR remains vague regarding the exceptional cases and the compelling reason, it might be the best not to disregard the right to object outright.

Data in use protection improvements are difficult to assess considering the technical changes to the application. As was mentioned in the previous section, the protection of data in use is so dependent on the users of the application that small technical solutions are not enough. As the application logs actions taken extensively and a trace is left of the user's action so if something goes wrong, the culprit can at least be found.

Currently, even though the connection via HTTPS from the client to the web-server is a possibility and adequate level of security through encryption is implemented in the software system, the connection from said web-server to the database is unencrypted. This means that it is possible for an eavesdropper to listen on the connections from the web-server to the database. The situation can be improved with a TLS-connection (Transport Layer Security) that should with proper encryption algorithm give adequate protection for the database connections. .NET-framework supports this functionality natively, so this change does not require any significant functional changes in the application [31].

As for the data at rest, SQL Server provides suitable features. Microsoft's SQL server has a built-in feature called Transparent Data Encryption (TDE) which encrypts and decrypts

the database, backups and transaction log files in real time. It should not cause any problems with the application itself, so it can be turned on without requiring any changes. TDE does not directly encrypt the data tables, so if the user has permissions, they can see the whole database unencrypted. Instead, TDE protects the physical data files and log files so that if they would be moved to another location, the files are encrypted and cannot be used. [32]

As TDE is a relatively lightweight to take into use, it would increase the data at rest protection level with relative ease. It works with older SQL Server versions and the free Community version too, so compatibility issues should not be an issue even if a site has an older version of SQL Server installed. It would prevent someone from directly accessing the database files on the server machine but would not interfere with legitimate connections coming from the application, therefore increasing the confidentiality and integrity of the stored personal data. Whether a malicious individual deliberately going to the database server and removing storage devices and then trying to extract data from them is considered likely depends mainly on the site where this application is located, but with the TDE such a possibility can be eliminated with relatively small changes to the application architecture.

Another solution Microsoft offers is Always Encrypted that is designed for highly sensitive data which does apply to the application in question. Always Encrypted feature is much more complex and robust than TDE. In Always Encrypted a driver is installed to client computers, and that driver determines from previously configured settings which information is to encrypt and what is not and then sends those columns encrypted to the server, which then decrypts the encrypted data. [33] Always Encrypted requires a newer version of the SQL Server and is likely too complicated solution regarding the analysed application. It potentially requires structural changes to database tables and additional setup. As the internal network the application is in offers a fair amount of protection, Always Encrypted seems to be excessive.

7.5 The result of the analysis

The overall suitableness of the analysed application regarding the GDPR is satisfactory, and the changes required for the application can be considered modest. Several of the functionalities the application already had before the GDPR could be repurposed quite easily to fit the requirements derived from the GDPR. The presented change proposals contain room for deciding the best approach which means that every proposed change is not mandatory.

Every software application is unique and differs in the amount of work needed to make the application and organisation compatible with the new regulation. Medical applications may have an advantage over regular consumer applications and services since medical applications have exemptions in the regulation and the usage of professional medical

applications is different to consumer applications. Based on the analysis, the GDPR affects the privacy landscape on a grander scale than mere feature implementations. Feature implementations are still needed so that the applications would be GDPR compliant but are not what the regulation emphasises.

Related to the national level data protection laws discussed in Chapter 5, the GDPR brings new requirements for the data controllers and software systems, although the most apparent new requirements, the right of erasure or data portability, are not applicable in the analysed application's case. Most of the GDPR's requirements cannot be summarised into a short presentable notion since they are foundational. Still, among the requirements of the GDPR, there is an evident number of previously established ones. Personal data considering health is still deemed to be of high risk by default, and the applications must be designed mirror that risk level. With the proposed changes, the analysed application fulfils the requirements presented by the wording of the GDPR and the best estimate that can be given at the time of writing this thesis.

What the GDPR seems to mean for the data controllers handling health data is that the controllers need to think more about what personal data they collect, how they process it and how to handle the lifecycle protection of the personal data. Organisations need to consider the risks that may be directed to the data that they process, and the risk analysis should include the data itself and not merely the means of protecting it. This is especially real when operating under the explicit consent given by the data subject as that consent can be revoked.

Considering the health care industry, the GDPR does not bring any significant changes with it that would affect the industry specifically. The organisations need to be more careful than those entities that process data that has a smaller risk level, but that was the case before the GDPR. The regulation has included health care specific rules and exclusions to the rights the data subjects have so that entities that have legal requirements based on them from other national and broader levels of legislations can exempt themselves from fulfilling the rights. An example of this is the right of erasure that was also shown to be in contradiction with the requirement of storing log files even if somehow it was applicable in health care scope.

The security rules presented by the regulation are not by themselves new, but the level the GDPR attempts to apply the rules is. The data controllers need to think more about the ways to increase their information security level as the GDPR instructs in Article 32 such as data minimisation and using pseudonyms [13]. With these new requirements, the information security level also increases when the contents of the data itself are considered on an equal level with the improvements made to technical solutions like encryption or access control.

8. DISCUSSION

Based on the analysis of one medical software application, the GDPR does not offer anything ground-breaking regarding strict requirements. The medical software may not be the best to estimate the overall impact of the regulation since medical records have been thought of as sensitive data even before the GDPR. As such the GDPR's more substantial impact is likely is consumer software applications and services that may not have had such extensive functionality and operate on the public internet. The GDPR does seem to extend the idea of risk management into broader business activity, where data controllers and processors must think about why what and how they process data subjects' data.

The analysed application was already in use with several years of development behind it. The GDPR's idea of "data protection by design and data protection by default" is not realised since redesigning a whole application from the ground up is not feasible. However, this applies to every application developed before the GDPR came into effect, with varying results. Of course, the thinking model presented by the "data protection by design and data protection by default" can be integrated into new features that are added to the existing systems.

It can be argued that it is difficult to draw any universal conclusions from one application's scope. Then again as the regulation is so new, this same approach can be used to analyse other applications as well. Also considering medical software applications the points presented in this thesis may help when thinking about which parts of the GDPR are relevant to them and how to ensure a suitable information security level in the application.

Some aspects, especially regarding the user rights presented in the GDPR, were brought upon and analysed here that were not necessarily mandatory if the text of the regulation was to be interpreted concretely. Such is, for example, the right to object. As the personal data contained in the application analysed is medical but at the same time collected partly through the explicit consent of the patient the data controller must think about whether the right applies and how to justify that decision to those entities that ask about it.

Koops and Leenes [25] mentioned that purely technical solutions should not be the only solutions when thinking about GDPR compliance. Not relying entirely on technical solutions makes the compliance with the GDPR complicated from the data controller's perspective since organisational measures such as training should be planned and executed alongside the required technical solutions. The possible complications can be thought of as a good thing and likely is what the makers of the regulation intended. While the technical aspects of the application were analysed, and their states quantified, there is still a gap considering how the sites using the application conform to GDPR.

8.1 Limitations and future research

The newness factor of the GDPR is an obvious limitation to this thesis since legal precedents are bound to start taking shape, and the emphases of the GDPR will become more apparent and the regulation more straightforward to implement. That is part speculation and part hopefulness, but since the spirit of the law may sometimes differ from the detailed text, the best practices related to the regulation are likely to mould soon. Then less speculative analyses can be completed, and more concrete results derived from them.

The laws in this thesis were almost purely analysed based on the official texts. Laws usually have interpretations and precedents that are used when the laws are enforced. The absence of these interpretations is an obvious limitation of this thesis as the analysis could have been more thorough if the ways of interpretation were also analysed. The approach of this thesis still has merit though since the precedents of the GDPR have not yet been formed so legal texts are compared on the same level when only the texts themselves are analysed.

Another limitation is the almost purely technical viewpoint since the human factor of the sites cannot be accounted for. However, many companies face the dilemma of making the product they sell as good as possible considering data protection and do not have much knowledge of the specific environments where their customers use the software they have developed. There is a separation of concerns where both parties must think about which solutions fit them the best. For those integrating new software into their broader network of systems, a guarantee from the developers of the application that the application conforms fully with the GDPR might be an advantage on the competitive market in the future.

In the future, the analysis method of this thesis could be refined more. The ISO/IEC 25010 standard provides with the power mean a suitable basis for numerical analysis of the application's current state related to the requirements. The requirements derived from the GDPR could be further refined and multiplied so all the needed aspects of the regulation could be analysed. In this thesis, the needed requirements were not significant in numbers since the analysis considered a developing company and their application. An analysis of an implementation of the GDPR in one of the sites using an application like the one analysed in this thesis would surely be beneficial.

As the GDPR becomes the norm and first legal cases are resolved the actual impact of the regulation will become more evident. Same applies for member state legislations as they get refined and introduce member state relevant measures into the legislation alongside the measures of the GDPR. As future research, it would be interesting to see how the GDPR is enforced. The study of the enforcement is to see how the regulation has changed the way companies think about privacy or whether the GDPR has changed anything. One

development that is also interesting is how the regulation changes the legislations of non-EU countries.

It would be fascinating to study what kinds of organisational measures are developed to improve the overall data protection level in companies and other institutions. A study where the gaps left by this thesis would be analysed, such as data in use protections or how the data subject communication with the data controller would be standardised would benefit many other organisations.

8.2 Analysis of the proposed changes to the application

The changes required to the application are relatively small. The result is, of course, good news since the advent of the GDPR does not cause any significant financial requirements from the company developing this application.

The proposed changes raise the rating of the application towards the maximum number four, but a new analysis would be needed after these changes are reviewed an implemented to find out a new grade. All the user rights are considered however which means that the aim of the regulation is reached. As was mentioned before, not all the rights are necessarily applicable regarding healthcare applications

A question related to the more information security-oriented change proposals is that how secure an internal network can be deemed to be? Even with a high level of risk, an internal network provides a level of protection from outside threats. So, are the detailed encryption proposals necessary in the case of an internal network? The answer is not unequivocal because a private network can be thought to increase data protection level by itself so much that encrypted connections are unnecessary. Then again, the changes required for the implementation of wholly encrypted connections and encryption of data at rest is relatively easy and with that small commitment increase the confidentiality and integrity of personal data and other data contained in the application by a significant amount. Encryption is also mentioned as a security-enhancing method several times in the literature, so it should not be forgotten purely because of an internal network set-up.

The implementation of more encrypted connections does bring with itself the need for a strategy concerning encryption keys and certificates. As the HTTPS connection was already implemented in the application, some strategies have already had to be formed. This means that the management strategy should not have to be made up from scratch and the older guidelines could be modified to use. TDE's key management is straightforward according to Microsoft since the key is stored in the database boot record [32].

Considering the application with the proposed improvements included, it is fitting to analyse it regarding the ISO/IEC 25010 standard's [46] security sub-characteristics with availability included from the reliability characteristic. The proposed changes include

room for further analysis and there is a possibility that even though some changes would seem to improve the security level, a direct opposite may happen, or some other aspect of the application may be worsened.

Van der Haak et al. [22] mention secure connections and authorisation concepts for ensuring confidentiality. As the proposed connections and the database are encrypted, and access control measures are in place, the confidentiality of the information the application contains is guaranteed to be better than before. Although van der Haak et al. [22] discuss cross-institutional electronic patient records in their study, this scope is not much different.

Encryption also helps with ensuring information integrity and authenticity. As the used protocol for transmitting the data from the web-server to the database is SSL/TLS, a message digest is calculated ensuring integrity if a proper hashing algorithm is used. Authenticity is improved because the message is unlikely to change while the information is transmitted, and the information is as the user intended it to be. Access control also helps with increasing the level of authenticity.

Accountability and non-repudiation were in good shape already in the application as the audit logs and trail was extensive. These changes do not affect the state of these subcharacteristics. Availability should also not be affected, and TDE should not change the functionality of the application or make data unavailable for legitimate users. Availability is also affected by measures like backups and resilient systems which are out of the scope of this thesis.

Internal networks and encryption are not automatically compatible. Since the network is internal, the appropriate question is that are the presented encryption solutions necessary. Fauri et al. [16], while discussing internal networks in industrial control systems, analyse the benefits of adding encryption to a network already protected from the public. Fauri et al. conclude that encryption does not automatically yield extra security, that encryption can, in fact, have negative consequences for security and that encryption can increase maintenance costs. Although industrial control systems are a completely different field than the analysed medical application, they share the commonality of an internal network and that they are both information systems.

Many attacks target the end points of the connections, and in those cases, encryption does not help increase security. As Fauri et al. [16] and van der Haak [22] mention encryption helps ensure confidentiality on the wire. On the other hand, if the potential attack or other defect is directed at the end points, the confidentiality on the wire does not matter as the messages can be decrypted by using the encryption keys on the endpoints [16].

Encryption might present threats to security and hinder the functionality of solutions. Fauri et al. [16] mention that while encryption obfuscates the data from potential attackers, it also obfuscates it from legitimate tools that monitor the network. These tools might

without the encryption notice irregularities in the network traffic and be more effective in finding issues. The last reason for doubts considering encryption, Fauri et al. argue that encryption could increase the cost and complexity of troubleshooting and recovery. Sometimes the best way to check, for example, congestion related issues is to check package contents, which would not be possible if the packages were encrypted. This issue might be more relevant in the industrial control system space, but it is something that should be considered.

Fauri et al. [16] do not discard encryption in internal networks entirely but merely want to remind that it is not a silver bullet. In the analysed application, encryption should not be a silver bullet either. Considering the relative easiness of adopting an encrypted connection in the application, adopting a secured connection to the database is still recommended based on the analysis conducted in this thesis. The need for encryption might be a relevant issue to discuss with the sites using the analysed application as the means of monitoring network traffic probably differ between sites, and some sites might want to sacrifice encryption to monitor the network more effectively.

9. CONCLUSIONS

This thesis attempted to analyse the in parts open-ended nature of the GDPR to give more concrete examples on how the regulation would affect software applications that contain data concerning health. The analysed application was deemed to be in a satisfactory state considering the requirements derived from the GDPR, although improvements are needed to comply with the regulation.

Through the analysis of a medical software application in this thesis, it has been shown that while the General Data Protection Regulation contains novel concepts that will undoubtedly improve data subjects' privacy, its impact on medical applications cannot be considered extensive. Although the amount of work needed to be on an acceptable level of data protection and information security depends on the application, it can be said that if best practices and guidelines have been followed during development, then the regulation should not present a significant need for changes.

The GDPR compares itself well to the national legislations that preceded it. In several sections, the regulation is more detailed even though there are not any single huge changes. The GDPR extends the idea of risk management into the contents of the personal data that entities store and process, so that companies must think about why they store data and what risks might be directed to the data. The means detailed by the GDPR such as data minimisation are needed to mitigate the risks. Other, more technical measures are also needed, but they are not the only answer to the threats targeted at personal data.

Most of the data subject rights presented by the GDPR have already been detailed in previous legislations in Europe and outside of it. On the rights, like the right of erasure, the regulation details many exceptions and ways for the data controllers to deny the requests made by data subjects. Medical applications are given freedoms in these exceptions as the other legal requirements weigh heavily on balance.

While the impact of the GDPR might not be extensive on the medical field, it will undoubtedly still improve the data protection of the patients. The unifying element of the regulation will likely help in brining all the member states' information security and data protection legislations closer to one another. Open questions remain on how the regulation will be enforced in practice and how the data subject rights end up affecting different fields.

REFERENCES

[1] J. Andress, Basics of Information Security, 1st ed. Syngress, 225 Wyman Street, Waltham, MA 02451, USA, 2011, 171 p.

[2] C. Bier, P. Birnstill, E. Krempel, H. Vagts, J. Beyerer, Enhancing privacy by design from a developer's perspective, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 73-85.

[3] P.E. Black, I. Bojanova, Defeating Buffer Overflow: A Trivial but Dangerous Bug, IT Professional, Vol. 18, Iss. 6, 2016, pp. 58-61.

[4] J. Botha, M. Grobler, M. Eloff, Global Data Breaches Responsible for the Disclosure of Personal Information: 2015 & 2016, European Conference on Cyber Warfare and Security, 2017, pp. 63-72.

[5] Cantrell, David W. and Weisstein, Eric W, "Power Mean," [Online]. Available: From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/PowerMean.html [Accessed 22 October 2018]

[6] A. Cavoukian, A. Fisher, S. Killen, D.A. Hoffman, Remote home health care technologies: how to ensure privacy? Build it in: Privacy by Design, Identity in the Information Society, Vol. 3, Iss. 2, 2010, pp. 363-378.

[7] J.Q. Chen, A. Benusa, HIPAA security compliance challenges: The case for small healthcare providers, International Journal of Healthcare Management, Vol. 10, Iss. 2, 2017, pp. 1-12.

[8] Choi, H., Park, J. & Jung, Y. (2018). The role of privacy fatigue in online privacy behavior, Computers in Human Behavior, Vol. 81 pp. 42-51.

[9] Commission Nationale de l' Informatique et des Libertés, "ACT N°78-17 OF 6 JANUARY 1978 ON INFORMATION TECHNOLOGY, DATA FILES AND CIVIL LIBERTIES," [Online]. Available: https://www.cnil.fr/sites/default/files/typo/document/Act78-17VA.pdf [Accessed 30 September 2018]

[10] data protection legislation, in: A Dictionary of Computer Science, 7th ed., Oxford University Press, 2016, .

[11] privacy, in: A Dictionary of Computer Science, 7th ed., Oxford University Press, 2016, .

[12] A. Esteve, The business of personal data: Google, Facebook, and privacy issues in the EU and the USA, International Data Privacy Law, Vol. 7, Iss. 1, 2017, pp. 36-47.

[13] European Council, "General Data Protection Regulation (2016/679)." [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:119:FULL&from=EN [Accessed 9 May 2018]

[14] European Council, "Charter of Fundamental Rights of the European Union," [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012P/TXT&from=EN [Accessed 4 June 2018]

[15] European Council, "Consolidated version of the Treaty on the Functioning of the European Union" pp. 171-172 [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012E/TXT&from=EN [Accessed 7 June 2018]

[16] D. Fauri, B. de Wijs, J. den Hartog, E. Costante, E. Zambon, S. Etalle, Encryption in ICS networks: A blessing or a curse? 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm), IEEE, pp. 289-294.

[17] Federal Trade Commission, "Federal Trade Commission Act," [Online]. Available: https://www.ftc.gov/enforcement/statutes/federal-trade-commission-act [Accessed 1 August 2018]

[18] Finland's Ministry of Justice, "Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä," [Online]. Available: https://www.finlex.fi/fi/laki/ajantasa/2007/20070159#L5P17 [Accessed 14 September 2018]

[19] Finland's Ministry of Justice, "Personal Data Act (523/1999)," [Online]. Available: https://www.finlex.fi/en/laki/kaannokset/1999/en19990523_20000986.pdf [Accessed 8 May 2018]

[20] Finland's Ministry of Justice, "Sosiaali- ja terveysministeriön asetus potilasasiakirjoista (298/2009)," [Online]. Available: https://www.finlex.fi/fi/laki/ajantasa/2009/20090298 [Accessed 6 August 2018]

[21] Y. Flaumenhaft, O. Ben-Assuli, Personal health records, global policy and regulation review, Health policy, Vol. 122, Iss. 8, 2018, pp. 815-826.

[22] M. van der Haak, A.C. Wolff, R. Brandner, P. Drings, M. Wannenmacher, T. Wetter, Data security and protection in cross-institutional electronic patient records, International journal of medical informatics, Vol. 70, Iss. 2, 2003, pp. 117-130.

[23] D. Hofman, L. Duranti, E. How, Trust in the Balance: Data Protection Laws as Tools for Privacy and Security in the Cloud, Algorithms, Vol. 10, Iss. 2, 2017, pp. 47.

[24] M. Kemp, Barbarians inside the gates: addressing internal security threats, Network Security, Vol. 2005, Iss. 6, 2005, pp. 11-13.

[25] B. Koops, R. Leenes, Privacy regulation cannot be hardcoded. A critical comment on the 'privacy by design' provision in data-protection law, International Review of Law, Computers and Technology, Vol. 28, Iss. 2, 2014, pp. 159-171.

[26] D. Liebwald, Law's Capacity for Vagueness, International Journal for the Semiotics of Law, Vol. 26, Iss. 2, 2013, pp. 391-423.

[27] A. Liptak, Polar suspends its global activity map after privacy concerns, The Verge, [Online]. Available: https://www.theverge.com/2018/7/8/17546224/polar-flow-smart-fitness-company-privacy-tracking-security [Accessed 22 September 2018]

[28] I.M. Lopes, P. Oliveira, Implementation of the general data protection regulation: A survey in health clinics, 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), AISTI, pp. 1-6.

[29] B. Lundgren, N. Möller, Defining Information Security, Science and Engineering Ethics, 2017, pp. 1-23.

[30] J. Mai, Personal information as communicative acts, Ethics and Information Technology, Vol. 18, Iss. 1, 2016, pp. 51-57.

[31] Microsoft, “Enable Encrypted Connections to the Database Engine,” [Online]. Available: https://docs.microsoft.com/en-us/sql/database-engine/configure-windows/enable-encrypted-connections-to-the-database-engine?view=sql-server-2017 [Accessed 16 September 2018]

[32] Microsoft, “Transparent Data Encryption (TDE),” [Online]. Available: https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/transparent-data-encryption?view=sql-server-2017 [Accessed 26 September 2018]

[33] Microsoft, “Always Encrypted (Database Engine),” [Online]. Available: https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/always-encrypted-database-engine?view=sql-server-2017 [Accessed 26 September 2018]

[34] The Open Web Application Security Project (OWASP), “Top 10-2017 A5-Broken Access Control,” [Online]. Available: https://www.owasp.org/index.php/Top_10-2017_A5-Broken_Access_Control [Accessed 26 September 2018]

[35] The Open Web Application Security Project (OWASP), “Buffer overflows,” [Online]. Available: https://www.owasp.org/index.php/Buffer_Overflows [Accessed 26 September 2018]

[36] C.P. Pfleeger, S.L. Pfleeger, Security in computing, 4th ed. Prentice Hall, Upper Saddle River, NJ, 2007, 850 p.

[37] A. Pfitzmann and M. Hansen, “A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management,” August 2010. [Online]. Available: http://dud.inf.tu-dresden.de/Anon_Terminology.shtml. [Accessed 14 June 2018]

[38] Privacy Rights Clearinghouse, “Data Breaches,” [Online]. Available: https://www.privacyrights.org/ [Accessed 19 October 2018]

[39] J.H. Saltzer, M.D. Schroeder, The protection of information in computer systems, Proceedings of the IEEE, Vol. 63, Iss. 9, 1975, pp. 1278-1308.

[40] I.N. Shu, H. Jahankhani, The Impact of the new European General Data Protection Regulation (GDPR) on the Information Governance Toolkit in Health and Social Care with Special Reference to Primary Care in England, 2017 Cybersecurity and Cyberforensics Conference (CCC), IEEE, pp. 31-37.

[41] D.J. Solove, A Taxonomy of Privacy, University of Pennsylvania Law Review, Vol. 154, Iss. 3, 2006, pp. 477-564.

[42] D.J. Solove, INTRODUCTION: PRIVACY SELF-MANAGEMENT AND THE CONSENT DILEMMA, Harvard law review, Vol. 126, Iss. 7, 2013, pp. 1880-1903.

[43] D.J. Solove, W. Hartzog, THE FTC AND THE NEW COMMON LAW OF PRIVACY, Columbia law review, Vol. 114, Iss. 3, 2014, pp. 583-676.

[44] S. Spiekermann, L.F. Cranor, Engineering Privacy, IEEE Transactions on Software Engineering, Vol. 35, Iss. 1, 2009, pp. 67-82.

[45] P. Stirparo, I.N. Fovino, I. Kounelis, Data-in-use leakages from Android memory - Test and analysis, 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), IEEE, pp. 701-708.

[46] Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models, 1. ed. 2011-03-01 ed. ISO, Geneva, 2011, 34 p.

[47] D. Tavrov, O. Chertov, Evolutionary approach to violating group anonymity using third-party data, SpringerPlus, Vol. 5, Iss. 1, 2016, pp. 1-32.

[48] C. Tikkinen-Piri, A. Rohunen, J. Markkula, EU General Data Protection Regulation: Changes and implications for personal data collecting companies, in: Computer Law & Security Review, 2018, pp. 134-153.

[49] P. Tsormpatzoudi, B. Berendt, F. Coudert, Privacy by design: From research and policy to practice – the challenge of multi-disciplinarity, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 199-212.

[50] U.S. Department of Health & Human Services, "Health Insurance Portability and Accountability Act," [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/index.html [Accessed 1 August 2018]

[51] U.S. Department of Health & Human Services, "Health Insurance Portability and Accountability Act Privacy Rule," [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html [Accessed 1 August 2018]

[52] U.S. Department of Health & Human Services, "Health Insurance Portability and Accountability Act Security Rule," [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html [Accessed 1 August 2018]

[53] U.S. Department of Health & Human Services, "Health Insurance Portability and Accountability Act Enforcement Rule," [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/special-topics/enforcement-rule/index.html [Accessed 1 August 2018]

[54] J. Wang, Z.A. Kissel, Introduction to Network Security : Theory and Practice, 2nd ed. , Wiley, 2015, 418 p.

[55] S.D. Warren, L.D. Brandeis, The Right to Privacy, Harvard law review, Vol. 4, Iss. 5, 1890, pp. 193-220

[56] E. Wheeler, I., Security Risk Management: Building an Information Security Risk Management Program from the Ground Up, Syngress Media Incorporated, US, 2011, 340 p.

[57] H. Xu, J. Heijmans, J. Visser, A Practical Model for Rating Software Security, 2013 IEEE Seventh International Conference on Software Security and Reliability Companion, IEEE, pp. 231-232.