Implementing a data protection impact assessment for the web-application on the piloting phase

University of Oulu
Department of Information Processing Science
Master's Thesis
Jaakko Tikka
06.05.2020

Abstract

The General Data Protection Regulation (GDPR) contains several obligations for the ones that are processing personal data of the EU citizens. The major obligations are to take data protection by design and by default, and to carry out a data protection impact assessment (DPIA) whenever there is a high risk to breach privacy. Some organizations and companies are still struggling to achieve these obligations. Violating these obligations may cause sanctions that are up to 4% of the annual turnover. This created the motivation to research how these obligations should be implemented to achieve better compliance with the GDPR.

The objective of this thesis work was to research how the GDPR should be considered in applications that are processing personal data. Based on the related work, it was possible to recognize that DPIA process was recommended to cover the obligations of the GDPR. Therefore, the purpose was to research how the DPIA process would affect to the case application. Case application was a web-application that was on the piloting phase.

Design science research was applied as a research method. It was decided to carry out a DPIA by applying the guidelines of the Information commissioner's office (ICO). The DPIA process was applied to the case application. After the DPIA was completed, it was possible to evaluate its impact on the case application. Evaluation was completed in three parts, by evaluating how well the process of the DPIA covered the requirements of the GDPR, by evaluating the technical advantages and costs of the process, and by evaluating how the DPIA was applied in practice.

The results of this thesis showed that applying the DPIA process improved data protection, privacy and technical features of the case application. It was possible to reduce the privacy risks associated with data processing activities. In addition, DPIA process improved the technical side of the case application. The data model was simplified and unnecessary information flows were eliminated. These improvements were estimated to increase the workload of the developers for 2.7%. This meant that DPIA process was suitable way to cover the obligations of the GDPR.

Keywords
General Data Protection Regulation, Privacy impact assessment, Data protection impact assessment, Data protection by design and default, GDPR, PIA, DPIA

Supervisor
Dr. Jouni Markkula

Contents

1. Introduction

The General Data Protection Regulation (GDPR) was adopted in December 2015 by the European Union and came into effect on May 2018. The GDPR applies for all the organizations and businesses in the European Union and for those who process the data of the EU citizens despite the location of their processing. The purpose of the new legislation was to clarify data protection legislation because technological innovations and use of the internet have been increased since the previous legislation, which was adopted in 1995. (European Commission, 2016; Tankard, 2016.)

One of the major requirements of the GDPR is that it obligates data protection by design and by default. This means that organizations are required to ensure that data protection has been taken into account already in the design phase by implementing appropriate measures. Organizations are required to demonstrate these measures whenever requested by the authors. (European Commission, 2016; Tankard, 2016; Crutzen, Ygram Peters & Mondschein, 2019.)

One major reform concerns sanctions, which can result from non-compliance with the legislation. The highest sanctions are usually related to the accountability, these major breaches can cause fines that are up to 4% of the annual turnover and even the minor breaches can cause fines that are up to 2% of the annual turnover. These sanctions are highly motivating reason for the organizations to understand what actions are required to compliance with the GDPR and who is responsible for taking these actions. (European Commission, 2016; Tankard, 2016; Reetz, 2019; Crutzen, Ygram Peters & Mondschein, 2019.)

GDPR defines two key actors that are responsible to ensure that the processing is GDPR compliant (European Commission, 2016; Crutzen, Ygram Peters & Mondschein, 2019). Data controller is "*the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data* (European Commission, 2016)". Data controller is the main actor that is responsible for ensuring that the GDPR is complied. Data controller is also responsible to report any privacy violations to the authorities within 72 hours. (Diamantopoulou, Tsohou, & Karyda, 2019.) Data processor is defined in the GDPR to be "*a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller* (European Commission, 2016)". Data processors are the ones that are processing on the behalf of the Data controller. They are also responsible to report any privacy violations to the data controllers (Crutzen, Ygram Peters & Mondschein, 2019; Diamantopoulou, Tsohou, & Karyda, 2019).

The motivation for this thesis work is that the compliance with the GDPR is challenge for many organizations that are dealing with the data of the EU citizens, as they are struggling to adapt these requirements to the existing services. Studies have shown that some organizations have had to close down their existing services for while, because of the fear of the sanctions. (Shastri, Wasserman & Chidambaram, 2019). Therefore, there is a need to research how the compliance with the GDPR can be achieved.

This thesis is carried out by delving into the legislation and prior research to identify the required obligations and measures that are necessary for the applications and services that are processing personal data. After this, it is possible to utilize existing information to specify research problem and derive research questions. As a research method, design science research is applied to the existing case application by implementing appropriate measures to achieve compliance with the GDPR. Case application is a web-based application that is on the piloting phase. The initial design of the application is not GDPR compliant. This allows to evaluate how the initial design is affected and how compliance with the GDPR is covered. It is desirable that organizations and companies, which are struggling with the GDPR with their applications may benefit from the findings.

2. Related work

The purpose of the related work is to delve into the main points of the General Data Protection Regulation (GDPR). The emphasis is on finding the requirements that need to be considered in the software development and to find related research to support how to put the requirements into practice. Therefore, this chapter is focused on the legislation of the GDPR as its main information source, supplemented by relevant research on the subject.

The Article 1 of the GDPR states "*This Regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data.*" (European Commission, 2016). For the applications storing personal data, this means that there are several requirements to the rights of the data subjects and how their personal data can be processed. Most essential rights for the data subjects are that the data subjects have right to access and erase all information about them from the systems, and they have the right to know where their information is used and how it is processed. In addition, system operators are responsible to report any data breaches within 72 hours. For the actual processing activities, there are several other requirements, which determine what kind of obligations and measures should be applied while processing personal data. (European Commission, 2016, Shao & Oinas-Kukkonen, 2019.)

To ensure that the processing activities are compliant, GDPR has an obligation for data protection by design and by default, which is set out in Article 25 of the GDPR as it requires "*to implement appropriate technical and organisational measures*" for a data protection purposes (European Commission, 2016). Studies have mentioned different approaches to achieve compliance with the obligation data protection by design and by default, as Dewitte et al. (2019) mentioned in their research that legal persons apply more Data protection impact assessments (DPIA) while technical persons are more relying on a privacy and security threat models. However, Dewitte et al. (2019) and Sion et al. (2019) had similar results to point out that DPIAs may lack in technical perspective and in turn software engineering approaches may lack in compliance with legal requirements.

2.1 Processing personal data

To protect personal information and its processing, it is necessary to understand the concept of personal data and legislation in the area of processing the personal data. The GDPR defines personal data to be *"any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person"* (European Commission, 2016).

The definition of the personal data is comprehensive. The direct data is easier to identify to be a name or a social security number; the indirect data may be more complex. Any

indirect data that may lead to identify a particular person, such as images or sound, are categorized as a personal data. For this reason, data controllers should be careful when assessing what personal data is processed in their systems. (Kelli et al., 2019.)

For processing the personal data, GDPR has set seven data protection principles in Article 5. Each organization, company or individual that is processing a personal data in any purposes, is required to respect and follow these principles to compliance with the GDPR. These data protection principles are show in the following Table 1. (European Commission, 2016).

Table 1. Data protection principles (European Commission, 2016).

Principle	Statement
1. Lawfulness, fairness and transparency	Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject
2. Purpose limitation	Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes
3. Data minimisation	Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed
4. Accuracy	Personal data shall be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay
5. Storage limitation	Personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject
6. Integrity and confidentiality	Personal data shall be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures
7. Accountability	The controller shall be responsible for, and be able to demonstrate compliance

The principles 1-6 set the requirements for the processing of the personal data, the principle 7 (Accountability) can be considered as a fundamental principle that is obligated to cover these requirements. According to the GDPR, a data controller is responsible to demonstrate and take the actions to ensure, that these principles are met during the lifecycle of the processing. (European Commission, 2016.)

2.2 Data protection by design and by default

To cover the data protection principles, GDPR requires data protection by design and by default (DpbDD). This means, that a data controller is responsible to carry out and

demonstrate suitable mechanisms to ensure a data protection during the lifecycle of the data processing. As described in GDPR, Article 25, a data controller should be *"Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing*" (European Commission, 2016).

To understand the requirement for DpbDD, it can be separated in two parts. The requirement for a data protection by design is to ensure, that organizational and technical measures are applied already during the design phase and early on development phases to ensure a data protection. Data protection by default can be understood so, that these existing measures are not applied only in the design phase, but continuously during the lifecycle of the processing of personal data. (Romansky, 2019.)

To fulfill the requirements of the DpbDD can be a challenge, as Jasmonite et al. (2018) mentioned in their study that, albeit the idea behind the DPbDD is understood, the principles of the legislation are more complicated to implement and understand practically. Some studies emphasize, that the one should not confuse DpbDD and Privacy by Design approach even the objectives are similar. Privacy by Design is initially for purposes of data minimization, transparency, and to guarantee more reliable and trustworthy systems for users. Instead, DpbDD is more upgraded version and requirement to take whole organization and its processes into account to protect personal data. (Jasmontaite, Kamara, Zanfir-Fortuna & Leucci, 2018). However, Bincoletto (2019) showed in the study on Electronic Health Records, that Privacy by Design is a relevant approach to ensure DpbDD, as long as, it is reflected to the data protection principles.

To accomplish this requirement for DpbDD, study proposes to implement Data Protection Impact Assessments (DPIA) (Sion et al., 2019). Some organizations argue that investing in a privacy may be expensive and unnecessary in certain cases (De Francesco, 2019). Studies show that even most of the new projects have integrations with different legacy systems; therefore, it may be expensive and laborious to achieve compliance with the GDPR (Sarrat & Brun, 2018). However, it is good to understand that non-compliance may cause greater harm for the organizations, both financially and reputably. Therefore, it is advisable to invest in privacy manners and update privacy solutions to meet GDPR requirements. (De Francesco, 2019.)

2.3 Data protection impact assessments

By the GDPR Article 35**,** Data protection impact assessments (DPIA) are required if there is a possibility in the data processing to result "*a high risk to the right and freedoms of natural persons* (European Commission, 2016)*"*. The legislation lists these cases to be whenever there is a systematic collection and evaluation of natural person data in the system, whenever processed data contains sensitive data of the natural person or when system or organization operates in the publicly accessible field. (European Commission, 2016.) DPIA is a risks analysis, which purpose is to identify and analyze risks towards individuals, which may occur while using organizations systems. To demonstrate accountability, organizations are also required to present the DPIA to the authorities upon a request. (Bieker, Friedewald, Hansen, Obersteller & Rost, 2016). Albeit, the DPIAs may lack in technical perspective, they are more designed to ensure the compliance with the GDPR. The more technical approaches, such as threat models,

can be useful together with the DPIAs to add technical benefits if necessary. (Sion et al., 2019; Dewitte et al., 2019.)

Several studies reference and suggest applying guidelines from Working Party 29 (WP29) to carry out DPIA. Demetzou (2019) mentions that the purpose of the DPIA is to cover requirement for the accountability, which was presented as a Principle 7 in Table 1. Dewitte et al. (2019) adds that WP29 offers list of criteria that contains nine bullet points to ensure quality of DPIA. It is mentioned that DPIA should be implemented whenever two bullet points out of nine are met. However, it is explained that whenever organization is not sure, whether these criteria are met the consultation of the legal expert is recommended (Article 29 Data Protection Working Party, 2017). According to research of Sarrat and Brun (2018), organizations should not always systematically follow these criteria, since in some scenarios there might be eight criteria met and yet there is no real need for a DPIA. In turn, one criteria may be sufficient to indicate that DPIA is required.

Some organizations may not see a need to carry out a DPIA. However, WP29 recommends that it is a good practice in any case, as it can improve compliance with the law. If organizations are not willing to carry out DPIA, the reason has to be well documented. It is mentioned in the WP29 guidelines that DPIA should be continuous process and it should be re-produced after every three years. The process of DPIA should be started as soon as possible, even if some of the processes were still unknown. (Article 29 Data Protection Working Party, 2017.)

GDPR does not require any specific framework or process to implement DPIA. Instead, GDPR has set minimum requirements for an acceptable DPIA. These requirements are shown in the Table 3. (European Commission, 2016.)

Table 3. Minimum requirements for a DPIA (European Commission, 2016).

Requirement	Statement
1	A description of the envisage processing operations
2	An assessment of the necessity and proportionality of the processing
3	An assessment of the risks to the rights and freedoms of data subjects
4	The measures envisaged to address the risks and to demonstrate compliance with the regulation

Tikkinen-Piri, Rohunen and Markkula (2018) mentioned in their research that Privacy Impact Assesment (PIA) guidelines are a good starting point whenever organizations are starting to compile their DPIA documentations. However, they are also mentioning that organizations are able to create their own processes for this assessment. Similarly, WP29 recommends official EU generic frameworks such as PIA by the UK Information commissioner's office (ICO) and PIA by the French Commission Nationale de l'Informatique et des Libertes (CNIL) (Article 29 Data Protection Working Party, 2017).

2.4 DPIA by the UK Information commissioner's office

UK Information commissioner's office (ICO) provides an iterative process to carry out DPIA. This process consists of nine steps that are shown in the following Figure 1. (International Commissioner's Office, 2018.)

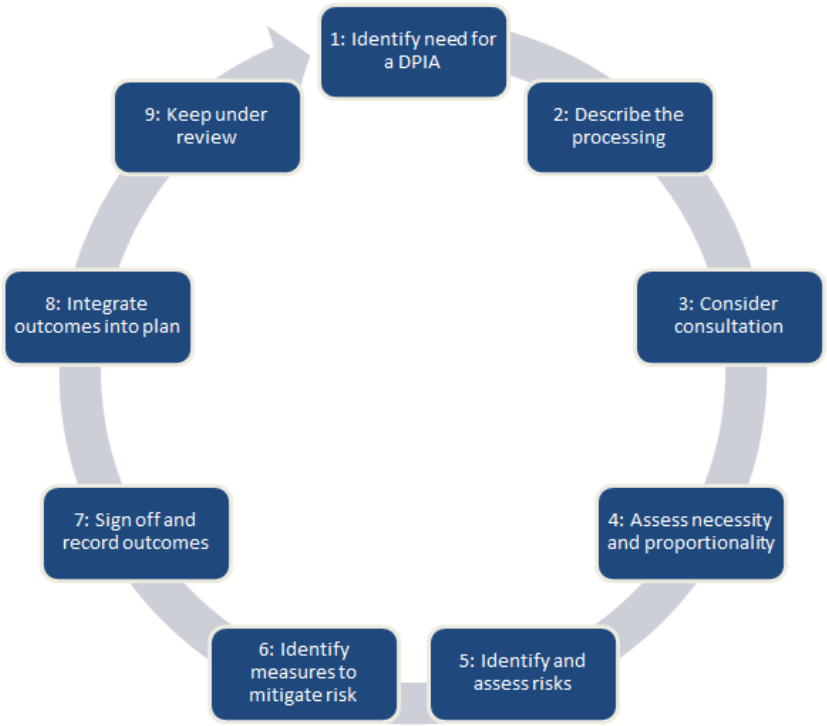

Figure 1. DPIA steps by ICO (International Commissioner's Office, 2018).

The first step is to identify the need for a DPIA. This means that at this step, it is required to describe what type of processing is involved in the project. The purpose is to utilize these descriptions to identify whether it is necessary to implement the DPIA. (International Commissioner's Office, 2018.)

The second step is to describe the processing. The purpose is "*to describe nature, scope, context and purposes of the processing* (International Commissioner's Office, 2018)". This means that at this step, it is required to describe how the data is being collected, stored, used and removed. It is recommended to visualize these descriptions in a flowchart or use some other appropriate method. (International Commissioner's Office, 2018.)

The third step is to consider consultation. This means that at this step, it is necessary to consider whether external or internal stakeholders should be consulted during the process. If so, it is required to describe when and how this happens. If there is no need for consultation, the reason must be justified. (International Commissioner's Office, 2018.)

The fourth step is to assess necessity and proportionality. The purpose is to assess whether the processing is compliant and proportional. In practice, this means that processing should be purposeful and lawful, including the data quality and data minimization should be ensured. If deficiencies are found, it is necessary to define how these are corrected. (International Commissioner's Office, 2018.)

The fifth step is to identify and assess risks. The purpose is to describe risks, and to assess how they will affect on individuals. It is required to assess likelihood, severity and overall status for each risk. The overall status can be either low, medium or high. Status is assessed based on the likelihood and severity as shown in the Figure 2.

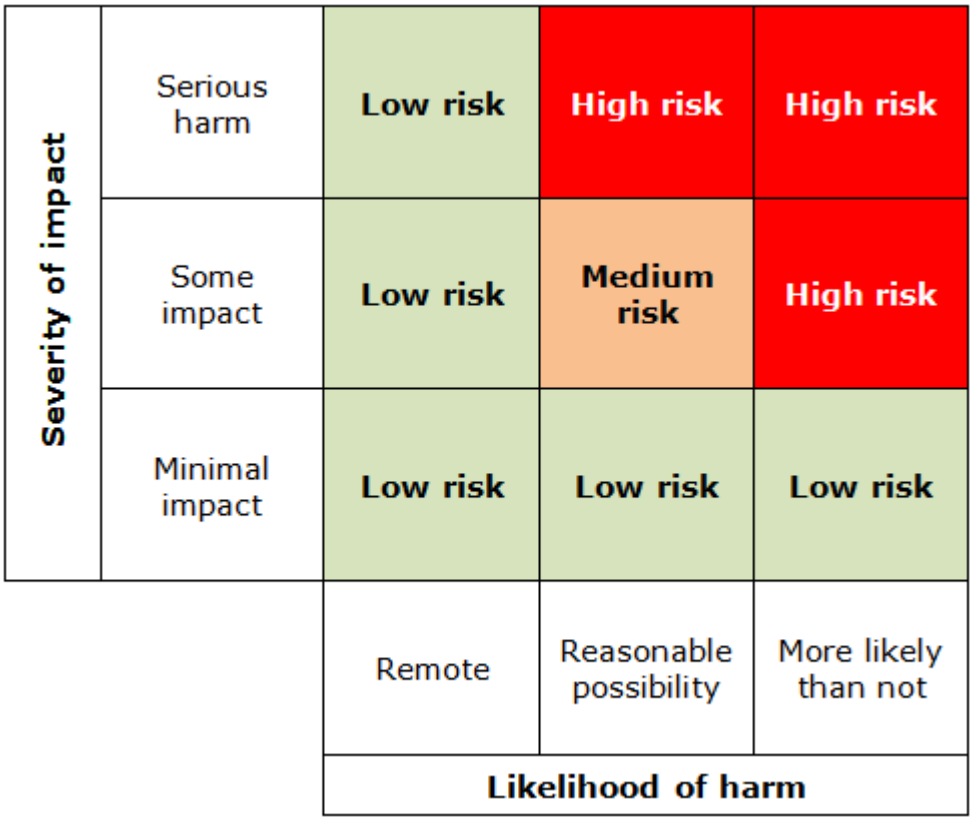

Figure 2. Risk assessment matrix by ICO (International Commissioner's Office, 2018).

The sixth step is to identify measures to mitigate risk. This means that medium and high risks must be either eliminated or reduced. If it is not possible to reduce or eliminate high risks, it is mandatory to consult the ICO for further instructions. In a similar situation, medium risks can be accepted. (International Commissioner's Office, 2018.)

The seventh step is to sign off and record outcomes. This means that each measure, which is identified in the sixth step is recorded, these are outcomes are integrated into project plan in the eight step. As the DPIA process should be continuous process, the ninth step is to keep DPIA under review and to repeat the process regularly. (International Commissioner's Office, 2018.)

2.5 DPIA by CNIL

Commission Nationale de l'Informatique et des Libertes (CNIL) provides a process to implement DPIA. This process consists of four main phases that are show in the following Figure 3. (Commission Nationale de l'Informatique et des Libertés, 2018).

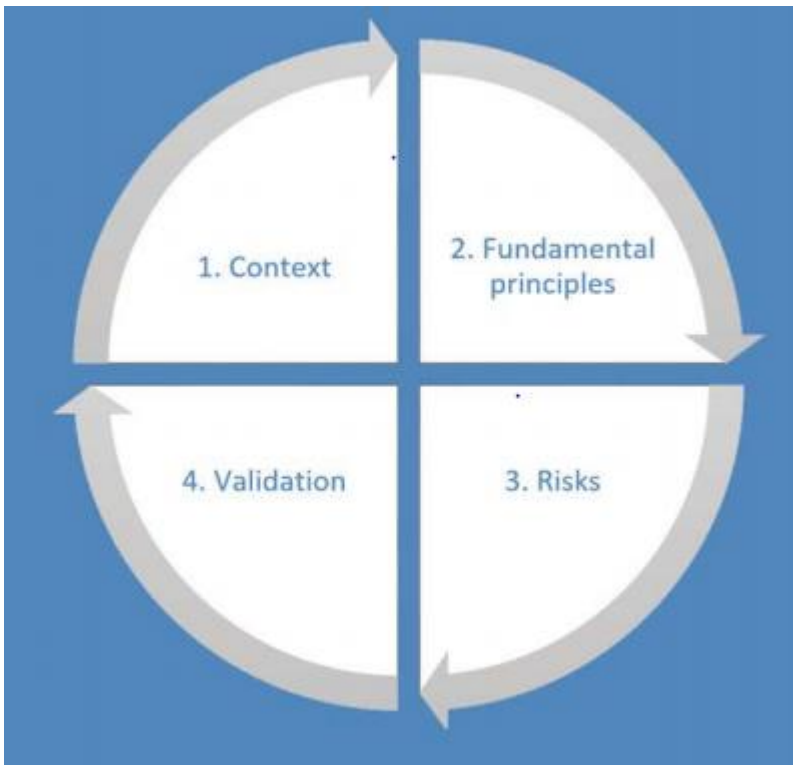

Figure 3. The structure of DPIA by CNIL (Commission Nationale de l'Informatique et des Libertés, 2018).

As in the ICOs DPIA process, DPIA by CNIL is started by describing processing operations. This is a common structure for DPIA processes. (Dewitte, Wuyts, Sion, Van Landuyt, Emanuilov, Valcke & Joosen, 2019.) The first phase to carry out DPIA by CNIL is a Context, it consists of two parts, by providing overview of activities where individual data is being processed and by describing *Data, processes and supporting assets* where the scope is described more accurately. (Commission Nationale de l'Informatique et des Libertés, 2018.)

The second phase is Fundamental principles. The goal is to make sure that the application is designed with the respect for data protection principles. (Commission Nationale de l'Informatique et des Libertés, 2018.).

The third phase Risks consists of two parts, *Assessment of existing or planned controls* and *Risk assessment: potential privacy breaches*. The goal of the first part is to understand the controls that are contributing for the security. Risk assessments purpose is to identify what is causing the risk and what are the potential effects of these risks. (Commission Nationale de l'Informatique et des Libertés, 2018.)

Last phase of the DPIA by the CNIL is Validation, the goal is to validate whether DPIA can be accepted. This phase consists of two parts, *Preparation of the material required for validation* and *Formal validation*. The purpose is to validate whether the DPIA can be accepted based on the findings of previous phases. (Commission Nationale de l'Informatique et des Libertés, 2018.)

2.6 Existing DPIA applications

In addition to the existing DPIA frameworks, CNIL provides free application that can be utilized to carry out DPIA. This application contains comprehensive information on

how DPIA should be implemented technically and legally. The advantages of the application are its modularity and visual tools that can be utilized to identify privacy risks. Source code is available for the users, which allows users to customize the application by adding new features or editing existing features. (Commission Nationale de l'Informatique et des Libertés, 2018.) The application is easy to use and it covers all the main steps to implement DPIA, therefore it is recommended for those who are not familiar with the process of the DPIA (Sarrat & Brun, 2018).

Other relevant applications that can be utilized to carry out DPIA are very industry-specific. As Piatowska et al. (2017) developed a tool which was designed for a smart grid systems and Alnemr et al. (2015) designed a tool for cloud. These tools have similarities, as they both are web applications, which are designed to help produce a proper DPIA. Basic structure of these tools is also similar, as they are based on questions, which are answered by the user who receives evaluation regarding the answers. These questions are based on ICOs Code of Practices. Each of these tools have pre-phase, which is meant to solve whether process should be continued. Even if tool introduced by Piatowska et al. (2017) seems to be a relevant choice when conducting DPIA, it suits mainly in purposes for smart grid systems. In turn, tool introduced by Alnemr et al. (2015) is more suitable for the general use, but it is not possible to access for this tool for further research.

3. Case application

Case application is a web-application that will enable inter-organizational interaction. System users can be individual actors from organizations or institution, which are seeking or offering business opportunities to other users of this ecosystem. By this system, users get possibility to find or share their own information and needs to others to find interesting match with other users or user groups.

Case application is designed to provide digital services with real-time access for the users of the system. System will require registration to the system, so that outsiders cannot join without approval. This will allow users of the system to create content in the system. Naturally, this content will be stored in the database, including these users contact and organization information.

Collected data contains of cases, organization information, contact information, tags and supplements. Cases are created by users, these cases can be any business need or project that creator is looking to start. Cases can be private or public, and case owner has responsibility for this. Organization and contact information from users are public for every user in the system. Organization information contains organization name, description, size and its contacts. Contact information requires contacts name, email address, phone number and organization. Supplements are comments and notes that users are able to write for organizations, public cases, or private cases when they are participated in the case.

There are two different access levels in the system. Admin users are able to access every content in the system, when basic users are only able to see all the public data that is visible in the system.

3.1 Project team, management and stakeholders

Project team consists of product owner, three software developers, one software architect and one UX designer. All members in the project team including pilot users perform software testing. Case applications stakeholders are pilot users, organizations and project members.

Project team uses Jira as a project management tool. This allows adding development tasks to the backlog. Each task are scheduled for the developers from the backlog. Scrum master is responsible for this action. The scrum master is one of the members of the project.

3.2 Stored data

Stored data consists of user accounts, contacts, organizations and cases. In addition, users are able to add comments and notes to the case application.

3.2.1 User account

User is one who has a registered user account in the system. One with the user account is able to create content in the system based on the security level. The following attributes were identified with the User Account: Username along with password captures the login credentials of the user. These are not accessible by anyone apart from the user him/herself and the users associated with System administrator security level. User email is a unique in the system. It is required information to make sure there are no duplicates in the system. User may receive confirmation or other requests from the system to their email address. User security level identifies the security clearance of the users, i.e. what entries they are able to access in the system. Security levels are divided in basic and moderator users. Basic users are able to use the system, they are able to create and see content in the system. However, they have limitations in the system. Moderator users have access to every content in the system and they are administrators in the system. Created identifies the user who created the user account and date when it was created. Modified identifies the user who most recently modified the user account and date when it was modified.

3.2.2 Contact

Contact is user accounts viable part in the system. This is the user that other users in the system are able to see. Contact information contains all the personal information from the individuals in the system. The following attributes were identified with the contact: Real name includes adding contacts first- and last name. Email address is same address that user gave to the user account in the registration phase. Phone number is an optional addition to the contact information. Organization is required for contact. Contacts can be added to existing organization in the system or they are able to create new organization in the system. Tags are required and attached to describe the contact, for example, with listing of expertise areas. Created identifies the user who created the contact and date when it was created. Modified identifies the user who most recently modified the contact and date when it was modified. Access security level defines who can edit contact profile.

3.2.3 Organization

Organization captures information related to any organization such as a company, university or funding organization. The following attributes were identified with the organization: Name of the organization. Contact info stores contact list for the organization. General info consists of description of the organization (required), size of the organization (optional) and location information (optional). Tags are collection of what the organization has to offer. These could be skills, expertise or experience that the organization wishes to make visible to others. Created identifies the user who created the organization and date when it was created. Modified identifies the user who most recently modified the organization and date when it was modified. Access security level defines who can edit organization profile.

3.2.4 Case

Case refers to a joint endeavor of several organizations for a well-defined purpose. One example of a case is a joint effort to build a new product for international markets. The

following attributes were identified with the case: Title is the name of the case. Description is a description of the case. Status refers to state of the case. Case statuses are new, active, last chance and closed. Owner is the user who is in charge of the case and controls the access rights for the case. The creator of the case becomes the first owner. Partners is a list of organizations that are participating in the case. Contacts is a list of contact that are participating in the case, these are contacts from the participating organizations. Revenue denotes potential business value of case. Validity period is a time period when case is open. Tags describe the case, for example listing the business niche associated with the case. Privacy means, that case can be either public or private. Owner can set case to private when only participants are able to see the content. Created identifies the user who created the case and date when it was created. Modified identifies the user who most recently modified the case and date when it was modified. Access security level defines who can access the case.

3.2.5 Supplement

Supplement is a data item – notes and comments. Supplements are associated with contacts, cases and organizations. The following attributes were identified with the supplement: Title describes the content of the supplement. Access security level restricts access to the supplement when necessary. Owner is the user who created supplement and is in charge of the supplement and controls the access right for the supplement. Created identifies the creator of the supplement and date when it was created. Modified identifies the user who most recently modified the supplement and date when it was modified. Content stores the actual content of the supplement. This could be a document or a link to a web page.

4. Research problem and methodology

The objectives of this thesis work are introduced, followed by design science research that is applied to achieve these objectives.

4.1 Research problem

As stated in the introduction, the objective of this thesis work is to research, how the GDPR should be considered in the case application that is on the piloting phase. Based on the related work chapter, it was possible to recognize that the GDPR has imposed a number of obligations on the processing of a personal data. Therefore, it is understandable that it is necessary to take a data protection by design and by default into account on the early stage of the project to ensure that processing complies with the GDPR. In order to research how to meet these obligations at the later phase of the projects, it was decided to implement a DPIA draft for the case application. Obvious reasons are that, as the case application contains personal data, GDPR has requirement to implement a DPIA and prior research argued that DPIA is a useful approach to cover obligation for a data protection by Design and by Default specifically from a legal point of view. Because the case application is on the piloting phase, it is possible to research what kind of technical advantages there are to implement a DPIA at this phase of the project and what are the costs of the possible changes.

Based on the related work it was possible to detect that there are no specific requirement for which framework should be applied while implementing a DPIA, instead there were existing guidelines from the official authors that are designed to meet the requirements of the GDPR. As the objective of this research is not to compare different approaches or frameworks, it is convenient to select to follow guidelines from the ICO. Motivation for this selection is that the structure of the ICOs DPIA process seemed to be clear, effective and easy to follow. In addition, based on the findings from the prior research each official guideline follows similar structure, therefore it is challenging to compare which framework should be applied and it would not add sufficiently value for this research.

By responding to the following research questions, it is possible to achieve the research objectives of this thesis work:

- How well compliance with the GDPR is covered by utilizing DPIA**?**
- What are the technical advantages and costs on implementing DPIA on the piloting phase?
- How the DPIA is applied in practice?

As the initial design of the case application did not cover the requirements of the GDPR, by answering to the first research question, it is possible to evaluate how effectively the process of the DPIA can be utilized to achieve the compliance with the GDPR. As the first research question focuses on the legal side, the second research question focuses on the technical side of the process. Therefore, by answering to the second research question, it is possible to identify the technical advantages of the process and to measure

the costs of the possible implications. By answering to the third research question, it is possible to evaluate how the DPIA process is applied in practice.

4.2 Research method

Design science research (DSR) is chosen as a research method, since it is appropriate method to the given research objectives. In the field of information systems research, Hevner et al., (2004) introduced seven guidelines to follow in the DSR, these guidelines are applied in this thesis work and the accomplishment is evaluated after the design process is completed.

The aim of the DSR is to produce a purposeful and viable IT artifact to solve organizational problems. This artifact can be a model, instantiation, method or construct that can be utilized in the design, development, use or analysis of the IT systems. (Hevner, March, Salvatore, Park & Sudha, 2004). To apply this research method, the purpose is to improve the case application by carrying out a DPIA draft. Therefore, the IT artifact is the case application itself. According to Hevner (2007) high quality of DSR contains three different cycles, where relevance cycle links the DSR processes to the environment, rigor cycle to the knowledge base and design cycle iterates with the implementation and evaluation of the IT artifacts. The objective is to follow the design, relevance and rigor cycles, as these cycles are presented in the Figure 4.

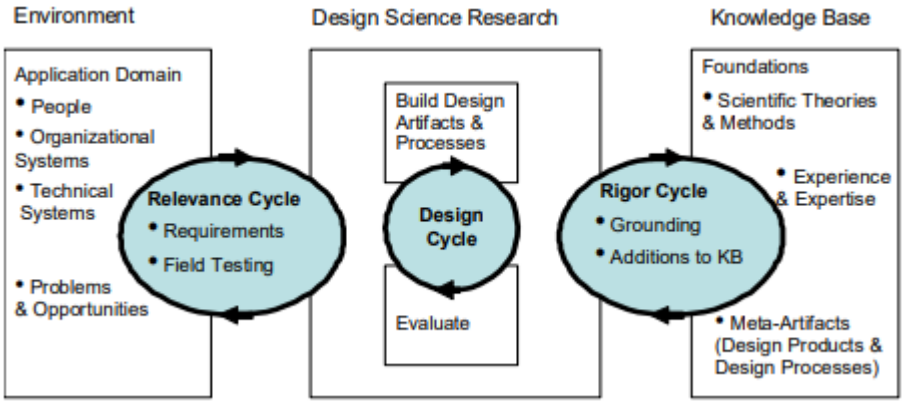

Figure 4. Design science research cycles (Hevner, 2007).

As the people, different organizational systems and technical systems create the environment itself, it is mentioned that environmental problems and opportunities are the ones that create the motivation (Hevner, 2007). In the context of this thesis work, there is a case application, which the project team is working on. The relevance cycle includes technical requirements for the case application. The case application is intended for business-to-business collaboration, which is why there are personal data processed. Therefore, the motivation is to improve the GDPR compliance of the case application.

Knowledge base consists of the prior research that is known on the subject-area. This includes known theories, experiences, expertise and different design processes. As in the context of this thesis work, the knowledge base contains the existing requirements of the GDPR and different processes to carry out a DPIA. As the rigor cycle that connects the knowledge base and DSR is an iterative process, the idea is to get familiar with the knowledge base and to utilize it in the study by creating value back to the knowledge base. This means, that in good DSR, the one does not only select existing methods or processes and design an artifact, but is able to add value to the knowledge base by

creating something new. (Hevner, 2007). The objective of the design phase is to add value to the knowledge base by evaluating how the process of DPIA is improving the compliance with the GDPR of case application on the piloting phase.

The design phase follows nine steps of the ICOs DPIA. These steps are applied on the case application in the following order:

Step 1: Identify a need for a DPIA
The objective is to identify a need for a DPIA. This step is performed by describing the personal data that is stored in the case application. Personal data is recognized by going through the content of each data object in the case application.

Step 2: Describe the processing
The objective is to describe the data processing activities involving personal data. Other project members are involved in this step. This step is performed by sending an email to each project member. In the email, each member are requested to send back their own descriptions of the information flows that concern personal data. By doing this, it is possible to find and describe accurately the most essential processing activities involving personal data.

Step 3: Consider consultation
The objective is to identify and describe the need for consultation during the development of the case application. This step is performed in collaboration with the internal stakeholders of the case application. Internal stakeholders are consulted in the meetings during the development process.

Step 4: Assess necessity and proportionality
The objective is to assess necessity and proportionality to the described data processing activities. This step is performed by making a decision on how to ensure the lawfulness and necessity of the processing activities in the case application. The decision is made in a discussion with the project team.

Step 5: Identify and assess risks
The objective is to identify risks towards individuals based on the results of the Step 2. Therefore, trivial risks are left outside. In order to identify and assess risks, other project members are consulted in the workshop and decisions made in Step 4 are utilized at this point.

Step 6: Identify measures to mitigate risks
The objective is to identify measures to mitigate risks that are identified in the Step 5. This step is performed in workshop together with project team.

Step 7: Sign off and record outcomes
The objective is to record outcomes that are resulted from the Step 6. This means that measures to mitigate risks are recorded in projects backlog.

Step 8: Integrate outcomes into plan
The objective is to integrate outcomes in to the project plan. This means that recorded outcomes from Step 6 are scheduled together with project team.

Step 9: Keep under review
The objective is to decide how the DPIA process should be continued. Decision is made together with project team.

After design phase is completed, the IT artifact can be evaluated as a whole. This means that it is possible to evaluate how the DPIA process affected to the case application and how the DPIA process practically succeed.

5. Design

Once the design phase was presented to follow a structure with 9 steps, the design process could be started.

5.1 Identifying the need for a DPIA

First step of the implementation was to identify the need for a DPIA. This was done by describing the entities that contained personal data. These entities were data objects that were stored in the SQL database and processed in the case application. These entities are presented in the following Table 4.

Table 4. Entities with personal data.

Entity (Data object)	**Properties (personal data)**
User account	• Username • Email address
Contact	• First name • Last name • Email address

As Table 4 demonstrates that there were two entities recognized with personal data properties. As the User account included properties such as username and email address, it was argued that these properties do not necessarily refer to a natural person, but in many cases username or email address may consists of the person's real name. The second entity with personal data properties was Contact. This entity was a clear choice, since contact information included properties such as person's first name, last name and email address.

Although, there were only two entities with personal data properties, several other entities that were stored and processed in the case application contained link to these entities. As each organization, case and supplement contained information about the creator and each organization and case had list of the involved contacts.

Identifying the entities with personal data was enough to prove that there were the need for a DPIA; therefore, it was decided to move to the next step to identify and describe the processing activities.

5.2 Describe the processing

To describe the processing it was important to identify all the key processes that included personal data directly or indirectly. This meant to identify all the processes that included contact or user account data that were described in the chapter 5.1. Based on the workshop it was possible to add each process to the list, results were naturally

similar with each member and there were multiple duplicates but each complemented the results. The outcome of the workshop is presented in the Table 5.

Table 5. Information flows of personal data.

Information flows	Description
1. Registering as new user	New user fills personal information in form. Information includes first name, last name, email address, organization, username and password. If information is valid and no duplicates are found from the system, new registration request is sent and stored to the system database with pending status and new user is notified that request is sent via email.
2. Accepting registration request	Admin user is able to accept registration requests from the system. All registration request related information is visible and editable for the admin user. Accepting request creates new user account and contact card in the system database. Pending user is notified that user account is created via email.
3. Rejecting registration request	Admin user is able to reject registration requests from the system. All registration request related information is visible and editable for the admin user. Rejecting request will be stored in the system database with rejected status. Pending user is notified that user account is rejected via email.
4. Managing security levels	Managing security levels means that admin can set user account role for basic or admin user. Basic user has limited access to system data.
5. Editing user account information	Admin user is able to access and edit all system user accounts in separate admin view.
6. Editing contact information	Admin user is able to access and edit all system contact details.
7. Adding new user account	Admin can create new user account to the system. It requires to fill user account information in form in separate admin view. New user account and contact card is stored to the system database.
8. Adding new contact	All system users are able to create new contact cards in the system. This includes to fill form with contact information.
9. Removing contact	Admin user is able to remove any contact information from the system, if contact has user account information, this information is deleted automatically.
10. Removing user account	Admin user is able to remove user account information from any contact from the system.
11. Adding tag to contact	Basic user is able to add tags to his/her contact card. Admin user is able to add or create new tags

	to any contact card in the system.
12. Removing tag from contact	Basic user is able to remove tags from his/her contact card. Admin user is able to add or create new tags to any contact card in the system.
13. Adding contact to organization	Any user is able to join to organization. Organization contacts are able to add existing contact cards to the organization. Admin user is able to add any contact to any organization.
14. Removing contact from organization	Organization contacts are able to remove other contacts from the organization. Admin user is able to remove any contact from any organization.
15. Adding contact to a case	Case owner and admin are able to add any contact to the case. After contact is added, contact is able to add his/her colleagues to a case. All contacts are able to join a public case.
16. Removing contact from a case	Case owner and admin is able to remove any contact from case. Basic user is able to remove own contact card from case.
17. Searching for contact	All system users are able to search for contact by using search field, which requires typing person name. User is also able to search tags, cases and organization.

As the Table 5 describes the identified information flows, it was seen that the outcome included several different cases where data is either added, removed or updated and most importantly, which user group has the right to carry out each process. Since user groups could be divided into admin users and public users with different authorization levels. After identifying the information flows, it was seen as good practice to create a flow diagrams from the most critical information flows to help illustrate the processes and to see if they add value to the existing information flows. The most critical information flows concerning personal data were identified to be in a registration process, therefore information flow 1. Registering as new user was illustrated first (Figure 5).

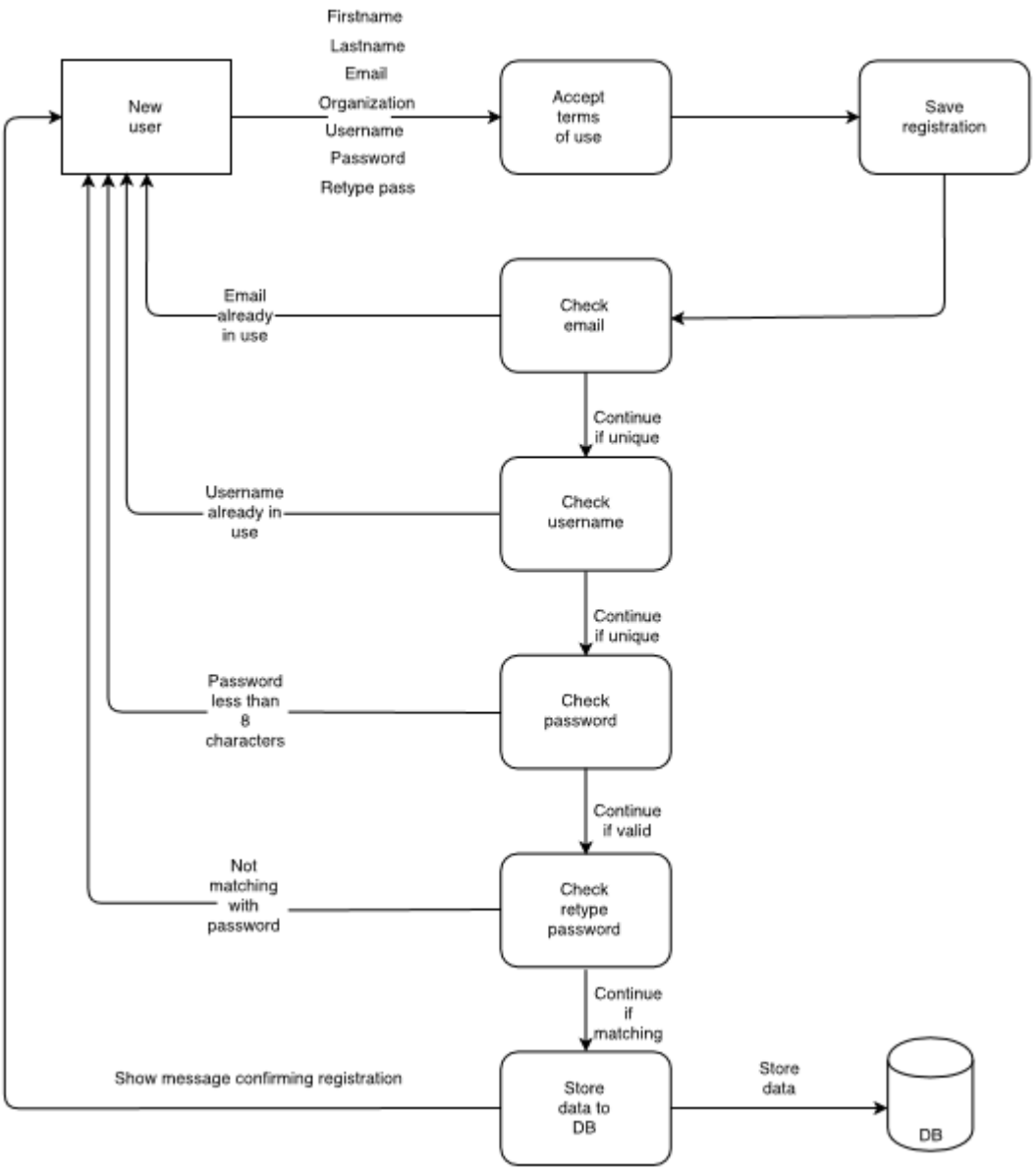

Figure 5. Flow diagram: Registering as new user.

As seen in the Figure 5. Registration process included accepting terms of use and adding valid information to the form. Email and username were checked in case of duplications before the new user could be saved in the database successfully. Also, password had to be minimum of 8 characters and it had to be retyped to match with the initial password.

As a continuum, when new user had sent acceptable registration request, the admin user had right to accept or reject the request (Fig 6). Therefore, it was seen as an important process to illustrate as a flow diagram.

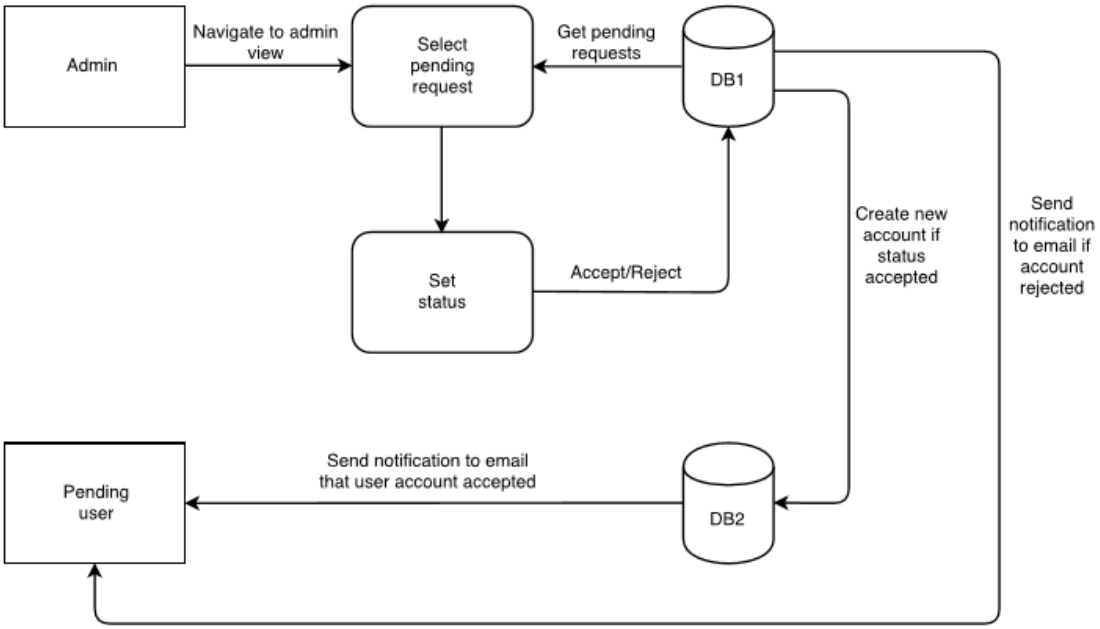

Figure 6. Flow diagram: Admin can accept or reject registration requests.

As presented in the Figure 6, admin user was able to see the list of the pending registration requests and either accept or reject the request. If the registration was accepted, new user account was created and stored in the database and notification from the accepted request was sent to the user email address. If request was rejected, notification from the rejected request was sent to the user email address.

In addition to the flow diagrams from the registration process, it was decided to merge three critical information flows as a one flow diagram. As Table 4 that was presented in the chapter 5.1 included two separate entities with personal data including user account and contact, admin user could edit and remove both entities. Therefore, information flows 6, 9 and 10 from the Table 5 (Fig 7) were illustrated to mainly describe the contact and user account removal process.

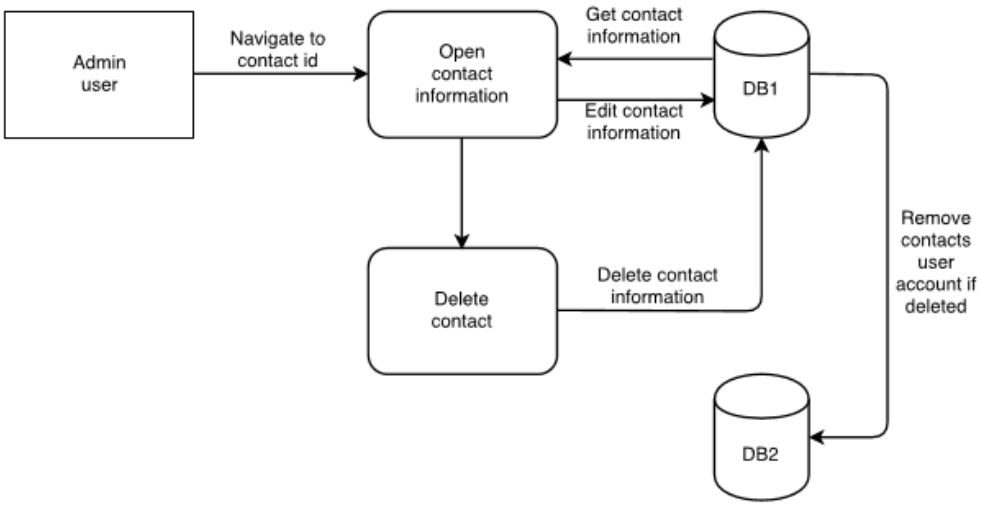

Figure 7. Flow diagram: Admin can edit and delete user accounts and contact information.

As presented in the Figure 7. The admin user had access to edit each contact card in the system, the process was seen important since whenever contact information was removed from the system, admin user was responsible to remove the user account information from the database.

After the information flows were described and the key processes were visualized it could be identified that it was possible to document processing by linking each information flow to the entities consisting personal data that were gathered in the Table 5 in chapter 5.1. To document and demonstrate this a table with three columns was created (Table 6). First column representing the name of the entity stored in the database, second column representing the chapter with specific description of the entity since case application was described more detailed in the chapter 3 and the third column representing the identified information flows that are concerning each entity.

Table 6. Information flows and descriptions of entities with personal data properties.

Entity	**Description**	**Information flows**
User account	User is one who has a registered user account in the system.	Registering as new user Accepting registration request Rejecting registration request Managing security levels Editing user account information Adding new user account Removing user account
Contact	Contact is user accounts viable part in the system.	Accepting registration request Editing contact information Adding new contact Removing contact Adding contact to organization Removing contact from organization Adding contact to a case Removing contact from a case Searching for contact

By adding each entity with personal data properties with specific description and then linking each information flow that is processing the entity, it was possible to perceive specific description of *the nature, scope, context and purpose of the processing* as it was required at this phase based on the literature. After the processing could be analyzed in more detailed, it was also possible to make assumption that since processing involved personal data, the need for a DPIA could be seen as a good practice to document processing and identify possible privacy risks related to a processing.

5.3 Consider consultation

Internal stakeholders were regularly consulted during the development of the case application. Demo sessions were arranged in every two weeks. The main purpose of these sessions was to introduce new features and changes of the case application to the internal stakeholders. Internal stakeholders provided feedback on the changes. Based on the feedback, it was possible to make improvements at an early stage. In addition, it was possible to detect design and development errors more effectively. Internal stakeholders were aware of the privacy policies. Therefore, the impact of each significant change on privacy could be assessed in larger group.

The demo sessions included a discussion on upcoming features of the case application. One of the upcoming features concerned adding appropriate content to the terms of use of the case application. The subject was discussed in demo session. The discussion concerned what should be included in the terms of use. As a result of the discussion, it was decided to consult a lawyer, because it was considered as a proper way to ensure legitimate content in the terms of use.

5.4 Assess necessity and proportionality

At this point, it was known that there were 17 information flows in the case application that processed personal data (Chapter 5.2, Table 5). A brief discussion was held with the project team. The topic of discussion was to decide how these 17 information flows should be assessed in order to ensure their necessity and lawfulness.

As a result of the discussion, it was decided that data protection principles should be utilized when identifying privacy risks in step 5. Data protection principles included all essential obligations for lawful processing, therefore, it was seen that this measure is the best option to ensure necessity and lawfulness of the processing activities.

5.5 Identify and assess risks

To identify and assess risks, the information gathered in the chapter 5.2 was utilized by evaluating each information flow and flow diagram to identify potential risks that may occur while processing. Each risk were added to the table (Table 7) that was applied from template of the ICO by adding separate columns for risk on individuals (ROI) and what data protection principle it violates. Therefore, data protection principles (Principle) that were presented in Table 1 in the chapter 2.1 were reflected to each risk. Corporate risks were left out of this step, since author did not have suitable schedule to evaluate them with the stakeholders. First column of the table was description of the risk, fourth column described the likelihood of the risk by adding each risk for either remote, possible or probable. Fifth column described the severity of the risk with the options minimal, significant or severe, depending on what would be the impact if the risk would be realized. The last column overall was evaluated based on the likelihood and severity of the risk, by adding overall status for low, medium or high.

Table 7. Identifying the privacy and related risks

Risk	ROI	Principle	Likelihood	Severity	Overall
1. Duplicated data in registration	Personal data is stored	Storage	Probable	Minimal	Medium

process	longer than purposes in multiple locations	limitation			
2. Duplicated user related data in the system (User account, Contact)	Redundant personal data is stored and that it is not accurate	Storage limitation	Probable	Significant	High
3. Too many users are able to modify personal data	Personal data is wrongly modified	Accuracy	Remote	Significant	Medium
4. Risk of allowing invalid (=too large) user access (role) allows access to service data	New user is given admin rights accidentally. Able to view, edit and delete all content of the system	Accuracy	Remote	Significant	Medium
5. Personal data is not accurate	Personal data is outdated	Accuracy	Possible	Severe	Medium
6. Personal data is processed in analytics/3rd party system	Personal data is leaked to outside system and used in purposes not originally intended	Lawfulness, fairness and transparency, purpose limitation, data minimisation	Remote	Significant	Low
7. User is able to create new contact cards in the system	Created contact might not be aware that he/she is created in the system	Integrity and confidentiality	Probable	Severe	High

The outcome of the risk assessment resulted to identify seven risks, each risk presented in the Table 7. First risks that were identified were related to duplicated data, even if they were similar they could be separated in two different risks. As it was initially identified based on the flow diagram that was presented in the Figure 4 that during the registration process the user account with status pending was stored to the own table in the database. Whenever admin accepted the pending user request, the data was copied to the other table with accepted user accounts. Therefore, it was seen that there is a risk to breach principle for storage limitation since initial data for the pending user account was no more needed and it shouldn't be stored longer than its purposes. Risk was seen to occur frequently and its likelihood was set to probable. The severity was seen minimal since it did not affect to the performance of the application and it did not have any negative impact on users. However, since it was seen to breach principle for storage limitation, the overall status was set for value medium.

The second risk for duplicate user related data was identified based on information flows 7 and 8. It was added to the table since it was recognized that whenever user account was accepted to the system, admin user could add and link separate contact information to the user account which was visible for the other public users. Therefore,

it was possible that there were already existing contact card in the system and it was seen that it could lead for more duplicates and redundant contact cards and that could cause that principle for storage limitation could be breached. The likelihood of this risk was set for probable with significant severity, therefore the overall status was set high.

Third recognized risk was, that too many users were able to modify personal data, causing that personal data might be modified wrongly. Based on the information flows 5 and 6, we could recognize that admin users were able to edit any contact and user account information in the system. This privacy issue could be trivial to every system, but it was recognized that there are too many users with admin rights, therefore this risk had to be listed, because it was seen that it might violate principle for accuracy. Therefore, the likelihood for this risk was set to be remote, but if the admin user would abuse the application the severity would be significant, even the probability was seen low, the overall status was set for medium. Fourth risk added to the table was a continuum to the third risk. Since, it was noticed based on the information flows 2 and 4 that whenever admin user accepts new registration request, there is a very small chance that new user is given admin rights accidentally. Likelihood for this scenario was seen remote, but similarly to the third risk, if this risk would realize the impacts could be significant, therefore, the overall status was set medium.

The fifth identified risk, personal data is not accurate, was derived from the risks 1 and 2 and partially from the information flows 7 and 8, which were presented in the Table 5. It could be recognized that user account and contact information are stored in the separate tables in the database. It may cause, that even if contact information were updated, the user account information might not inherit the information. Therefore, there can be outdated data in the system database. This could cause violating data protection principle for accuracy. Likelihood of the risk was set to possible and the severity for severe, therefore the overall status was set to medium.

Sixth recorded risk was that personal data is processed in analytics/ 3rd party systems. This risk was seen to be more general since there were no particular third parties involved, but since source code included using libraries provided by third parties the possibility should be taken into account. In case, that third parties would be utilized more in the future, it could cause that personal data is leaked to outside system and used in purposes not originally intended. Therefore, it could violate data protection principles for lawfulness, fairness and transparency, purpose limitation and data minimisation. As this risk was not seen threating according to circumstances, the likelihood was set for remote and the overall status for low, but in worst case scenario the severity would be significant.

Seventh risk was identified based on the information flow 8 (Table 5), as the public users were able to create contact cards to the system, therefore the created contact might not be aware that he/she is added to the system. This act was seen to violating principle for integrity and confidentiality, and therefore the overall status was set automatically high, as the likelihood was set for probable and severity to severe.

5.6 Identifying measures to mitigate risk

Once the risks were identified, it was possible to identify measures in the workshop to mitigate each risk. The results of the risk mitigation are shown in the following Table 8.

Table 8. Identifying the privacy and related risks

Risk	Options to reduce or eliminate risk	Effect on risk	Residual risk	Approved by
1. Risk of duplicated data in registration process	Remove unnecessary data	Reduced	Low	PO
2. Risk of duplicated user related data in the system (User account, Contact)	Merge user account and contact card concepts in the system	Eliminated	Low	PO
3. Too many users are able to modify personal data	Implement stricter access rules	Reduced	Low	PO
4. Risk of allowing invalid (= too large) user access (role) allows access to service data	Providing admin role for user should be separated from basic user account creation. Or add additional phase ("are you sure?") to UI process	Accepted	Medium	PO
5. Personal data is not accurate	Remove periodically unused personal data	Reduced	Low	PO
6. Personal data is processed in analytics/3rd party system	Do not allow access to personal data to 3rd parties, stricter access rules	Eliminated	Low	PO
7. User is able to create new contact cards in the system	Do not allow creating empty contacts in the system	Eliminated	Low	PO

Identifying the measures led to reduce overall status of each risk, as there were only one risk left with status medium. Also, three risks could be eliminated and three reduced, leaving one accepted risk. As there were no official DPO, each risk was approved by project owner (PO) as the changes were agreed with project team.

To mitigate the Risk 1, it was proposed that unnecessary data was being removed regularly. This meant that table that stored the user account requests were cleared for those user accounts that were already accepted or rejected. Therefore, this action would reduce the risk and the residual risk could be set low. For some technical reasons related to relational database, the risk could not be eliminated by removing each request when status was set for accepted or rejected.

To mitigate the Risk 2, it was decided to make major changes to the implementation, this meant that the user account and contact cards are not separate entities anymore,

since it was decided to merge these tables for the better performance. Therefore, this action would eliminate the risk permanently and the residual risk could be set to low.

To mitigate the Risk 3, it was proposed to implement stricter access rules, since too many admins were able to manage data in the system. This meant that there should be only necessary amount of admin users in the system and some admin rights could be removed. Therefore, the risk could be reduced and the residual risk could be set low.

To mitigate the Risk 4, it was proposed to either separate the authorization selection from the view where user account is being accepted to ensure that each new user accounts are added with basic user rights or to add confirmation whenever the user right is set to the admin rights. As the action was not yet decided, the risk was accepted for now and the status was left for medium.

To mitigate the Risk 5 for accuracy of the personal data, the proposed solution was similar as in the Risk 1, as it was decided that there will be periodical check to remove unused personal data from the system, this would not eliminate the risk, but it would reduce it, therefore the status could be set to low.

To mitigate the Risk 6, the proposed solution was to ensure that 3rd party systems that may gather information are not used in the system. As an example, the analytics tool that was utilized in the application was part of the in-house project. Decision to leave 3rd parties out eliminated the risk and the status was set to low.

To mitigate the Risk 7, it was proposed that the possibility to create empty contact cards without user accounts should be removed. Therefore, the risk could be eliminated and status could be set to low.

5.7 Sign off and record outcomes

Outcomes of the DPIA process were recorded to project backlog by adding each action that required further actions. This included risks that could be either reduced or eliminated, therefore accepted risks were not added at this step. The risk that concerned 3rd parties was left out, since it was seen to be more trivial risk and the risk had already overall status low without any measures. The recorded outcomes are presented in the following table 9.

Table 9. Recorded outcomes

Risk	**Options to reduce or eliminate risk**	**Effect on risk**
1. Risk of duplicated data in registration process	Remove unnecessary data	Reduced
2. Risk of duplicated user related data in the system (User account, Contact)	Merge user account and contact card concepts in the system	Reduced
3. Too many users are able to modify personal data	Implement stricter access rules	Reduced
4. Personal data is not accurate	Remove periodically unused personal data	Reduced
5. User is able to create new	Do not allow creating empty	Eliminated

contact cards in the system	contacts in the system	

As an outcome, five risks required further actions. By taking necessary actions, four of the risks could be reduced and one could be eliminated.

5.8 Integrate outcomes into plan

As the outcomes were recorded, it was possible to integrate them into project plan. This meant that each recorded outcome had to be analyzed separately. These outcomes are presented in the following table 10.

Table 10. Adding outcomes into plan

Action to be taken	Date for completion of actions	Responsibility for action
1. Remove unnecessary data	To be decided in sprint planning	Development team / Scrum master
2. Merge user account and contact card concepts in the system	To be decided in sprint planning	Development team / Scrum master
3. Implement stricter access rules	To be decided in sprint planning	Development team / Scrum master
4. Remove periodically unused personal data	To be decided in sprint planning	Development team / Scrum master
5. Do not allow creating empty contacts in the system	To be decided in sprint planning	Development team / Scrum master

As the actions were recognized to be technical tasks and they required more detailed design and specifications before they could be implemented, they were integrated into project plan by scheduling each action from projects backlog. Each action were added with information “To be decided in sprint planning” and responsible author was set to development team/ scrum master, as it was decided to add them to a projects natural development cycle in the upcoming sprint planning.

5.9 Keep under review

After the design phase was completed, it was necessary to decide next steps to maintain the case applications compliance with the GDPR. It was decided that the DPIA process should be repeated after each action (Table 10) are implemented. This would ensure that the case application stays GDPR compliant in the future.

6. Evaluation

After design phase was completed, it was possible to evaluate the DPIA as a whole. To measure the success of the DPIA and its impact on the case application, the evaluation was completed in three parts.

Compliance with the GDPR was evaluated by analytically evaluating how the process of the DPIA covered the data protection principles. Technical advantages and costs of the process were evaluated by evaluating, how the process of the GDPR affected to the initial design of the application and by estimating the costs of these implications. The implementation of the DPIA process was evaluated by evaluating the success of each step of the DPIA process.

6.1 Evaluation against the data protection principles

To achieve GDPR compliant application it was necessary that data protection principles were respected during the lifecycle. As in the design process, risks were identified by reflecting the information flows on the data protection principles. Therefore, it was possible to evaluate the efficiency and success of the design process, by analysing how well the process itself helped to ensure that each principle were being followed.

Based on the 17 information flows that were described in the design phase, it was possible to summarize, that there were 7 risks identified that would either violate one or more data protection principles (Principle).

Principle 1: Lawfulness, fairness and transparency
Risk(s): Personal data is processed in analytics/3rd party system
Solution: Eliminated
Evaluation: It was seen that using 3rd party systems would have minor potential to breach Principle 1. As disclosing information to third parties could cause unfairly or unlawfully usage of the personal data, therefore this risk was eliminated by creating own in-house tool for the analytics to minimize third parties.

Principle 2: Purpose limitation
Risk(s): Personal data is processed in analytics/3rd party system
Solution: Eliminated
Evaluation: It was seen that using 3rd party systems would have minor potential to breach Principle 2. As disclosing information to third parties could cause further processing of the personal data, therefore this risk was eliminated by creating own tool for the analytics to minimize third parties.

Principle 3: Data minimisation
Risk(s): Personal data is processed in analytics/3rd party system
Solution: Eliminated
Evaluation: It was seen that using 3rd party systems would have minor potential to breach Principle 3. As disclosing information to third parties could cause irrelevant

usage of the personal data, therefore this risk was eliminated by creating own tool for the analytics to minimize third parties.

Principle 4: Accuracy
Risk(s): Too many users are able to modify personal data, Personal data is not accurate
Solution: Reduced
Evaluation: Two risks were seen to have potential to breach Principle 4. As a first, too many users were able to modify personal data. Therefore, this risk was reduced by implementing stricter access rules for the users. Second identified risk was that personal data is not accurate. Risk was reduced by periodically removing unused personal data.

Principle 5: Storage limitation
Risk(s): Risk of duplicated data in registration process, Risk of duplicated user related data in the system (User account, Contact)
Solution: Reduced, Eliminated
Evaluation: Based on the information flows it was possible to recognize risks "Risk of duplicated data in registration process" and "Risk of duplicated user related data in the system (User account, Contact)". These risks were found to have potential to breach Principle 5. Risks were either reduced or eliminated by removing unnecessary data and by merging user account and contact card concepts in system.

Principle 6: Integrity and confidentiality
Risk(s): User is able to create new contact cards in the system
Solution: Eliminated
Evaluation: Based on the information flows it was possible to recognize the risk "User is able to create new contact cards in the system". This risk was found to have potential to breach Principle 6. Risk was eliminated by preventing creating empty contacts in the system.

Principle 7: Accountability
Risk(s): -
Solution: Continuous process
Evaluation: As the data protection was not taken into account by design, the appropriate technical and organizational measure to ensure data protection by default was to carry out appropriate DPIA. Therefore, by keeping DPIA up to date it is possible to ensure, that appropriate measures are applied during the lifecycle of the project.

Based on the evaluation against the data protection principles, it was possible to recognize that the process of the DPIA was successful in the legal perspective. As the process helped to identify risks towards data protection principles, each of these risk were either eliminated or reduced. Therefore, through the process it was possible to ensure and demonstrate the GDPR compliance as required in the data protection principle for accountability. This also meant, that the requirements for the DPIA and Data protection by Design and by Default were covered with the assumption that the process remains continuous.

6.2 Evaluation of the technical advantages and costs

After the design phase it was possible to confirm, that five risks required further actions to be either reduced or eliminated. This meant, that these planned actions affected to the architecture and initial design of the software and they required re-design, development

work and testing. To evaluate these changes, it was possible to evaluate technical advantages and costs by analyzing how each action affected to the software. Advantages were evaluated in terms of technical improvements. Costs were calculated by estimating how many working days each actions required from the developer, and by calculating their total percentage of the total workload of the project which was 540 person working days. The calculation was considered from the developer's point of view because previous workload for developers were available. Also, it was seen that this was the most practical way to measure technical costs in the current situation.

Action: Remove unnecessary data
Evaluation: Removing unnecessary data due the registration process was seen to require an adjusted procedure for cleaning the rejected registration requests from the database. Therefore, it was estimated to require 1 day of the planning and preparations and 2 days of the actual development work. The Advantages are that the database is cleared regularly from the unnecessary data; therefore, data is not stored longer than its necessary.
Cost: 3 person working days

Action: Merge user account and contact card concepts in the system
Evaluation: The process of merging two entities in the system was seen as a major change in the application, as it effects on multiple information flows. This action requires changes from the database to client. Therefore, the estimation of the workload was estimated to require total of 10 person working days, as it was estimated to require 2 days of planning, 6 days of development work and 2 days of regression testing. The technical advantages of these changes could be considered high, as it simplifies data model by allowing to maintain, store and process two major entities in same schema. Also, the amount of the information flows could be reduced within this action.
Cost: 10 person working days

Action: Implement stricter access rules
Evaluation: Implementing stricter access rules was estimated to require 5 person working days, as it was estimated to require 1 day of planning, 2 days of development work and 1 day of testing. The advantages are that the amount of the information flows could be reduced within this action.
Cost: 4 person working days

Action: Remove periodically unused personal data
Evaluation: Similarly to removing unnecessary data, the unused contact cards could be removed from the database with manual script. As the merging user account and contact card concepts and not allowing to create empty contact cards would remove the problem eventually, therefore the estimation for required person working days was not possible at the time. The advantages of this action were that the database was cleared from the unnecessary data to avoid storing personal data longer than it was necessary.
Cost: Could not be estimated

Action: Do not allow creating empty contacts in the system
Evaluation: Preventing creating empty contact cards was estimated to require total of 3 working days, requiring 1 day of planning and specifications, 1 day of development work and 1 day of testing. The advantage on preventing creating empty contact cards reduced the amount of information flows.
Cost: 3 person working days

To evaluate the technical advantages and costs, it was possible to recognize that several technical changes were required to the application. These changes could be demonstrated to reduce amount of the information flows, personal data and more coherent data model. Therefore, these could be counted as technical improvements. The costs were directly related to the workload, as they were estimated to require total of 20 person working days. This result was a sum of the four actions that were estimated, as one action could not be estimated. It was calculated that 15 days of these involved developer as these required development and design work. The estimated total workload for the development work was estimated to be 540 person working days. Therefore, it was calculated that the technical changes increased the total workload for the developers for 2,7%.

6.3 Evaluating the implementation of the DPIA process

After the design phase, it was possible to evaluate the practical implementation of the DPIA process. Each step from the DPIA process was evaluated separately. The purpose of the evaluation was to assess whether the practical implementation was successful, partially successful or unsuccessful. This made it possible to identify good and bad practices.

Step 1: Identify a need for a DPIA
Status: Successful
Evaluation: It was possible to recognize that user account and contact information included personal data. This was sufficient information to identify the need for a DPIA. By going through content of each data item was an effective way to recognize what personal data was processed in the case application.

Step 2: Describe the processing
Status: Successful
Evaluation: Involving other project members helped to find all essential information flows that included personal data. There were 17 information flows described. Based on them, it was possible to describe flow diagrams from the most critical information flows concerning personal data.

Step 3: Consider consultation
Status: Successful
Evaluation: Consultation was obtained when necessary. Internal stakeholders were consulted during the demo sessions. The need for external consultation was identified in demo sessions. As a result, lawyer was consulted for terms of use.

Step 4: Assess necessity and proportionality
Status: Successful
Evaluation: Decision was made together with project team to identify and assess risks by reflecting to data protection principles. This decision proved to be successful. This measure improved the necessity and lawfulness of the case application.

Step 5: Identify and assess risks
Status: Successful
Evaluation: Risks were identified in the workshop together with the project team. For each information flow that was described in the Step 2, it was examined what data protection principle it might violate. This proved to be successful, as it was possible to identify 7 potential privacy risks.

Step 6: Identify measures to mitigate risks
Status: Successful
Evaluation: Workshop was arranged to mitigate risks that were identified in the Step 5. In this way, it was possible to assess what technical changes were required to reduce each risk. As a result, each risk could be reduced.

Step 7: Sign off and record outcomes
Status: Successful
Evaluation: Each measure to mitigate risk that was identified in the Step 6 were recorded in projects backlog. As a result, a description of each measure was stored in the projects backlog.

Step 8: Integrate outcomes into plan
Status: Partially successful
Evaluation: Outcomes were recorded to the project backlog and total workload of each outcome was assessed. However, they were not scheduled properly into project plan.

Step 9: Keep under review
Status: Partially successful
Evaluation: It was possible to make decision how to continue with the process in the future. However, the plan to continue was neither scheduled nor finalized.

Evaluation of each step from the DPIA process showed that seven steps were successful. This meant that the practical implementation of these steps was success. Two steps were partially successful.

7. Discussion

The objective of this research was to study how the compliance with the GDPR could be achieved for the non-compliant application that was on the piloting phase. Design science research was applied to achieve the research objectives by implementing a DPIA to improve the GDPR compliance of the case application. The design process was structured to follow design phases that were based on the guidelines of the ICOs DPIA. The success of the design process was then evaluated by evaluating how the data protection principles were covered in the process. This made it possible to analyse how the process of DPIA could cover the requirements set by the GDPR and to find possible technical advantages that the process itself could cause. The technical advantages and costs were evaluated by evaluating how the process of the DPIA affected to the application on piloting phase. The practical success of the DPIA process was also evaluated.

How well compliance with the GDPR is covered by utilizing DPIA? It is possible to achieve compliance with the processing activities required by the GDPR by following the steps of this research as presented in the design phase. Based on the related work, it was possible to understand that the main obligation from the GDPR for the web-applications was to ensure Data protection by Design and by Default. This meant that data protection principles had to be respected during the lifecycle of any project or application, which was processing personal data. To achieve this obligation, studies proposed to utilize DPIAs, as DPIA process itself was recommended to cover the seventh data protection principle. DPIA was carried out by following the steps of the framework by the ICO to research how the compliance with the GDPR was covered. Each risk were identified by reflecting information flows and data flows to the data protection principles. By doing this, it was possible to ensure and evaluate that the data protection principles were respected in the application. Therefore, the DPIA process was successful as it fulfilled its purpose to eliminate and mitigate the data protection risks. In addition, by respecting data protection principles, the design process made it possible to achieve the obligation for data protection by design and the obligation for data protection by default can be achieved by maintaining and repeating the process as instructed.

Naturally, the DPIA process did not cover all the requirements obligated by the GDPR directly. The GDPR requires organizations and companies to allow data subjects to remove any data, which of them is stored in the system, and to take necessary steps to ensure that data subjects are aware of how their personal data is being processed. In addition, GDPR requires organizations to report any privacy incidents within 72h. These obligations require measures outside the DPIA process.

What are the technical advantages and costs on implementing DPIA on the piloting phase? The process of the DPIA was not primarily recommended to improve technical side of the application, as the prior research recommended different approaches such as privacy and security threat models; however, in addition to legal advantages it was possible to recognize that because of the process, some technical improvements could be detected. As the measures that were required to cover the compliance with the GDPR

led to reduce the amount of the information flows, the amount of processing and storing of personal data, and to create a more consistent data model.

The costs of these implications were 2.7% increase in total workload for the developers; there were no relevant studies to compare this finding, since prior research mainly expressed that the process may be expensive in certain systems. However, it is possible to assume that, if the current web-application had included any system integrations, these workloads would have increased. It would have emphasized that the Data protection by Design and by Default approach was not followed in the initial design.

As the process was implemented at the piloting phase, it is fair to speculate, that 2.7% increase in the workload for developers is not significant. It required some major changes in the initial design. In comparison to building a new GDPR compliant system from a scratch, the process proved that it is reasonable option to follow the steps of the DPIA to achieve compliance in the processing activities also at the later stage. In industry, an increase of 2.7% in the workload for the developers may be significant. The customer may expect that these requirements are already taken into account early in the project and these may be unnecessary expenses to them. However, in many cases, the process of the DPIA is a mandatory assessment; therefore, fines followed by non-compliance of the GDPR are greater than the costs of the process.

The benefits of the DPIA could be measured in both legal and technical sense, as it specifically improved technical processing activities of the application, which resulted better compliance with GDPR. When new applications are implemented, they usually contain strict specifications, comprehensive user stories, and technical descriptions, and naturally, workload estimates are based on these. However, the actual technical implementations often differ from these, as there may be several solutions to accomplish the task or there may be flaws in a specifications. Therefore, by implementing and maintaining DPIA during the development can reveal several breaches at an early stage and save extra work later.

How the DPIA is applied in practice? It is possible to carry out a DPIA by following the practices of this research. This research followed the guidelines of the ICOs DPIA. Those guidelines were applied to the case application successfully.

The first phase of the design process was to identify the need for a DPIA. It was seen that it was not possible to identify whether the case application had "*a high risk to the right and freedoms of natural persons* (European Commission, 2016)*"*. However, it was possible to confirm that the case application processed personal data. This was easily accomplished by going through each data object in the case application. This should be a sufficient reason to carry out a DPIA, because risks with status high were identified at a later stage in the DPIA process.

Describing the information flows was efficient way to describe the personal data processing activities. Members of the project were asked to describe their versions of the information flows. This made it possible to find all the essential information flows. Based on information flows, it was possible to illustrate most essential processing activities to flow diagrams. This helped to identify risks effectively at later steps. Reflecting information flows to the data protection principles proved to be effective and simple way to identify privacy risks. It is possible that implementing more flow diagrams would have resulted to identify more privacy risks. Therefore, in a more complex application, the need for flow diagrams would increase.

Internal consultation was received throughout the development process. The demo sessions were appropriate forum to discuss about the privacy issues during the development process. Therefore, there was no significant need for external consultation. However, external consultation and audit should be considered more in large-scale applications.

Outcomes of the DPIA process were recorded in the projects backlog as development tasks. This practice is necessary to maintain the DPIA process. This ensures that the outcomes end up in the project plan.

7.1 Design science research guidelines

The seven design science research guidelines introduced by Hevner (2004) were followed during the design process. The completion of these guidelines is presented below.

Guideline 1: Design as an Artifact
A viable IT artifact was produced in the form of a GDPR compliant case application. Case application was improved during the DPIA process in both legal and technical sense.

Guideline 2: Problem relevance
The research problem was to understand how the process of the DPIA would help to meet the obligations under GDPR. Therefore, it was natural to approach the problem by producing a DPIA draft.

Guideline 3: Design evaluation
The evaluation of the artifact was done analytically. Efficacy of the artifact was possible to rigorously demonstrate by analyzing how the process of the DPIA covered the data protection principles to improve GDPR compliance in the case application. Also, the implications were evaluated by analyzing the technical advantages and the cost of the implications. As a third, the success of the DPIA process itself was evaluated.

Guideline 4: Research contributions
The contribution was the case application and the understanding how the DPIA process itself can be utilized to result a GDPR compliant application.

Guideline 5: Research rigor
The research followed the guidelines of the design science research. The knowledge base provided methods that were applied in the design process. The design was then evaluated based on the requirements that were discovered from the knowledge base.

Guideline 6: Design as a search process
The knowledge base of the research was utilized to find suitable methods to carry out design process. This meant finding viable methods that were utilized to reach research objectives.

Guideline 7: Communication of research
The research was structured and presented so that both technical and legal entities are able to utilize the findings of the research.

7.2 Limitations

The design cycle consisted of one iteration round. There were significant evidence that by reflecting to the data protection principles, one design cycle improved compliance with the GDPR. However, it is not possible to demonstrate that one iteration round could lead to identify all possible privacy risks. For this reason, an additional round would have yielded more accurate results.

Similarly, one design cycle was not enough to evaluate technical disadvantages. The evaluation of the technical advantages was open to interpretation. There were no previous point of comparison to evaluate technical advantages. This limited the results of the research.

The increased total workload for developers was based on an estimate. Estimations were accurate, potential challenges were taken into account. However, the actual result may differ from the estimation.

The approach to the research problem ignored the comparison of the different DPIA frameworks. This was a justified decision since there were several existing studies that were focusing on comparing existing frameworks. However, by creating own framework or merging existing frameworks could have add value to the research.

7.3 Future research

Several directions for future research were found. They were derived from the limitations and findings of this research, by constructing new approaches to the research problem.

It would be valuable to research how much DPIA process increases the total workload in different areas of the projects. Current studies have pointed out that achieving GDPR compliance may be expensive. One of the findings of this research showed that the process of DPIA increased total workload for developers by 2.7%. There were no relevant studies, which would have shown similar results. Related studies were focusing on budgetary implications of the DPIA process. These findings would be generally valuable, because different organizations may have different cost structures.

Combining the best aspects of different DPIA processes should be researched. It is possible to use the basis of this research to compare ICOs DPIA framework to the other existing frameworks. By merging the best aspects of each process, it would be possible to construct a more effective DPIA process.

It should be researched how to construct hybrid frameworks from the DPIAs and more technical privacy and security threat models. These hybrid frameworks would increase transparency in project teams, and would ensure that technical and legal solutions go hand in hand.

8. Conclusions

The objective of this research was to study how the requirements of GDPR had to be taken into account when developing a web-application in the piloting phase. It was possible to identify that the GDPR contained a number of requirements that required measures. The most essential requirements were the obligation for data protection by design and by default, and the obligation to carry out DPIA. The related work was utilized to find best practices to implement these obligations.

To meet the requirements of the GDPR, it was decided to carry out a DPIA for the case application by following the guidelines of the ICOs DPIA. The purpose was to improve the case applications compliance with the GDPR via DPIA process. After the DPIA was completed, it was possible to evaluate its impact on the case application. Evaluation was completed in three parts, by evaluating how well the process of the DPIA covered the requirements of the GDPR, by evaluating the technical advantages and costs of the process, and by evaluating how the DPIA was applied in practice.

Implementing a DPIA improved the compliance with the GDPR. The process improved several processing activities in the case application that contained privacy risks. This was achieved by reflecting processing activities to the data protection principles during the design process. This made it possible to implement measures to reduce and eliminate identified privacy risks. This also led to technical improvements, as the data model became more consistent and unnecessary information flows containing personal data were reduced. The measures to achieve the compliance with the processing activities increased the total workload for the developers for 2.7%.

The process of the DPIA improved data protection, privacy and technical features of the case application. For this reason, the DPIA process is an efficient tool to achieve compliance with the GDPR. The DPIA process should be started as early as possible, but the findings show that it is possible to achieve compliance with the GDPR with reasonable amount of work also in later phases.

References

Alnemr, R., Cayirci, E., Dalla Corte, L., Garaga, A., Leenes, R., Mhungu, R., & Tetrimida, K. (2015, October). A data protection impact assessment methodology for cloud. In *Annual Privacy Forum* (pp. 60-92). Springer International Publishing.

Bieker, F., Friedewald, M., Hansen, M., Obersteller, H., & Rost, M. (2016, September). A Process for Data Protection Impact Assessment Under the European General Data Protection Regulation. In *Annual Privacy Forum* (pp. 21-37). Springer International Publishing.

Bieker, F., Martin, N., Friedewald, M., & Hansen, M. (2017, September). Data Protection Impact Assessment: A Hands-On Tour of the GDPR's Most Practical Tool. In *IFIP International Summer School on Privacy and Identity Management* (pp. 207-220). Springer, Cham.

Binns, R. (2017). Data protection impact assessments: a meta-regulatory approach. *International Data Privacy Law*, *7*(1), 22-35.

Bincoletto, G. (2019, June). A Data Protection by Design Model for Privacy Management in Electronic Health Records. In *Annual Privacy Forum* (pp. 161-181). Springer, Cham.

Commission Nationale de l'Informatique et des Libertés. 2018. Privacy Impact Assessment (PIA). https://www.cnil.fr/en/privacy-impact-assessment-pia

Commission Nationale de l'Informatique et des Libertés. 2018. The open source PIA software helps to carry out data protection impact assesment. https://www.cnil.fr/en/open-source-pia-software-helps-carry-out-dataprotection-impact-assesment

Crutzen, R., Ygram Peters, G. J., & Mondschein, C. (2019). Why and how we should care about the General Data Protection Regulation. *Psychology & health*, 1-11.

De Francesco, G. P. (2019). The General Data Protection Regulation's Practical Impact on Software Architecture. *Computer*, *52*(4), 32-39.

Demetzou, K. (2019). Data Protection Impact Assessment: A tool for accountability and the unclarified concept of 'high risk'in the General Data Protection Regulation. *Computer Law & Security Review*, 105342.

Dewitte, P., Wuyts, K., Sion, L., Van Landuyt, D., Emanuilov, I., Valcke, P., & Joosen, W. (2019, April). A comparison of system description models for data protection by design. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing* (pp. 1512-1515). ACM.

Diamantopoulou, V., Tsohou, A., & Karyda, M. (2019). From ISO/IEC 27002: 2013 Information Security Controls to Personal Data Protection Controls: Guidelines for GDPR Compliance. In *Computer Security* (pp. 238-257). Springer, Cham.

European Commission. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing the Directive 95/46/EC (General Data Protection Regulation). Official Journal L119, 04/05/2016; 2016b

Gellert, R. (2018). Understanding the notion of risk in the General Data Protection Regulation. *Computer Law & Security Review*, *34*(2), 279-288.

González, E. G., & de Hert, P. (2019, April). Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles. In *ERA Forum* (Vol. 19, No. 4, pp. 597-621). Springer Berlin Heidelberg.

Hoofnagle, C. J., van der Sloot, B., & Borgesius, F. Z. (2019). The European Union general data protection regulation: what it is and what it means. *Information & Communications Technology Law*, *28*(1), 65-98.

International Commissioner's Office (ICO). 2018. How do we carry out a DPIA? https://ico.org.uk/for-organisations/guide-to-the-general-data-protectionregulation-gdpr/data-protection-impact-assessments-dpias/how-do-we-carryout-a-dpia/

Jasmontaite, L., Kamara, I., Zanfir-Fortuna, G., & Leucci, S. (2018). Data protection by design and by default: Framing guiding principles into legal obligations in the GDPR. *Eur. Data Prot. L. Rev.*, *4*, 168.

Kelli, A., Lindén, K., Vider, K., Kamocki, P., Birštonas, R., Calamai, S., ... & Gavriilidou, M. (2019, May). Processing personal data without the consent of the data subject for the development and use of language resources. In *Selected papers from the CLARIN Annual Conference 2018, Pisa, 8-10 October 2018* (No. 159, pp. 72-77). Linköping University Electronic Press.

Kovacs, Z. (2019). The impact of the GDPR on data privacy experience.

Lovell, M., & Foy, M. A. (2018). General Data Protection Regulation May 2018 (GDPR) How does it affect us? *Bone & Joint 360*, *7*(4), 41-42.

Piatkowska, E., Bajraktari, A., Chhajed, D., & Smith, P. (2017). Tool support for data protection impact assessment in the smart grid. *e & i Elektrotechnik und Informationstechnik*, *134*(1), 26-29.

Reetz, M. (2019). GDPR: Does Coverage Exist for Fines and Penalties for Noncompliance? *TortSource*, *21*(3).

Renault, A. (2019). Designing a Data Protection Process Assessment Model Based on the GDPR. In *Systems, Software and Services Process Improvement: 26th European Conference, EuroSPI 2019, Edinburgh, UK, September 18–20, 2019, Proceedings* (p. 136). Springer Nature.

Romansky, R. (2019). A Survey of Informatization and Privacy in the Digital Age and Basic Principles of the New Regulation. *International Journal on Information Technologies and Security*, *1*(11), 95-106.

Sarrat, J., & Brun, R. (2018, June). DPIA: How to Carry Out One of the Key Principles of Accountability. In *Annual Privacy Forum* (pp. 172-182). Springer, Cham.

Shao, X., & Oinas-Kukkonen, H. (2019, April). How does GDPR (General Data Protection Regulation) affect persuasive system design: Design requirements and cost implications. In *International Conference on Persuasive Technology* (pp. 168-173). Springer, Cham.

Shastri, S., Wasserman, M., & Chidambaram, V. (2019). The Seven Sins of Personal-Data Processing Systems under GDPR. *USENIX HotCloud*.

Sion, L., Dewitte, P., Van Landuyt, D., Wuyts, K., Emanuilov, I., Valcke, P., & Joosen, W. (2019, March). An Architectural View for Data Protection by Design. In *2019 IEEE International Conference on Software Architecture (ICSA)* (pp. 11-20). IEEE.

Tankard, C. (2016). What the GDPR means for businesses. *Network Security*, *2016*(6), 5-8.

Tikkinen-Piri, C., Rohunen, A., & Markkula, J. (2018). EU General Data Protection Regulation: Changes and implications for personal data collecting companies. *Computer Law & Security Review*, *34*(1), 134-153.

Wilker, S., Meisel, M., Piatkowska, E., Sauter, T., & Jung, O. (2018, June). Smart Grid Reference Architecture, an Approach on a Secure and Model-Driven Implementation. In *2018 IEEE 27th International Symposium on Industrial Electronics (ISIE)* (pp. 74-79). IEEE.