GDPR Privacy by Design

From Legal Requirements to Technical Solutions

Salah Addin ElShekeil & Saran Laoyookhong

Department of Computer and Systems Sciences

Master's degree project of 30 HE credits
Degree subject: Computer and Systems Sciences (Information Security)
Spring term 2017
Supervisor: Fredrik Blix
Reviewer: Paul Johannesson
Swedish title: GDPR: Integritet genom Design

Stockholm University

GDPR Privacy by Design

From Legal Requirements to Technical Solutions

Salah Addin ElShekeil & Saran Laoyookhong

Abstract

Respect for human rights and freedom are fundamental elements of any democratic society. In today's world, digitisation is contributing and making our life easy, however, at the same time poses a lot of threats to the fundamental rights of individuals. Privacy should be preserved and respected for individuals regardless of being in physical space or cyberspace. The European Union has issued the new general data protection regulations which will be enforced by all European states starting from May 2018. This research aims to translate the legal requirements of GDPR provision 25 (Data Protection by Design and by Default) into technical solutions by building a framework using the design science as the methodology. The proposed framework consists of three phases and is evaluated by a case study on an artificial intelligence application, ChatBot. Upon applying the framework, the IT system shall be compliant, and with the individual rights preserved, the society shall be flourished.

Keywords

Data Protection Principles, GDPR Compliance, Privacy by Design

Abstract - Svensk översättning

Respekt för mänskliga rättigheter och frihet är grundläggande faktorer i alla demokratiska samhällen. Nuförtiden bidrar digitaliseringen till ett enklare liv, trots det utgör den även ett hot mot mänskliga rättigheter. Människans integritet måste bevaras oberoende om det är i det verkliga livet eller på Internet. Europeiska Unionen har publicerat den nya allmänna dataskyddsbestämmelsen som kommer att verkställas av alla europeiska länder från och med maj 2018. Syftet med den här studien är att förvandla lagkraven av GDPR provision 25 (Data Protection by Design and by Default) till tekniska lösningar genom att bygga ett ramverk med användningen av design och tillämpning av vetenskaplig metodik. Förslaget på ramverket består av tre faser och utvärderas av en fallstudie via en artificiell intelligens-applikation vid namn ChatBot. Vid appliceringen av ramverket ska IT-systemet vara kompatibelt med GDPR provision 25, och med de individuella rättigheterna bevarade ska samhället blomstra.

Nyckelord

Data Skydd Principer, GDPR Eftergivenhet, Integritet genom Design

Acknowledgement

We would like to thank our supervisor Dr Fredrik Blix for his constant and endless support during our thesis journey. We are also grateful to Dr Paul Johannesson for reviewing our work.

We would like to thank our internship supervisors for providing us with the opportunity to write a thesis with business and technical team of ChatBot.

Last but not least, we would like to thank our families for their support and understanding throughout writing this thesis.

Synopsis

BACKGROUND	In today's world, digitisation is transforming everything around us to contribute to better quality of life. However, it also created the privacy issue which threatens the fundamental rights of individuals. One of the initiatives to tackle this issue is the introduction of Privacy by Design foundation principles. European Union introduced new regulations General Data Protection Regulations (GDPR) to provide EU citizens with privacy rights and introduced obligations and penalties on entities who process personal data. Part of this new regulations is the requirements to implement privacy measure into IT systems which are the topic of this research study.
PROBLEM	One of the requirements of all public and private sectors who process citizen's data is the implementation of Data Protection by Design and by Default in IT System. Affected parties will be subjected to substantial penalty if violated, however, the knowledge on how to achieve compliance is very limited.

RESEARCH QUESTION	There have been several attempts to operationalize Privacy by Design in IT systems including the old principles introduced by both Ann Cavoukian and Peter Schaa. However, there are no works conducted to help entities to comply with the new GDPR. This research tries to answer the following questions:
	• How should the European General Data Protection Regulation (GDPR) Privacy by Design by Default principles concretely be implemented in IT systems?
	  o What are the legal requirements relating to the Privacy by Design by Default principles?
	  o How can technical measures be mapped to these legal requirements?

METHOD	This research uses design science as a research method and conduct case study after building the artefact to evaluate the artefact. The creative method and five activities in design science: explicate problem, define requirements, design and develop artefact, demonstrate artefact, and evaluate artefact were followed to develop a comprehensive framework for translating GDPR into IT system requirements. The primary data collection methods include conducting interviews to understand the problem and existing system as well as performing literature review in parallel throughout the research period around the topics of case studies, existing implementation and standards related to Privacy by Design principles.
RESULT	The result of this study is a framework that identified the legal requirements of the General Data Protection Regulation GDPR and translated these legal requirements into technical solutions. This framework answered the research questions. Moreover, when evaluating

	the framework by applying it to a case study, the researchers could improve upon the results following the design science methodology.
DISCUSSION	Due to the limitation of the resources, this study only evaluates the artefact in one case study which could cause a restriction on the generalizability when using the framework. This framework has only positive implication on the society, whereby preserving human rights by applying it. Finally, the framework is original and valuable for all data controllers and processors who process EU citizen's personal data and who are subjected to the General Data Protection Regulations (GDPR) by allowing them to comply with article 25 of the regulations.

Table of Contents

Tables and Figures

1. Introduction

In today's world, digitisation is transforming everything around us. This evolvement is contributing to a better quality of life. However, at the same time, it created other problems such as the privacy issue which resulted in threats to the fundamental rights of individuals. One of the initiatives to tackle the privacy issues is the introduction of Privacy by Design concept by Ann Covoukian in the mid 90's to address the growing issues of the information and communication technologies (Cavoukian and Stoianov, 2007). The same author also introduced seven Privacy by Design principles at an abstract level to be used when developing new IT Systems.

On the legal side, there have been many attempts to provide citizens with privacy rights when processing their data both in the automated or manual way. In 1981, the Council of Europe introduced a convention for the protection of personal data, where many rights were granted for a citizen of member states who participated in this agreement. Afterwards, the Directive 95/46/EC of the European Parliament and the Council granted additional rights for all EU citizens. Subsequently, in April 2016, the new general data protection regulations (GDPR) was approved by the European parliament to come into effect in two years' time, starting from May 2018. GDPR is the EU-wide law applicable to all EU States. It provided more rights to EU Citizens and introduced more constraints on controllers and processors who process EU Citizens data for any purposes.

All public, and private sectors who process citizen's data will be subjected to this new regulation, where the bar has been raised in both offering the citizens more privacy rights and introducing constraints on controllers and processors. One of these constraints is the implementation of Data Protection by Design and by Default in IT System per Article 25 of GDPR. Affected parties cannot take privacy lightly as they will be penalised by 4% of the annual turnover or 20 Million Euros, whichever is higher in the case of violating the regulations.

Since then, there have been several attempts to operationalize Privacy by Design principles in IT systems according to the old principles that have been introduced by (Cavoukian and Stoianov, 2007) or by (Schaar, 2010). However, there are no works conducted to help entities to comply with the new data protection by design and by default GDPR regulations requirements. This research tries to help entities implement data protection principles into IT systems by answering the following questions:

- How should the European General Data Protection Regulation (GDPR) Privacy by Design by Default principles concretely be implemented in IT systems?
    - What are the legal requirements relating to the Privacy by Design by Default principles?
    - How can technical measures be mapped to these legal requirements?

This design science aims at translating GDPR legal requirements into technical solutions by developing a framework to concretely apply Privacy by Design and by Default principles into IT systems and its supporting processes to comply with the data protection regulations.

2. Methodology

In this section, the research method including data collection methods and research ethics are presented. Discussion about alternative research methods and data collection methods is available in the Appendix section.

2.1. Research Method (Design Science)

The research method that was selected for this research work is design science with the case study that will be conducted after building the artefact to evaluate it. The researchers argue that the GDPR compliance is a strategic goal and business needs for public and private sectors who process EU Citizens personal data to avoid penalties and to protect the reputation of the respective organisations as well as the privacy of the individuals. Arnould noted that, to achieve efficient translation of the strategy into the robust information system infrastructure, a full design activity is required (Arnould et al., 2004). The same author also articulated that design science addresses the business needs by building and evaluating artefacts. Therefore, design science research was selected to be the method for this research work to meet these requirements.

The protection of personal data is a challenging topic. The research aim is to solve this problem by building a framework that will help the both entities to comply with the new EU regulations. According to (Johannesson and Perjons, 2014), design research is not about describing, explaining or predicting; it aims at improving the world by developing an artefact that helps in overcoming problems. Morandi and Camargo also added that design science research is used when the desired goal is an artefact or recommendation (Dresch et al., 2015). Moreover, design research has an origin in the areas of IT systems, thus aiming to create an artefact will help in developing, using and maintaining IT Solutions (Johannesson and Perjons, 2014).

Privacy and Data Protection by Design and by Default is a practical problem for those who are processing EU Citizens personal data at this period since the regulations require demonstration of the compliance by May 2018. According to Johannesson, the definition of a practical problem is "*an undesirable state of affairs, or more precisely, a gap between the current state and a desirable state, as perceived by the participants in practice*" (Johannesson and Perjons, 2012). The gap between the current state and the desired state is apparent in the problem of this research. Both sectors must change their current practices and systems to comply with the new regulations. However, they do not know how to be compliant. During the initiation of this research, a CEO of an IT security company the researchers interviewed said, "*Companies are struggling to achieve GDPR requirements, they don't know how to implement it in their current processes and systems*". According to (Arnould et al., 2004), (Dresch et al., 2015) and (Johannesson and Perjons, 2012), artefacts can be used to solve practical problems experienced by people and, at the same time, strengthen the existing knowledge about these challenges.

Morandi and Camargo compared 13 authors' main elements of design science methods and reported them as follows: problem definition, literature review, a suggestion for a possible

solution, development, evaluation, the decision about the best solution, reflection and learning, and communication of results (Dresch et al., 2015). Based on the comparison, most authors included the following elements in their works: problem definition, suggestions for a possible solution, development, and evaluation of the artefact. Also, (Johannesson and Perjons, 2014) noted that the framework of design science consists of five steps, explicate problem, define requirements, design and develop artefact, demonstrate, and evaluate artefact. However, not all design science research undertakes these five activities.

Although there are several attempts to translate Privacy by Design principles into IT system requirements, there is no comprehensive framework for translating GDPR into system requirements. Therefore, the focus was on the development and the evaluation of the artefact. Johannesson noted that it is common for design science research to focus on development and evaluation, by using both the research and creative methods, and assess the artefact using different research strategies such as case studies or experiments (Johannesson and Perjons, 2014). The following figure is representing the main activities of developing the artefact along with the methods that were used for each stage.

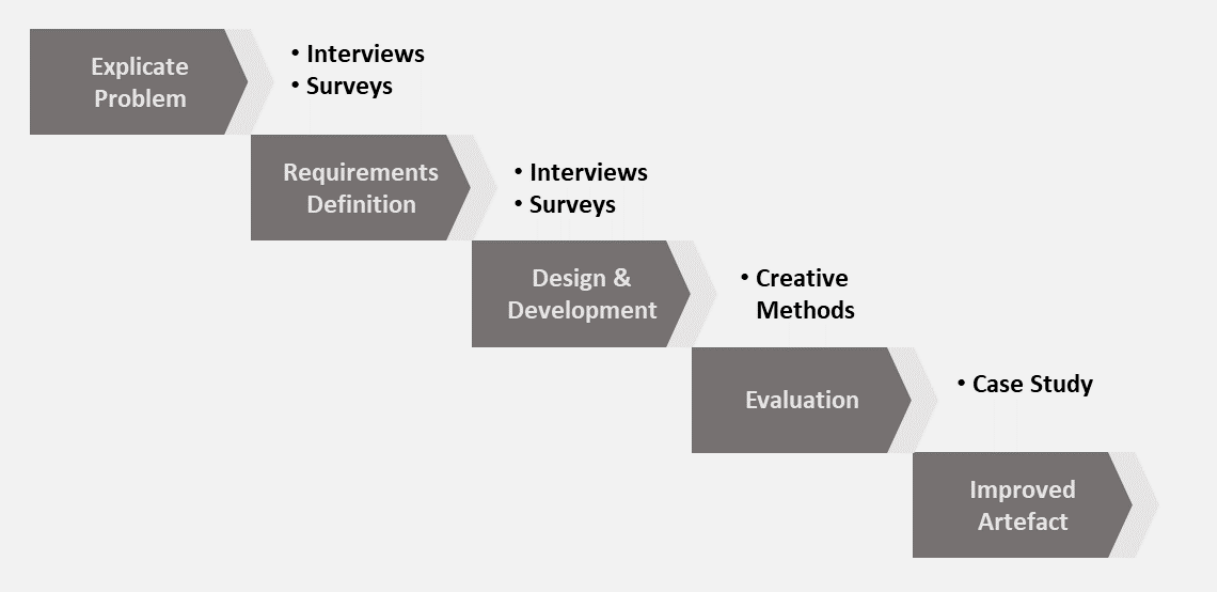


Figure 1 Design Science Method

2.2. Application of Design Science

This section discusses the activities that were conducted in this research to create the artefact. For the data collection methods, interviews have been carried out during the Explicate Problem, Requirements Definition activity and as data collection method of the Case Study during the Evaluation activity. There was a total of 5 interviews with both management team from IT security company and members from start-up company who is developing a system that processes the sensitive data. Furthermore, literature reviews were performed in parallel throughout, from the Explicate Problem until the artefact was developed in the Design and Development activity. Further discussion about the literature review can be found in section 3 and discussion around Design Science Method can be found in section 7.1 of the Appendix.

2.2.1. Explicate Problem

The first interview was conducted with the management team to exchange the knowledge about the Privacy by Design in GDPR, and to discuss a potential system that faces the problem. The management team revealed that the main issue is that despite the official publication of the GDPR, many companies are struggling to achieve the GDPR requirements because they do not know how to implement it concretely. The researchers were also told that many companies realise the challenges that there is no practical solution and pressure as the regulation will take effect in May 2018. The result from the interview concluded that the lack of knowledge to interpret the GDPR in a technical approach is the primary issue which prevents companies from understanding what specific measures needs to be implemented.

2.2.2. Requirements Definition

To outline and elicit the requirements of the artefact, the second interview where both management team and the CEO of the start-up company attended was conducted. The researchers began by reviewing the mutual problem *"how to concretely implement the GDPR data protection by design by default in IT Systems?",* and the possible approaches to solve the problem. The participants reached to a consensus that a framework should be developed to help the practitioners to translate the GDPR requirements into concrete IT systems requirements. Since overcoming the insufficient knowledge is considered to be cornerstone to build the framework, the participants and the researchers agreed that the framework is required to be usable, versatile and flexible. Therefore, the ultimate goal is to have a framework that helps practitioners with GDPR requirements in article 25 which is data protection by design and by default.

2.2.3. Design and Development

When started to develop the artefact, the researchers used creative methods, a brainstorming technique to collect ideas to develop the framework. This approach created different components necessary to build the framework and ensured the compatibility to define the requirements. All collected ideas were then grouped and selected according to their usages: *'how the framework should be built'* and *'how the developed framework will be used in the research work'*. The diagram showing the building and the usage of the framework can be seen in section 4.1. Both researchers followed this diagram to develop the framework as shown in Figure 5.

2.2.4. Evaluation

The researchers applied the developed framework into the case study for evaluation. Another three interviews were conducted to collect information needed to implement the framework. Following each phase of the framework, the first interview was carried out to understand the system and business context. The results were used to select appropriate measures for the system. After that, the researchers invited developers to participate in two more interviews to assess the system and to define data protection requirements. Finally, the researchers recorded the findings from the evaluation and provided recommended requirements for the system owner.

2.2.5. Improved Artefact

Upon applying the framework to the case study, the researchers learned about the necessity to improve the framework. The improvements were recorded as part of the result from the evaluation and were extensively studied. The researchers synthesised these new findings and created the improved framework as shown in section 4.4.

2.3. Research Ethics

The researchers were aware of the risks associated with this research. Since the research is about privacy and data protection principles, the researchers applied the principles itself on the process when conducting the research. Collected data from the interviews were anonymized, and stored in local drives of the researcher's computers and measures such as encryption and access control measures were also implemented.

The researchers also understood the sensitivity of the topic during the case study. Therefore, no sensitive data were reported per the signed non-disclosure agreements. The participations for both case study and interviews were totally voluntarily with the possibility to withdraw at any point in time during the research. Finally, the researchers have not dealt with any underage participants at any stage of this research.

3. Literature Review

In this section, the researchers conducted the literature reviews to answer questions around the "Privacy & Data Protection by Design by Default" topic. These questions are:

- What is the existing Privacy by Design principles?
- How can Privacy by Design principles be implemented in IT Systems?
- Are there any case studies for implementing Privacy by Design?
- Are there any established standards for implementing Privacy by Design?

Since the GDPR is a new regulation which will be effective starting from May 2018, there was not enough resources and literature that tackles the Privacy by Design and data protection principles provisioned in the regulation. Therefore, the data was collected from all possible resources including online journals, conference papers, books, white papers, industry reports, magazines, and any other type of documents found on the Internet.

To conduct the literature review, the researchers followed (Randolph, 2009) to construct the literature review section of this thesis. The following figure summarises the steps that were followed by the researchers of this thesis to conduct the literature review.

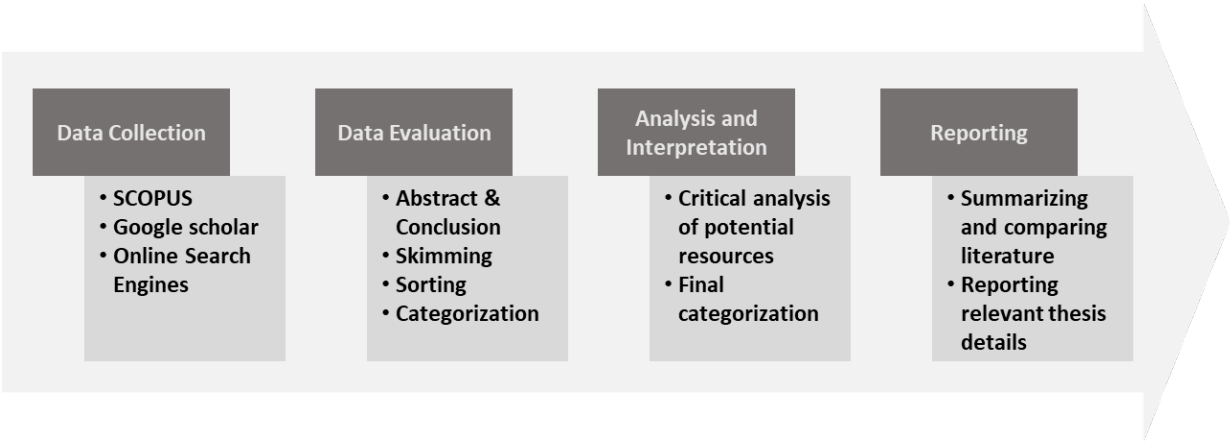


Figure 2 Literature review process based on (Randolph, 2009)

The process started by searching for the topic title "Privacy by design" and then continue looking for more specific keywords related to the topic such as "Implementing Privacy by Design" and "Operationalizing Privacy by design". All relevant literature was imported into a reference manager software "Mendeley" for post analysis. Then, all resources captured in the first iteration were carefully reviewed and were categorised per their contents and usage. After that, the categorised literature was critically analysed and further categorised. Finally, the results of the analysed papers were summarised, compared, and reported.

3.1. Definition of Privacy by Design

Privacy has been a controversial topic for a long time; there were many attempts to define what privacy is according to the context and the environment (Iorio and Carinci, 2013). Warren and Brandeis defined privacy term as the "*right to be left alone*" back in 1890 (Warren and Brandeis, 1890). After that, there were numerous of different definitions of privacy. On the legal side, privacy was perceived and defined differently from country to country, especially in the EU where privacy was lawfully granted for EU citizens in the European Commission of Human rights. Also, privacy was included on different upcoming conventions and directives such as the Convention for the Protection of Individuals 1980, the directive of data protection, and the recent general data protection regulations which repealed the EU/95 directive. During these evolvements, many rights were granted to the EU citizens, and, at the same time, introduce constraints to the entities who process EU Citizens data.

Privacy by design was introduced in the 90's by Ann Cavoukian who defined the concept as "Refers to the philosophy and approach of embedding privacy into design specification of various technologies" (Cavoukian and Stoianov, 2007). The seven principles were introduced and became the foundation of Privacy by Design concept for many research. Furthermore, (Schaar, 2010) introduced six different objectives to consider when designing processing systems, and argued that these principles should be bonded for technology designers, and producers. Also, the 11 privacy principles were in introduced according to the standard ISO/IEC 29100 (ISO/EIC, 2011).

Table 1 summarises the privacy principles in general and the Privacy by Design principles that were presented in the literature.

Table 1 Comparison of Privacy Principles

#	**Ann Cayoukian (Privacy by design)**	**Schaar (Privacy by design)**	**PriS Method goals**	**ISO/IEC 29100 (Privacy)**	**LINDDUN "Threats/ goals"**	**M. Rost (Goals)**	**Privacy targets "M. Caroline 2014. D. Science**	**OECD Guidelines**	**EU/95**	**GDPR**
1	Proactive not Reactive; Preventative not Remedial	Data minimization	Authentication	Consent and Choice	Likability	Integrity "Accountability."	Data Quality	Collection Limitations	Processed fairly and lawfully	Lawfulness consent, fairness, and Transparency
2	Privacy as default setting	Measurability	Authorization	Purpose legitimacy and specification	Identifiability	Availability "Bindingness."	Processing Legitimacy	Data Quality	Collected for specified, explicit and legitimate purposes	Purpose limitation
3	Privacy embedded into design	Transparency	Identification	Collection limitation	Non-repudiation	Transparency	Information right of data subject	Purpose Specification	Adequate, relevant, and not excessive about purposes	Data Minimization (Limited to the purpose)
4	Full functionality – Positive-sum, not Zero-sum	Data Confidentiality	Data Protection	Data minimization	Detectability	Interpretability	Access right of data subject	Use limitation	Kept in a form which permits identification of data subjects for not longer than necessary	Storage limitation (No longer than necessary)
5	End-to-End Security – Full Lifecycle Protection	Data Quality	Anonymity	User, retention and disclosure limitation	Information Disclosure	Confidentiality "Anonymity""	Data subject's right to object	Security safeguards		Integrity and confidentiality (unauthorised, unlawful, accidental loss, destruction, damage, technical or org measures)
6	Visibility and Transparency – Keep it Open	Possibility of segregation	Pseudonymity	Accuracy and quality	Content Unawareness	Concealment "Unobservability""	Security of Data	Openness		Accuracy (Accurate, up to date, erased or rectified)
7	Respect for User Privacy – Keep it User-Centric		Unlinkability	Openness, transparency and notice	Non-compliance	Unlinkability	Accountability	Individual participation		Accountability (Responsibility + Findability Demonstrate compliance)
8			Un-observability	Individual participation and access		Findability "Ascertainability."		Accountability		Data Protection by Design and by Default
9				Accountability						
10				Information security						
11				Privacy compliance						

According toI

Table 1According to Table 1, there is no consensus on the privacy principles or goals. Since the focus of this thesis is GDPR compliance, only the seven principles under the GDPR will be used. Moreover, GDPR article 25 noted that data controllers and processors have to infuse data protection principles in their products. However, these data protection principles were not mentioned in the same article, instead they are listed and defined in article 5 of the GDPR. Therefore, the researchers concluded the GDPR data protection principles as seven data protection principles and presented them with notations for each one in Table 2. These principles will be used as a foundation and the main requirements to achieve the data protection by design and by default in IT system.

#	Notation	Data Protection Principle
1.	DPP1	Lawfulness, Fairness, and Transparency
2.	DPP2	Purpose Limitation
3.	DPP3	Data Minimization
4.	DPP4	Storage Limitation
5.	DPP5	Integrity and Confidentiality
6.	DPP6	Accuracy
7.	DPP7	Accountability

Table 2 GDPR Data Protection Principles

More discussion and comparison about privacy by design approaches and data protection impact assessment in appendix 7.6 and 7.7.

3.2. Approaches to Privacy Implementation

There is a gap between the principles and the adoption of the implementation of IT Systems. The practitioners and engineers find it difficult to implement privacy principles into their system (Notario et al., 2015). Different approaches and some case studies were introduced to implement privacy on management, processes and technology level. Figure 3 articulates and summarises the previous approaches to implementing privacy at these levels.

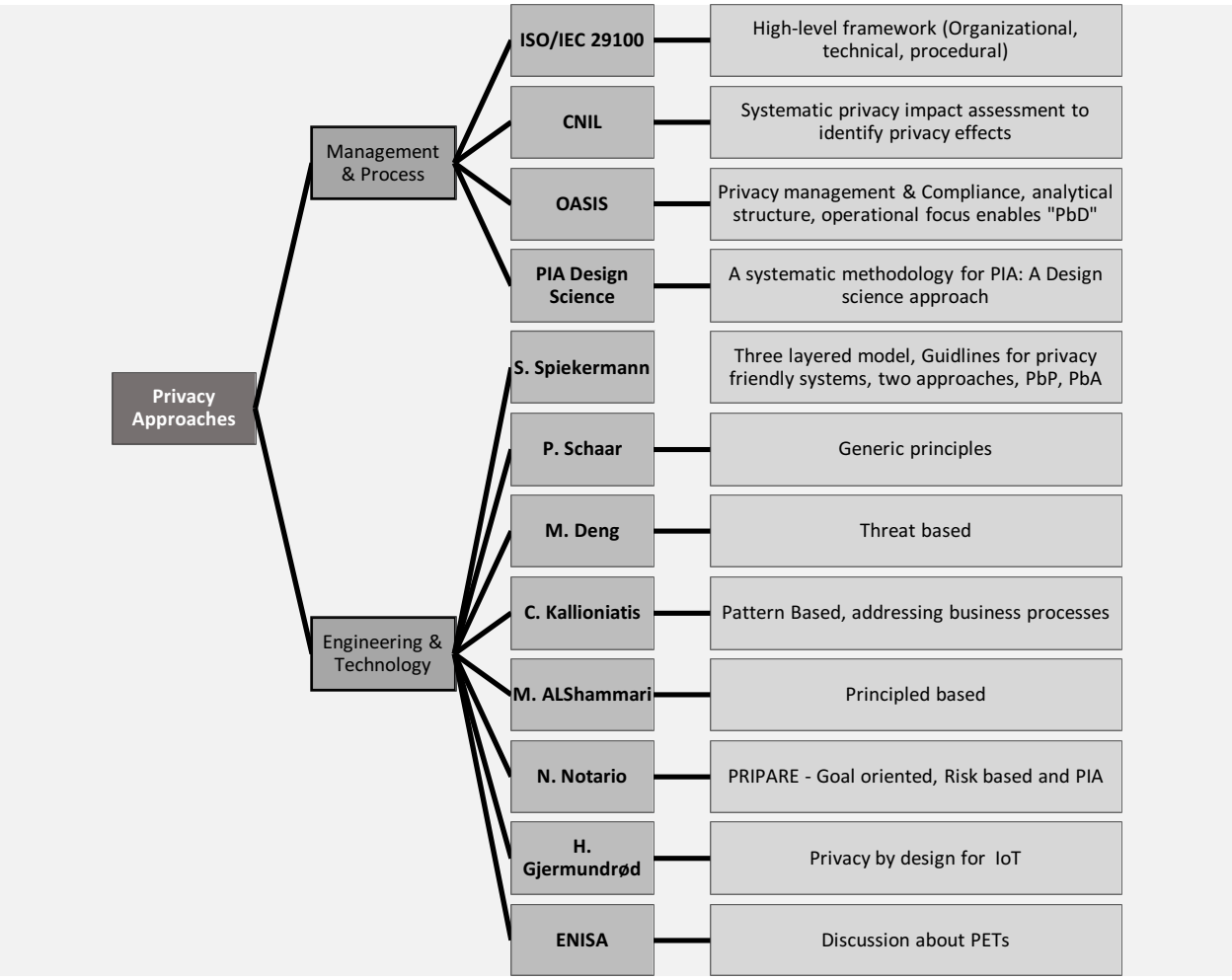


Figure 3 Privacy Implementations Approaches

3.3. Approaches to Privacy Impact Assessment (PIA) and Other Approaches

Wright defined that Privacy Impact Assessment (PIA) is *"a form of risk assessment, an integral part of risk management"* (Wright et al., 2011), while (Wadhwa and Rodrigues, 2013) viewed the PIA as a privacy management tool. These definitions can be implied that the PIA can manage the risk concerning the privacy and include a similar process to the risk management. Moreover, CNIL defined a privacy risk given the processing of personal data that has impacts on the privacy of data subjects and also stated that another aspect of the PIA is to determine suitable measures that complied with the legal requirements (CNIL, 2015).

The development of PIA went back since 1996 when the US Internal Revenue Service issued its IRS PIA, endorsed by the Federal Chief Information Officers Council (CNIL, 2015). Throughout the years, public sectors concerning information and privacy commissioner from various countries like Hong Kong, Canada, New Zealand and Austria has released their version of PIA to suit each need. Then, in December 2007, the Information Commissioner's Office (ICO)'s PIA Handbook was released and is considered the first PIA Handbook in Europe. After

that, in 2008, the International Organization for Standardization (ISO) published the PIA 22307:2008 standard. In the following year, the European Commission (EC) issued a recommendation on the implementation of privacy and data protection principles for RFID application which was endorsed by Article 29 (European Commission, 2011) and later became EC's PIA framework for RFID in February 2011. Also, there are similar tools such as Data Protection Impact Assessment (DPIA) which was introduced by the European Commission as an evaluation and decision-making tool. A further discussion comparing differing approaches can be found in section 7.6 of the appendix.

3.4. Challenges of Privacy by Design

Privacy by design is a new topic, with divergent views on both the principles and the approaches. There is no comprehensive framework that covers all aspects of privacy and risk-based approach. The lack of framework might be attributed to the controversial topic of privacy and no consensus on the privacy principles. As a result, there are different views on how privacy can be realised in organisations in general and in IT systems. At the same time, few case studies represent privacy implementation in IT systems. In conclusion, no established comprehensive risk-based standards or framework covers legal, technical and organisational requirements.

4. Results

4.1. Building the Framework

Based on the results from 3.1 in the literature review, the researchers started by identifying the requirements of GDPR Data Protection by design by Default which resulted in the 7 Data Protection Principles as presented in Table 2. Then, explored other approaches to implement privacy and to assess the impact on individuals in 3.2 and 3.3 of the literature reviews, the possible risks related to these requirements were identified. After that, the researcher identified the list of all feasible measures which fulfil the data protection principles and mapped them to each principle. Finally, the researchers identified possible approaches to realise the security requirements of the IT system.

Figure 4 Building of framework

At this stage, the researchers realised that while the framework can guide users to have their system complied with GDPR Principles, the framework was not versatile nor flexible enough to use in different situations. Also, in the real practice, other internal and external factors should be considered when selecting the measures. Therefore, the researchers referred to the seven factors mentioned in GDPR and added four other factors to compliment them. These additional refinements found their place in the framework as shown in Figure 5. Furthermore, the researchers analysed the literature reviews result, the legal requirements, and brainstormed to build the framework.

4.2. APSIDAL Framework

One of the core element of the framework is the Seven Data Protection Principles (7DPP) consists of **A**ccountability, **P**urpose Limitation, **S**torage Limitation, **I**ntegrity and Confidentiality, **D**ata Minimization, **A**ccuracy, **L**awfulness, Fairness and Transparent, hence **APSIDAL** Framework. This section discusses the components of the framework by dividing into three phases: Preparation Phase, Assessment Phase and Implementation Phase as shown in Figure 5.

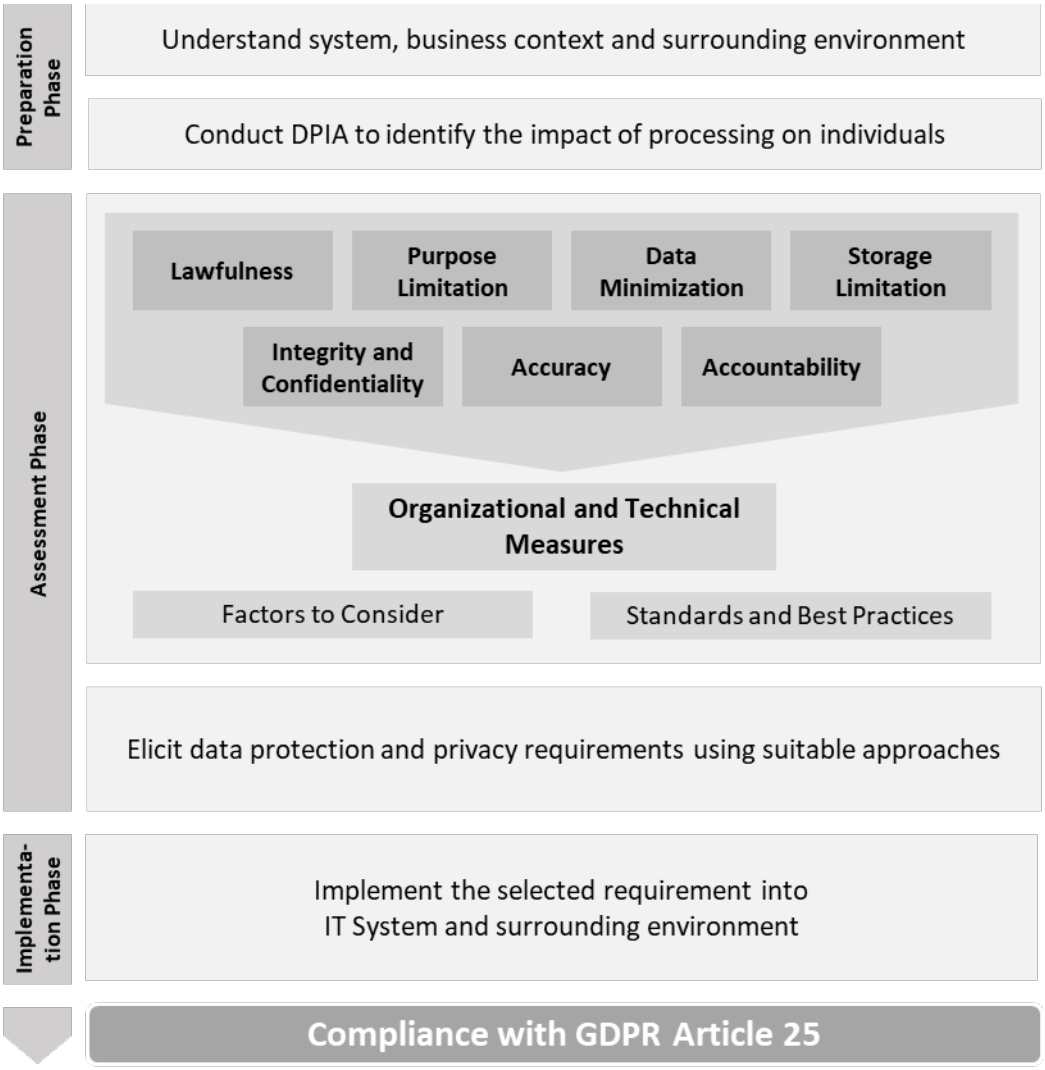


Figure 5 APSIDAL Framework

4.2.1. Preparation Phase

The preparation phase consists of two processes. First, the system context, business context and the surrounding environment are studied to have a sound understanding of the scope. The outcome will be the information that helps with the understanding of the system including, but not limited to, system diagram, business model, data flow diagram and data inventory.

The second process is to utilise the information from the previous process to conduct the DPIA. As stated in 3.3, there is no approach to conducting the DPIA since each gives a different result based on the purpose it designed for. The expected result from this process is the impact on individuals towards the seven principles. The severity of this impact will influence the selection of the measures.

4.2.2. Assessment Phase

At this phase, before any processing to personal data, data controllers must consider the legal basis for the processing. The data controllers should have the legal resources necessary to make sure that the processing is on par with the legal requirements of GDPR, such as legal consent or another legal basis for the processing. Since the scope of this thesis is the IT systems, it is necessary to point out any legal requirements should be handled separately by the legal specialists.

The core element of this phase is the Seven Data Protection Principles shown from Table 3 to Table 9. Each principle includes the GDPR Provision, objective, and measures. The GDPR Provision is from Article 5 with the objectives derived from the same article. Each table is dedicated to each principle and presents both technical and organisational measures to achieve compliance with the GDPR Data Protection by Design and by Default principles. The outcomes of this process are the list of general measures to fulfil the principle objectives. Also, the term data controller(s) in this framework implies both the data controllers and data processors.

Table 3 Lawfulness, Fairness and Transparency (DPP01)

DPP1	Lawfulness, Fairness and Transparency
GDPR Provision	*"processed lawfully, fairly and in a transparent manner in relation to the data subject"* GDPR Art.5.1(a)
Objective	No matter how minor the processing is, processing of personal data must stand on a firm legal basis. The data controllers must provide clear views on how the processing works and the consequences on the data subject before collecting and processing the data.

Organizational Measures

Policies, Processes, and Procedures: Data controllers must have strong policies, processes, and procedures in place to assess, and evaluate whether the processing of personal data has strong legal basis, fair and transparent to the data subject. These measures should be available in a simple language that an average user can read and understand. No negative consequences should be the norm for opting out of any service or engagement. That is, all user rights must be maintained before, during and after the processing. The purpose of the processing of the data, including any planned sharing, selling, disclosing, or further processing of the data, should be available to all data subjects. Policies for sharing and distributing the information is essential and must be implemented.

Legal Measures: Legal measures must be maintained by the data controllers. GDPR has different provisions concerning the legal basis for processing personal data. Consents are not the only legal basis to process personal data. Therefore, it is advised to consult the legal specialists before starting any processing activities.

Technical Measures

Embedded Transparency Measures: Data controllers should consider embedding the necessary forms, dialogues, policies, procedures, data processing and user rights into the design of the information systems. For example, when an application wants to get access to the location coordinates for the data subject, it should prompt the user for approval before accessing this information. Moreover, the consequences of this approval should be clearly communicated to the user. At the same time, the user has the right to opt out of this service and request the information to be deleted. These measures must be considered in when designing the system, not afterwards. d

Embedded Legal Measures: The legal measures can be technically implemented in IT systems by embedding the legal requirements into the systems. For example, a system with traceability capability can capture the consent and trace it back to the date and time when the data subject agreed.

Non-Repudiation Services: Depending on the level of the risk and the impact on the data subject, data controllers should consider implementing the non-repudiation service from the data subject. For example, the digital signatures shall be implemented when the collected data is sensitive.

Table 4 Purpose Limitation (DPP2)

DPP2	Purpose Limitation
GDPR Provision	"*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*" GDPR Art. 5.1(b)
Objective	Purpose limitation requires the data controllers to define the reasons for data collection and processing. Data controllers need to ensure that the data is processed according to the original purpose. The data should not be processed when the purpose has changed without firm legal ground. Also, data traceability throughout the data lifecycle is needed to know when the processing is not according to the original purpose anymore.

Organizational Measures	Technical Measures
Policies, Processes, and Procedures: The policy shall require the data controller to collect data only for a specific and explicit purpose. For some case, however, the personal data can be collected to prevent life-threatening emergencies, when law requires it, or when it relates solely to non-profit organization members or individuals. Moreover, while data collection can have any purpose, it is necessary to check whether that the purpose is legitimate. That is, the purpose shall confine with the legal concerned, and further processing shall be suspended when it is no longer compatible. For example, the flight ticket website starts to sell the information about the number of checked luggage to luggage vendor without legal ground causing the data collection purpose to deviate.	**Data Inventory Measures:** when collecting, processing and storing personal data, the data controllers should be able to trace back the purpose of the data collection. This measure can be implemented in the IT system by tagging or adding metadata to describe why the personal data was collected, for what purpose and for how long it should be stored. This measure can also be linked to the data minimization and storage limitation principles where the data controllers can wipe any extra data that were collected or stored and did not serve the purpose of the processing.

Table 5 Data Minimization (DPP3)

DPP3	Data Minimization
GDPR Provision	"*adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed*" GDPR Art. 5.1(c)
Objective	Data minimization involves reducing the amount of the collected data to what is necessary that the interaction with the data subject is satisfied without.

Organizational Measures	Technical Measures
Policies, Processes, and Procedures: Data controllers should strive to adopt, design, implement, and continuously improve policies and procedures to protect data subjects by preventing the unnecessary collecting, processing, storing or sharing any personal data that is not necessary for any interaction with the data subjects. **Access Limitation:** Any data should be minimised and made available only to those who cannot perform their duties without accessing it. Therefore, the location where the data is physically stored is vital from a data minimization stand point. The data storage and backup shall be limited to the fewer location where the data availability is necessary.	**Centralise Storage**: Store the data in a central location and use some technical measures such as access controls lists which can contribute to the data minimization by reducing the risk of disclosure and disruption of the personal data. **Data Pseudonymisation:** This technique can be used to replace part of sensitive data to prevent the link to the data subject. For example, by encrypting name and birthday data, even if the system is breached, the stolen data is not linkable to the data subject. However, the encryption key must be kept in secure place to prevent decrypting data back to the original state. **Strip Unused Metadata**: The unused data can be a burden on the data controllers. Deleting these data from storage and backup systems help to comply with the data minimization principle. The action can involve manual intervention from data controllers by having rigid procedures in place to make sure no unused are left in the system. It is recommended to automate this process and to conduct data inspection on a regular basis. **Intermediary Proxies**: The data collection and data transmission can be done through intermediary proxies where data can be filtered out and sent as anonymized data. This measure is usually applied to the transmission containing data that can be linked to the data subject such as IP addresses and cookies. Onion routeing is one of the example for this measure.

Table 6 Accuracy (DPP4)

DPP4	Accuracy
GDPR Provision	"*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*" GDPR Art. 5.1(d)
Objective	Data controllers must take necessary steps to ensure the accuracy of data obtained and to verify the data source. Furthermore, any challenges to the accuracy of information shall be considered and keep up to date when necessary. In the IT system, the accuracy of a digital record can also be measured by the ability of anyone to understand what the record says.

Organizational Measures	Technical Measures
Data Completeness Awareness: An awareness when entering data must be raised for all data controllers. When the same data record can be accessed by different employees, the problem might arise when each employee doesn't realise the importance of filling data correctly in all required fields. The incorrect or missing personal data could cause negative consequences not only to the data subject, but also the trustworthiness of the data controllers. **Data Normalization Policy:** Data collected from various sources might include different spelling variation. Therefore, there should be a guideline for all employees to followed. For example, the term 'artefact' in British English and 'artifact' in US English may seem to be understandable for human eyes. However, the system cannot distinguish between the two which might cause the inconsistency or segmentation problem. **Data Management**: There should be policies, procedures, and processes to define, and to assign roles and responsibilities for the data controllers to ensure the data quality. The data accuracy should also be checked when entering the new account to make sure the data is recorded correctly and completely. Furthermore, data controllers shall keep the personal data up-to-date and valid.	**Data Dispute Handling**: Data subjects shall be notified of their right to object or change personal data for legitimate reasons and the system should allow a communication channel for the user to inform about data disputation. After received a request from a user, the request handler should notify data controllers to investigate and solve the issue. **Data Cleansing:** Improving data quality overall using this technique can also increase the accuracy of the data. The data shall be analysed for its correctness, completeness and consistency as well as removing the 'dirty data' (inaccurate or incomplete data) in the system. This improvement can be made by using the tools or instruct the employees to perform the process manually.

Table 7 Storage Limitation (DPP5)

DPP5	Storage Limitation
GDPR Provision	*"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" GDPR Art.* 5.1(e)
Objective	Storage limitation principle focuses on keeping the identifiable data for only the period that the data serve its purpose. The data controllers have full responsibility to maintain data tracking and remove the data when it is no longer being processed for its original purpose.

Organizational Measures	Technical Measures
Training and Awareness: While putting a limitation on storage mainly rely on the software that manages the data and the hardware that it is stored on, the knowledge of the staff is also necessary. Having the data that no longer serve its original purpose is no different than processing unnecessary data. The training needs to guide data controllers to realise the risk of storing data longer than the intended processing time. **Data Lifespan:** When personal data is created, its lifespan shall be defined so that it can be destroyed automatically. However, this limitation must be in accordance with the period that data is still in use. This measure is done to prevent human error losing track of the data over time.	**Traceability**: An information technology system with good storage management would be able to show the relation and tracking route from the actual data to the backup data or other distributed copies. These locations are where the data are often overlooked. Enable the traceability provides a manageable way to remove the personal data along with its backup or copies in distributed storage after the processing finished.

Table 8 Integrity & Confidentiality (DPP6)

DPP6	Integrity & Confidentiality
GDPR Provision	"*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*" *GDPR Art.* 5.1(f)
Objective	Integrity and confidentiality are part of the foundation of information security. Protecting the privacy of the data subject by maintaining its integrity is to maintain the accuracy and consistency of stored data. Also, the confidentiality of the data is maintained by protecting the information from disclosure to unauthorised access. The measures for this principle shall be implemented and operated throughout the data lifecycle.

Organizational Measures	Technical Measures
Identity and Access Management: Identity and access management is one of the main components to uphold the confidentiality of personal data. Without proper identification, authentication and authorization processes, it will be impossible to make sure the right data subject has the adequate access level to the correspondence data. **Encryption and Key Management Policies, Process, and Procedures:** The use of cryptography is one of the strategic goals of any organisation dealing with personal data. This strategy must be translated into concrete initiatives across the organisation where personal data is collected, processed and stored. The encryption methods, libraries, specifications, and attributes of the encryption key is the critical aspect of the strong encryption and must be specified at the organisational level policies and procedures, both for data in transit and at rest. **Restriction on Data Processing:** The personal data shall be processed only on the approved devices or locations. If the organization has the Bring-Your-Own-Device policy which allows employees to personal device for work, proper controls shall be in place to allow only non-sensitive data to be accessed. **Physical Security:** The physical access to the facility, devices and other location where the personal data is stored shall be managed by data controllers.	**End to End Encryption:** Personal data should be encrypted end to end when it is transmitted on the network. An adequate level of encryption should be implemented on the systems depending on the level of risks associated with processing the personal data. Encryption can also be achieved by, but not limited to, different approaches including anonymization and Pseudonymisation. **Data Validation:** A technique to validate the data shall be used to preserve the integrity of the data when transmitting, or transfer in the network. By validating the data, the inconsistency and inaccurate will be reduced. This measure also contributes to the accuracy principle in overall. For example, input validation and hashes can be used to preserve and maintain the integrity of the data. **Authentication:** An adequate level of authentication is required when processing personal data. Based on the known categories of authentication (something you know, something you have, something you are) and the sensitivity of personal data, the data controllers must select the suitable type of authentication. In certain situations where the data sensitivity is high, the two-factor authentications with distinct categories of authentication shall be implemented. **Authorization:** An access rights management must be implemented into the systems that handle the processing of personal data. Role-based access control, or any access that can mitigate the risk of data disclosure to unauthorised persons, should be adequately implemented into the systems and the supporting environment.

Table 9 Accountability (DPP7)

DPP7	Accountability
GDPR Provision	*"The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1"* GDPR Art. 5.2
Objective	Accountability is a new concept introduced in GDPR. The data controllers must be accountable and be able to demonstrate compliance with the provisions of the regulations. The demonstration can be achieved in several ways, from not processing un-legal personal data, to implement the privacy principles into IT systems.

Organizational Measures

Strategy: Data controllers must consider privacy aspects in their business and operational strategies by demonstrating strong commitment to GDPR.

Standards and Best Practices: Data controllers and processors should strive to adopt and implement state of the art standards and best practices.

Policies, Processes and Procedures: The policies and procedures must be in place to govern and manage the organisation. The data controllers should consider aspects such as people, process and technologies when designing any measures. It is vital for compliance demonstration to frequently revisiting these measures to make sure they are harmonised with the controllers and processors environments.

Awareness and Education: Since the environment of controllers and processors are always changing, awareness and education for both data subjects and employees must be embedded as part of the processes performed by the data controllers. Constant improvement to the content is highly required to be on par with the new threats and changes in the environment.

Certification on IT Products and Services: One way to demonstrate the compliance is to have certification on hardware, software as well as on the service itself. Many organization or authorities offer privacy-specific certification including EuroPriSe (Jarno J. Vanto, 2009), which claims to recognise both IT products and services compliance with the new GDPR.

Technical Measures

Authentication and Authorization: An adequate authentication and authorization are considered a crucial aspect of the compliance of GDPR. Data controllers must consider the strength of the authentication, and the level of granularity of authorization based on the sensitivity of the personal data. Two or more factors authentication is recommended for very sensitive data which could cause danger to data subject's life if leaked.

Tampere Proof Audit Trails: When designing, building and maintaining the IT system, it is necessary to have audit trails which have sufficient details that can answer when, where, why, whom, and how questions of the user and system actions. The reliable audits trails must be protected from tampering. Moreover, measures such as encryption shall be applied in the audit trails where the sensitive personal data might be stored.

Monitoring: Since the outcome of this framework is compliance with GDPR Data Protection by Design by Default. Any identified threats on data subject must be prevented. Monitoring could be viewed as reactive activity; however, it can be useful if it is adequately implemented to prevent more risks.

Data Loss Prevention: The data loss prevention is yet another measure that can be implemented to detect and prevent data leakage at the last stage in the way out. It is a reactive rather than proactive approach to protecting data. This measure should be implemented as close as possible to the data for better protection.

Factors

The severity of the impact on the data subject and the understanding of the system from the previous phase determine how comprehensive the measures shall be implemented. Also, there are aspects to consider when choosing the measures. Provision 25 of the GDPR stated that the controller/processor should consider several factors when designing data protection in the systems or products. Those seven factors are state of the art technology, cost, nature, scope, context, purposes of processing, and the risks associated with the processing. In this research, both researchers synthesised these seven factors with the following additional factors to create a basis to find common ground among business, technical and data subjects' requirements.

1) Complexity – The implementation and operation of these measures should not be too complex that the data controllers and data subjects have a difficulty using it.
2) Usability – The measures with higher usability are more likely to have data controllers efficiently operate it.
3) Efficiency – Even with the same objective, different measures might generate different results. The efficiency determines how smooth the measures are operating.
4) Effectiveness – The effectiveness of the control determines how through the control achieve its purpose. The effective control shall fulfil its objective extensively compare to the weak one.

Altogether, these 11 factors will be considering in the framework to guide the selection of the measures.

Standards and best practices

The table below presents standards and best practices that the users of this framework can refer to find a complete list of controls tailored to their environment.

Table 10 Sources for Data protection, privacy and security measures

Domain	Standards and Best Practices	Description
Organizational	ISO 27000	Set of standards that help organisations to secure their information assets
	ISO 29100	Framework and set of controls for organisations who process personally identifiable information
	ISO 27018	Set of controls that meant to protect personally identifiable information in public cloud
	UCF	A unified compliance framework is a library that maintains a comprehensive set of controls applicable to both organisational and technical needs
	COBIT 5	A business framework produced by ISACA which includes various enablers for Securing Sensitive Personal Data or Information (SPDI) (Kadam and Vutha, 2012)

Technical	NIST 800-53	Technical Security controls for information systems
	ENISA	The Privacy enhancing technologies Guide, *"Privacy and Data Protection by Design from policy to engineering 2015",* contains a set of technical privacy measures to build privacy into systems and services

Data Protection and Privacy Requirements Elicitations for IT Systems

In this process of the assessment phase, the goal is to elicit and identify the data protection and privacy requirements of the IT systems. To make this framework applicable for any IT systems under developments and operating IT systems, the researchers propose two approaches. The first approach is to use tools such as threat modelling or LINDDUN (Deng et al., 2011) during the design phase of the IT system. In case the IT system is already in operation, the second approach is to use static code review, effective code review, or penetration testing to identify the vulnerabilities that can be fixed by applying add-on measures. The following table explains the different approaches for data protection and privacy requirements elicitations.

Table 11 Approaches for Data protection, privacy, and security requirements elicitation

Approach	Technique/Method	Description
System Under Design	LINDDUN	LINDDUN is a comprehensive framework for identifying privacy threats in software based systems. It used methods such as data flow diagrams and STRIDE to analyse the systems and identify privacy threats. Also, it contains a mapping for each privacy threat and the existing countermeasures to mitigate it (Deng et al., 2011).
	Threat Modelling	Threat Modelling is often used in the design phase of the systems where threats can be identified. Threat modelling can also use different techniques such as STRIDE, DREAD, misuse cases attack trees, and ASF to identify and analyse the potential threats (OWASP, 2015).
	Common Criteria (ISO/IEC 15408)	Common Criteria is an international standard for security requirements in information systems. It has re-used repository of security requirements. These requirements allow both the developers and the end-users to evaluate and gain trust on the products according to the implemented security functions (Mellado et al., 2007).
Operational System	Penetration and Vulnerability testing	When the system is in operations phase, penetration testing and vulnerability scanning can be used to identify system deficiencies from the attacker points of view. Therefore, system developers can fix the identified vulnerabilities which enhanced the security. Penetration testing is a too little too late

		method and should not be the only way to assess the security and identify security requirements (Arkin et al., 2005).
	Static and Dynamic Code Reviews	Static and dynamic code reviews is another way of reviewing the software. Static code reviews can be conducted by a specialised security expert or tools to identify vulnerabilities in the code. However, not all vulnerabilities can be identified using this method (OWASP, 2017).

4.2.3. Implementation Phase

Results from the previous phase are the concrete Data Protection, Privacy and Security Requirements that were 1) derived from the seven data protection principles, 2) based on the general measures after considering factors and impacts on the data subjects, and 3) attuned to the system using assessment tools. Therefore, by implementing these measures, the outcome is compliance with GDPR Data Protection by Design by Default.

4.3. Case Study for Artefact Evaluation

The purpose of conducting case study is to evaluate the artefact that was created to answer the research question. The researchers tried to find a suitable practical case for an IT system under development that aims to process personal data. In this section, the researchers present the case study and the application of the framework.

4.3.1. ChatBot

The researchers use the term “ChatBot” to refer to the IT system in this case study. A ChatBot is an IT system that is under development by a start-up company in Sweden. The ChatBot is a mobile application that acts as a virtual friend to the users which allow them to share memories, stories, and any other conversation in both texts and pictures format. These data are stored in the server to generate conversations about different topics to interact with the user. The collected data are sensitive according to the owners of the application. Since there are no restrictions on what type of the data can be entered, the users have the ultimate freedom to enter any text about any topic in the ChatBot. The owners of the application believe that privacy and data protection is not only the requirements to comply with, but also a business enabler that will allow them to meet the users’ expectations, as well as sustain and improve their business. Our engagement and scope are to help the company to comply with the GDPR requirements by implementing the data protection principles in their system and the surrounding processes. Our goal is to assess whether the framework can achieve its objectives and provide a useful guide for the business owner to comply with the data protection principles of the GDPR. In the following sections, the researchers explain the steps followed to apply the framework to the ChatBot system.

4.3.2. Applying the Framework

Preparation phase

To apply the framework, the researchers conducted several interviews both with the business owners and designers and developers of the system. The researcher started with the understanding of the business context of ChatBot and its business model. By knowing that ChatBot is meant to collect and process sensitive data of the data subjects, the impact of data disclosure is very high. From the interviews, the researchers found out that the owners already know the impact of the personal processing data. Furthermore, they said, *"if data got leaked, the impact could be severe, data subjects might lose their jobs or their beloved ones"*. The business owners provided the researcher with all necessary information to understand the business context and requirements. Then, the data protection impact assessment (DPIA) was considered together with the business owner of ChatBot.

Assessment phase

The understanding of the business context, the surrounding environment of ChatBot together with the impact on the data subjects assists the researcher to conduct the assessment phase. The researcher used this information as an input to the Seven Data Protection Principles (7DPP) to select the appropriate organisational and technical measures that meet the objectives of each principle. Moreover, during the selection process, other factors were also considered. Figure 6 shows how the inputs from 7DPP were combined with other aspects and factors to select the Data Protection by Design by Default measures relevant for ChatBot.

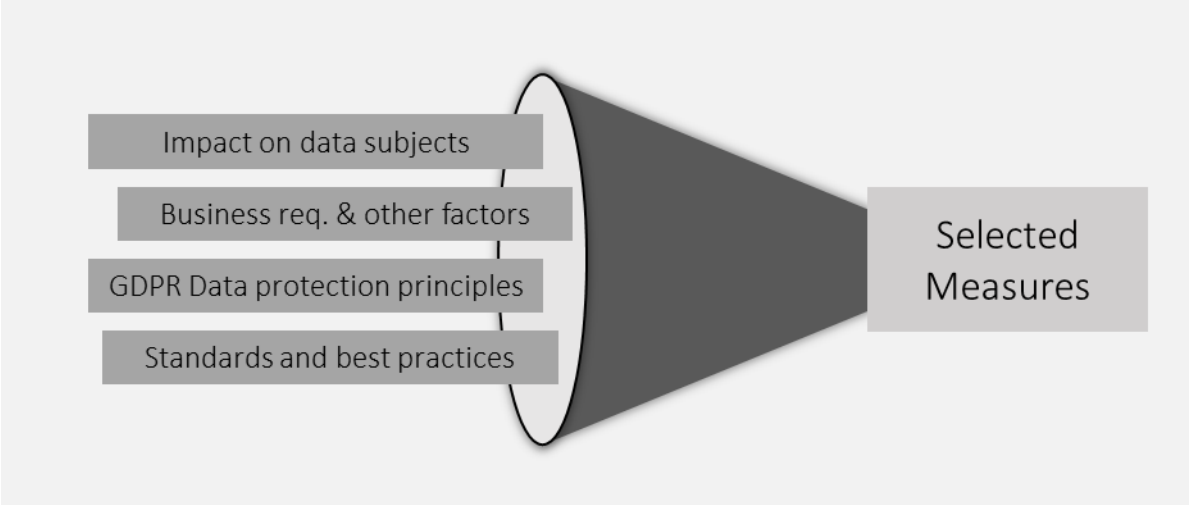


Figure 6 Applying the framework on ChatBot

These selected measures are generic measures which need to be translated into concrete system requirements. The translation can be achieved by several ways, for example, threat modelling or LINDDUN when the system is under design, or penetration testing when the system is in operation. In the case of ChatBot the system it is under design, so the researchers suggested to use tools that meant to elicit data protection, privacy and security requirements of the IT system.

Implementation Phase

The result from the assessment phase is the concrete list of data protection, privacy and security requirements. These requirements must be implemented and tested in the IT System. Due to the limited time, resources and the status of the ChatBot development, the outcome of the requirements implementation cannot be verified.

4.3.3. Case Study Results

The results of applying the framework on ChatBot provided the stakeholders with the knowledge on how to implement Data Protection by Design by Default on IT systems according to GDPR requirements. The knowledge includes the suitable measures for the environment that fulfil each principle as well as suggestions on how to implement these measures in a real environment.
However, during the process of applying the framework on ChatBot, the researchers found other aspects that need to be incorporated into the framework. These discovered aspects were used to improve the framework. Therefore, the case study fulfils its purpose of evaluating the artefact and improving it.

4.4. Framework Improvement

Based on the framework evaluation by using ChatBot case study, the researchers improved the framework by considering Software Development Life Cycle, Operating Environment and IT System Operations. The following figure shows the improved framework taking these aspects into consideration.

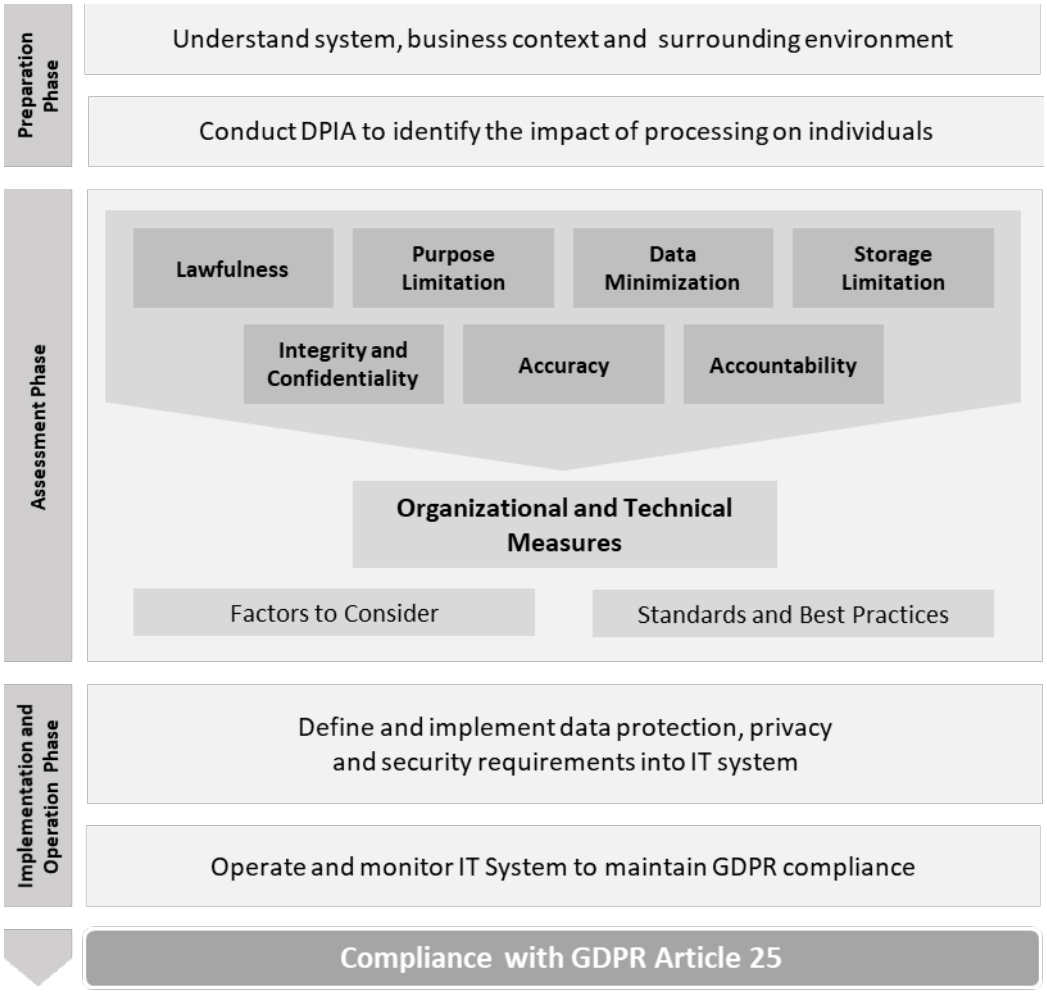


Figure 7 Improved version of APSIDAL Framework

Software Development Life Cycle (SDLC)

In this improved APSIDAL Framework, the outcome of the assessment phase can be used as an input to existing SDLC. Several organisations have established and matured SDLC where the security processes and tools are infused, whereby data protection, security, and privacy requirements can be realised. Furthermore, these requirements must be checked throughout the SDLC to assure the GDPR Compliance.

Operating Environment

When implementing Data Protection by Design, the operating environment and the supporting components of the system such as a database, operating system and hardware should be considered. Figure 8 represent the system components and the relying infrastructure.

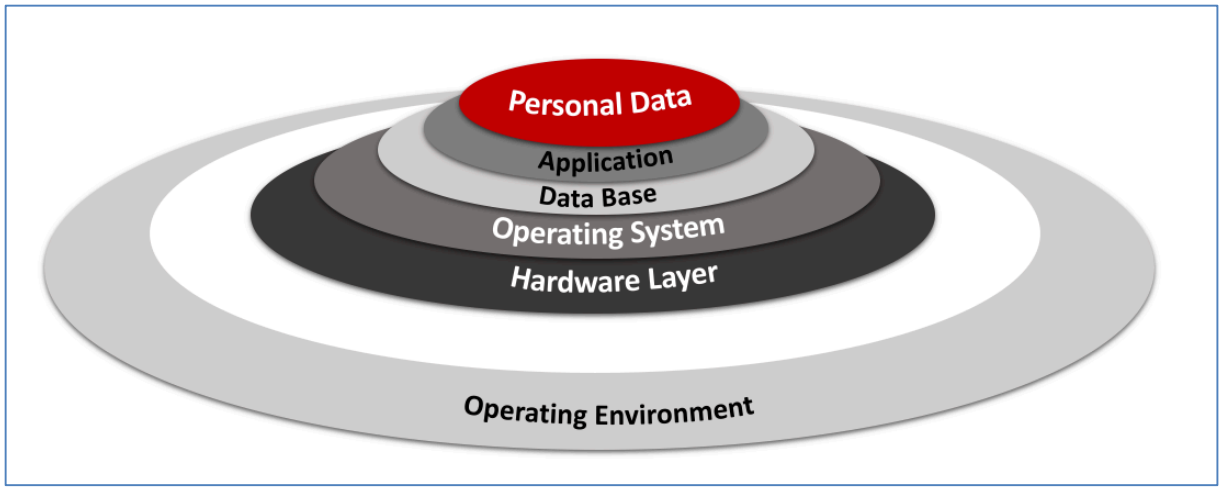


Figure 8 Operating Environment

IT System Operations

Monitoring the operation of the IT System is considered a vital aspect to maintain the data protection of data subjects. Since changes and threats in the environment may arise, timely intervention to prevent unwanted events and risks is required. Besides the IT system, the underlying infrastructure should be maintained and receive the proper level of attention as well. The measures such as access control, security baselines, applying latest patches and other best practices can assure the highest protection levels of personal data. Moreover, not only the technical measures but also organisational measures shall be implemented.

Besides these three distinct, the framework also covers other areas such as documentations and data protection mapping to data lifecycle to help the practitioners. The discussion of each area can be found in the following:

Documentation

Demonstration of the compliance is considered as one of the fundamental principles that were introduced into GDPR under DPP7 (Accountability) principle. Implementing Data Protection by Design and by Default into IT systems is contributing to the accountability principle. However, the data controllers should consider creating structured and comprehensive documentation for each step of this APSIDAL Framework to be able to demonstrate compliance with supervisory authorities. This documentation is also useful for the controllers to sustain and operate the IT system.

Data Protection Principles Mapped to Data Lifecycle

When applying he framework on ChatBot, the researchers realised the importance of mapping the data protection principles into the data life cycle phases of an IT system. The lifecycle of the data has been designed in various models. According to (Danezis et al., 2014), data protection by design shall cover the entire lifecycle management of personal data from collection to processing to deletion. Moreover, Hoel and Chen discussed the influence of GDPR requirement on the system development using the learning analytics processes (Hoel and Chen, 2016). This learning analysis process cycle was defined in ISO/IEC 20748-1 which includes

Learning & Teaching Activity, Data Collection, Data Processing & Storing, Analysing, Visualization, Feedback & Recommendation phase (ISO/IEC, 2016). Evidently, depending on the application, the lifecycle is differently divided into phases to accommodate the usage of data and to map those phases with the controls or action to be done on the data. In this research, three data lifecycle phases: data collection, data processing and data storage were used.

The mapping will help the practitioner of this framework to have a clear understanding of the data protection principles and will aid in the selection of the adequate technical and organisational measures for each phase. The data life cycle starts with data collection phase, where data minimization principle should be considered. The data controllers should collect only the necessary personal data that the processing cannot be performed without. In the data processing phase, purpose limitation is relevant, whereby the data controllers should process these collected data for the purposes they were collected for. Data storage is the third phase, where personal data should not be stored longer that it is required. Moreover, accuracy, integrity and confidentiality and accountability principles apply to all data lifecycle phases and the corresponding principles. Finally, it is essential for the data controllers to base all activities on the lawfulness, fairness and transparency principle. Figure 9 visualises the mapping between data life cycle in IT systems and seven data protection principles.

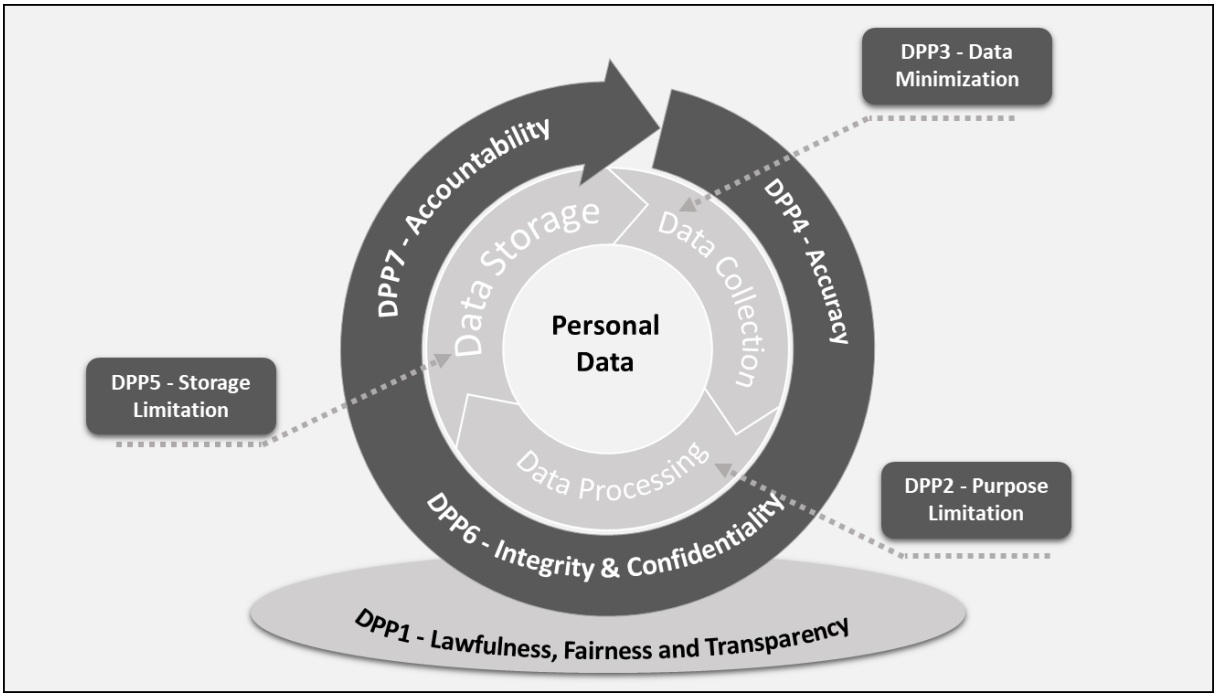


Figure 9 Date Protection Principles Mapped to Data Lifecycle

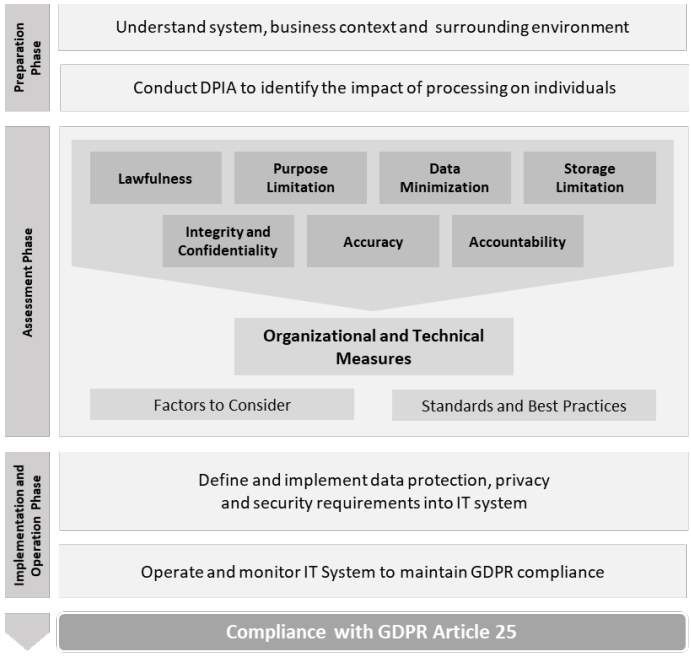
5. Conclusion

To answer the research question *"How should the European General Data Protection Regulation (GDPR) Privacy by Design by Default principles concretely be implemented in IT systems?",* researchers created APSIDAL Framework to implement GDPR Data Protection by design and by Default into IT systems. The researchers identified and translated the legal requirements to the into technical solutions and incorporating the necessary factors to achieve a comprehensive yet flexible framework. Seven data protection principles were elicited, defined and an objective for each one was set. To fulfil, these goals, both organisational and technical measures were identified and mapped for each principle along with examples.

Contribution

The framework consists of three phases: 1) preparation phase, 2) assessment phase and 3) implementation and operation phase. In preparation phase, the business, system context and the surrounding environment are studied to have a sound understanding of the scope. Then, DPIA is conducted to find the impact of the personal data processing on data subjects. The outcome of the first phase assists the practitioners in the measures selections. In the assessment phase, other factors, standards and best practices should also be considered to find the relevant and comprehensive measures to fulfil the legal requirements and to suit the environment. After that, in the implementation and operation phase, these selected measures are translated into the data protection, privacy and security requirements using existing SDLC or other assessment tools. Finally, the IT system must ensure these requirements during the operation to maintain GDPR compliance.

The researchers evaluated the framework by applying it to a virtual friend mobile application referred to as *'ChatBot'*, as a case study. Upon using the framework, the researchers realised the necessity to improve the framework. The following figure shows the framework with the improvement on generalisation, flexibility and usability area.



Also, Respect for human rights and freedom are fundamental elements of any democratic society. This research contributes to the society by building a framework to realise the privacy and protect data subject's data in IT systems. By applying this framework, the individual rights will be preserved, whereby the society can flourish and grow.

Limitation

The limitation of design science may arise in the evaluation stage when the artefact is evaluated in only one environment. In this research, due to the limitation in resources, one case study was conducted to assess the artefact. It can be argued that generalisation is not applicable for all. However, the framework can be assessed in future research in a different environment to test the applicability, efficiency and effectiveness in solving the research problem.

This research is limited to translate the legal requirements of GDPR to technical solutions. The researchers interpreted the provisions according to their understanding and based on the literature review. However, these requirements might change when the regulations enforced in 2018. Also, different methods and tools to achieve compliance with GDPR data protection principles were suggested, but there are no discussion and evaluation around the suggested measure. The framework leaves the decision to the practitioners to decide upon which methods or tools suit their needs.

Moreover, the research is not profoundly studying the applicability of the suggested technical measures on IT systems in real practice. It is also not discussing nor comparing different Privacy Enhancing Tools (PETs) in details, due to the limitation in resources.

The framework highlighted the integration with the existing software development life cycles. However, there is no differentiation between distinct types of development methods such as agile or extreme programming. This limitation can be addressed in future works.

Finally, the culture and the perception of privacy within the data controller's environment is another aspect that was not covered in this research. It can be argued that the culture can play a significant role in realising privacy in both organisational and IT system levels. However, this research focus on the compliance of legal obligations which leave no choice to the organisation whether to comply or not.

Future Research

Future research can cover the limitations of this thesis. It may include applying the framework in different environments and on different IT systems to evaluate and improve the framework. Furthermore, expanding the framework to cover and integrate it with the existing organisation practices and management systems could be another area of future research.

Also, profound identification, analysis and testing of technical measures and implementing them on real IT systems could be another area of research. Although Data Protection by Design and by Default is mentioned only in the EU GDPR, future research can expand to other jurisdictions and compare them with the EU GDPR. Therefore, comprehensive framework that covers multiple jurisdictions can be built.

In many organisation, there are legacy systems it may be challenging to implement privacy; this could also be another future research area. Also, the future research on the system that was already can be useful since this is the case for many data controllers.
Finally, the researchers plan to expand this research and study various aspects around this framework including assessing the applicability and usability in a different environment with different systems. This future work may turn into a standard to realise privacy in IT systems in the future.

6. Reference

ARKIN, B., STENDER, S., and MCGRAW, G., 2005. Software Penetration Testing. *THE IEEE COMPUTER SOCIETY*.

ARNOULD, E.J., HEVNER, A.R., MARCH, S.T., and PARK, J., 2004. Design Science in Information Systems. *MIS Quarterly*, Vol. 28, No. 08.09.2007, p. 75–105.

BOOTE, D.N. and BEILE, P., 2005. Full-Text. *Educational Researcher*, Vol. 34, No. 6, p. 3–15.

CAVOUKIAN, A. and STOIANOV, A., 2007. Biometric encryption. *Biometric Technology Today*, Vol. 15, No. 3, p. 11.

CNIL, 2015. Privacy Impact Assessment. *Privacy Impact Assessment*.

DENG, M., WUYTS, K., SCANDARIATO, R., PRENEEL, B., and JOOSEN, W., 2011. A privacy threat analysis framework: Supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering*, Vol. 16, No. 1, p. 3–32.

DENSCOMBE, M., 2014. *The Good Research Guide: For Small-Scale Social Research Projects*. 5th ed. Berkshire: Open University Press.

DRESCH, A., LACERDA, D.P., and JR., J.A.V.A., 2015. *Design Science Research*. Springer. Springer International Publishing Switzerland.

EUROPEAN COMMISSION, 2011. Privacy and Data Protection Impact Assessment Framework for RFID Applications, No. January, p. 1–24.

HESLIN, P.A., 2009. Better than brainstorming? Potential contextual boundary conditions to brain writing for idea generation in organizations. *Journal of Occupational and Organizational Psychology*, Vol. 82, p. 129–145.

ICO, 2009. Privacy Impact Assessment Handbook Version 2.0. *Information Commissioner's Office*, p. 86.

IORIO, C.T. DI and CARINCI, F., 2013. Privacy and Health Care Information Systems : Where Is the Balance ?, p. 77–105.

JARNO J. VANTO, 2009. EuroPriSe - the New European Privacy Certification [online]. Available: https://iapp.org/news/a/2009-11-europrise-the-new-european-privacy-certification/ [Accessed 2017-4-17].

JOHANNESSON, P. and PERJONS, E., 2012. *A Design Science Primer*. CreateSpace.

JOHANNESSON, P. and PERJONS, E., 2014. *An Introduction to Design Science*. Springer International Publishing Switzerland. Kista: Springer International Publishing Switzerland.

KADAM, A.W. and VUTHA, S., 2012. Securing Sensitive Personal Data or Information : Using COBIT5 for India's IT Act.

MARKOPOULOS, P., MARTENS, J.-B., MALINS, J., CONINX, K., and LIAPIS, A., 2016. *Collaboration in Creative Design*. Springer International Publishing Switzerland.

MELLADO, D., FERNÁNDEZ-MEDINA, E., and PIATTINI, M., 2007. A common criteria based security requirements engineering process for the development of secure information systems. *Computer Standards & Interfaces*, Vol. 29, No. 2, p. 244–253.

MONREALE, A., RINZIVILLO, S., PRATESI, F., GIANNOTTI, F., and PEDRESCHI, D., 2014. Privacy-by-design in big data analytics and social mining. *EPJ Data Science*, Vol. 3, No. 1, p. 1–26.

NOTARIO, N., CRESPO, A., MARTIN, Y.S., DEL ALAMO, J.M., METAYER, D. LE, ANTIGNAC, T., KUNG, A., KROENER, I., and WRIGHT, D., 2015. PRIPARE: Integrating privacy best practices into a privacy engineering methodology. *Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015*, p. 151–158.

OETZEL, M.C. and SPIEKERMANN, S., 2014. A systematic methodology for privacy impact assessments: a design science approach. *European Journal of Information Systems*, Vol. 23, No. 2, p. 126–150.

OWASP, 2015. Application Threat Modeling - OWASP [online]. Available: https://www.owasp.org/index.php/Application_Threat_Modeling [Accessed 2017-4-17].

OWASP, 2017. Static Code Analysis - OWASP [online]. Available: https://www.owasp.org/index.php/Static_Code_Analysis [Accessed 2017-4-17].

RANDOLPH, J.J., 2009. A Guide to Writing the Dissertation Literature Review. *Practical Assessment, Research & Evaluation*, Vol. 14, No. 13, p. 1–13.

SCHAAR, P., 2010. Privacy by Design. *Identity in the Information Society - Special Issue*, Vol. 3, No. 2, p. 267–274.

SMART GRID TASK FORCE, 2014. Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems, p. 74.

WADHWA, K. and RODRIGUES, R., 2013. Evaluating privacy impact assessments. *Innovation: The European Journal of Social Science Research*, Vol. 26, p. 161–180.

WARREN, S.D. and BRANDEIS, L.D., 1890. The Right to Privacy, Vol. 4, No. 5, p. 193–220.

WRIGHT, D., GELLERT, R., GUTWIRTH, S., and FRIEDEWALD, M., 2011. Minimizing technology risks with PIAs, precaution, and participation. *IEEE Technology and Society Magazine*, Vol. 30, No. 4, p. 47–54.

WRIGHT, D. and RAAB, C., 2014. Privacy principles, risks and harms. *International Review of Law, Computers & Technology*, Vol. 28, No. 3, p. 277–298.

7. Appendix

7.1. Glossary

GDPR: General Data Protection regulations, will be in effect for all EU states starting from May 2018

Privacy by Design: Privacy and data protection were used interchangeably in this thesis, it refers to the same meaning

Data Controller: Data controllers and data processors are used interchangeably in this thesis, which refers to the entities who process, store or handle EU citizen's personal data. The full definition of the controller and the processor is available in GDPR article 4.

Organisational Measure: Any activities or processes the controller conduct to govern, manage, and maintain personal data

Technical Measure: Technical controls that can be implemented in IT systems

CNIL: An independent French administrative under the name of Commission Nationale de l'informatique et des libertés

7.2. Discussion on Design Science Method

In this section, the researchers critically discussed and analysed the methodology on how it can be applicable in this thesis.

7.2.1. Explicate the problem

In the design science, the first activity is the Explicate Problem which is to define, justify and find the causes of the problem (Johannesson and Perjons, 2014). The interview was selected to understand the problem from interviewees' aspect, clarify the scope of the problem, and explore the root causes to reach the better result. Also, since it is important to ensure the original results in Explicate Problem (Johannesson and Perjons, 2014), the literature review was performed to understand the EU regulations, business needs as well as other works that tried to define concrete steps to solve the issue. The detail can be found in section 3.

7.2.2. Requirements definitions

The goal of this second activity is finding a solution to close the gaps between problem and solution. This activity starts by identifying 1) what artefact should be developed, 2) what are the factors that can be considered the cornerstones to build the artefact, and 3) what are the requirements that are important for the stakeholders (Johannesson and Perjons, 2014) (Johannesson and Perjons, 2012). There are different types of requirements: functional, structural, environmental, and non-functional requirements according to (Johannesson and Perjons, 2014). Moreover, the same authors articulated the stages of defining the requirements of an artefact into two distinct stages, first is to outline the artefact, and second is to elicit the requirements. Due to the characteristics of the problem and the dependencies on several factors such as the law, the context, the risks associated with the processing the personal data, and the variety of the IT systems, a framework would be the outline of the artefact.

To define the requirements, the researcher can use any research method or research strategy (Johannesson and Perjons, 2014). Also, the researcher has to base the artefact on the existing literature that developed artefacts to solve similar problems to show significance and originality in the artefact (Johannesson and Perjons, 2014). In this research, the main requirements are derived from the EU General Data Protection Regulations. Therefore, the interview was once again selected as the method of collecting data to ensure the continuity from the business-orientated interview in Explicate Problem activity to a more technical interview in Requirements Definitions activity. Being the most common method for gathering requirements (Johannesson and Perjons, 2014), interviews allow the interviewer to ask stakeholder directly about what should be included in the artefact, hence explicated the requirements needed. Moreover, the literature review was used to find related solutions to the problem.

7.2.3. Design and development

To address the explicated problem and to fulfil the requirement from the previous activities, this third activity is the actual work where the artefact is designed and developed. As mentioned in (Johannesson and Perjons, 2014), the result of this activity will be prescriptive knowledge built-in the artefact with the descriptive knowledge via its design. In (Johannesson and Perjons, 2012), two sub-activities: 1) Generate, and 2) Search and select were used. However, in the newer book from the same author, this concept is redefined, and the sub-activities are divided into 1) Imagine and Brainstorm, 2) Assess and Select, 3) Sketch and Build, and 4) Justify and Reflect (Johannesson and Perjons, 2014). The researchers chose to follow the new sub-activities to provide more concise detail how the artefact was developed.

Imagine and Brainstorm

The Imagine and Brainstorm sub-activities aim to generate as many ideas as possible via the divergent thinking as well as convergent thinking (Johannesson and Perjons, 2014). According to (Markopoulos et al., 2016), the design processes usually involve these two models of thinking since the Divergent thinking allows new information and external information while the convergent thinking is used for analytically reduce the options to elaborate the design structurally. Furthermore, the idea generation can be both by individuals and by the group and provide the opportunity for the creative methods such as brainstorming and brainwriting to come to fruition.

Brainstorming is a useful method for the group because starts by allowing all participants to generate idea without any criticism. Then, the ideas are grouped and organised into some larger ones. Finally, participants prioritise the idea to select the useful ones in the next sub-activity, Assess and Select. On the other hand, brainwriting is a similar method to brainstorming but emphasises more on the creativity among individuals (Heslin, 2009). In this method, each participant needs to write down their idea in a certain time limit. Then, the ideas are passed around so that another participant can use the ideas as inspiration to create greater ideas (Johannesson and Perjons, 2014).

Assess and Select

All ideas generated and categorised in the previous sub-activity, Imagine and Brainstorm, are assessed based on the requirements defined in the previous activity, Define Requirements. The assessment aims to narrow down the solution space using decision models such as rational decision-making model where all relevant ideas are evaluated to find the most optimal one, or bounded rationality where the decision-making come to a halt when sufficiently good ideas are identified (Johannesson and Perjons, 2014). Also, various form of bias was taken into consideration during the idea selection process such as anchor bias which was considered by not put more weight on the first idea. Both researchers coherently made the decision to avoid personal belief as confirmation bias. Finally, each idea was assessed individually to make a right decision in contrast to sunk cost bias.

Sketch and Build

Selected idea was used to sketch an overview of the core functions and overall structure of the artefact. To sketch the functions of an artefact, the researchers used the use case diagram to layout input and output from the user perspective. Furthermore, the framework was built via the walk-through where both researchers can provide feedbacks throughout the development process similar to peer review (Johannesson and Perjons, 2012).

Justify and Reflect

This sub-activity aims to help designer prepare for future design and development by keeping track of design decisions made throughout the design process. This list of decisions is part of design rationale and contain reasons, justifications as well as alternative decisions, all help designers to communicate between the different project and served as profound knowledge for the subsequent projects. Also, the reflection about the procedures used for designing and developing the artefact was created to improve working practices in the future work.

7.2.4. Evaluation

The evaluation is a cornerstone in the design science research. According to (Johannesson and Perjons, 2014), there are six goals of evaluating an artefact; the primary goal is to determine the effectiveness of the developed artefact in solving the problem that it was intended to solve. One of the evaluation methods is to study the artefact in depth in the business environment by using case study (Arnould et al., 2004). In this research, a case study was used to evaluate the artefact in bona fide business environment. The system that the artefact will be evaluated on is an Artificial Intelligence platform that is used to provide services to EU Citizens. Evaluation types can be ex-ante, and ex-post; with ex-post considered to be more rigorous and require more time and resources than the ex-ante (Johannesson and Perjons, 2014). A case study is an ex-post method of evaluating an artefact under the design science research. Due to the limitation, however, the case study was conducted at a single site, also can be biased to the researcher's interests (Johannesson and Perjons, 2014). The results of the case study were used to improve the artefact.

7.2.5. Limitations

Arnould defined seven guidelines related to design science research (Arnould et al., 2004). These guidelines cover all aspects related to producing effective, and efficient artefact which solve the problem and contribute to the knowledge. The guidelines were considered during research phases starting from collecting the requirements to the evaluation of the requirements. Moreover, (Johannesson and Perjons, 2014) noted that the researcher should strive to understand and be able to answer *'How'* and *'Why'* questions of the artefact. Also, besides evaluating whether the developed artefact is working, the technological, psychological, and social factors should be considered. The limitation of design science may also arise in the evaluation stage when the artefact is evaluated in one environment or using one case study. Generalizability could be an issue since it will be difficult to generalise the results for all based the results of one case study.

7.3. Alternative Research Method

In this section, the researchers present and discuss alternate research method to the Design Science method that can be used to answer the research questions of this research paper.

#	Research Method	Applicability to the Research
1	Surveys and sampling	According to (Denscombe, 2014), surveys as a method for research can be selected when the researcher aims to measure parts of a social phenomenon or trending. Moreover, it can be used to gather facts from the field to test or validate a theory. Since this research paper aims to solve a practical problem, the surveys research method is not applicable.
2	Case studies	According to (Denscombe, 2014), case studies are often used to study the relationship between certain factors operating within particular settings, which is also considered not suitable research method for this research. Instead, case studies according to (Johannesson and Perjons, 2012) can be used to evaluate artefacts built using design research as a method. In this paper, the researchers used the case study as a method to evaluate the artefact and improve it based on the evaluation results.
3	Experiments	Experiments tend to be used in the controlled environment to identify the reasons what are the cause and to observe how these causes can impact other factors (Denscombe, 2014). In this research paper, the aim is not to solve or find the cause of the problem in a controlled environment, but to solve a practical problem for the public interest. Therefore, this method was not selected for this research.
6	Grounds Theory	Grounds theory can be used when the aim of the research is to clarify theories or establish new ones. It can also be used to discover new topics and provide new visions (Denscombe, 2014). It can be argued that Data Protection by Design and by Default can be based on

		theoretical concepts. However, the problem this research is trying to solve is a practical problem. Hence, grounds theory was not selected.
7	Action research	Action research can be used as an alternative research method for this research. According to (Denscombe, 2014), action research can be used to solve the particular practical problem in the defined environment. The same sources also mention that action research can also produce best practices and guidelines to solve the practical problem. However, there are criticisms of action research that the results can only apply to the research case and cannot be generalised beyond (Denscombe, 2014). Since the aim of this study is to solve general problem and keep the generalisation, action research was not selected as a research method for this work.

7.4. Discussion on Data Collection Method

The interview is one of the data collection methods that allow the researchers to interactively discuss with the interviewees by asking the follow-up questions based on the initial answers. However, there are some drawbacks when using interviews as the respondents' perspective, interests, as well as the personal attributes of the researchers could affect the outcome of the interview (Johannesson and Perjons, 2012, 2014). The researchers overcame these disadvantages by interviewing several respondents who shared some responsibility as well as having two interviewers asking a question about the same topic from a different aspect. According to (Denscombe, 2014), when collecting data from the interview, the feasibility should be considered. The attributes such as the accessibility to potential interviewees, cost and type of data play an important role when considering the form, structure as well as the quality of the data being collected. There are many forms of interviews such as a one-to-one interview, group interview and focus group interview (Denscombe, 2014). This research mainly implemented group interview and focused group interview because different opinions from people with different responsibilities are needed to understand the problem from a distinct perspective (Johannesson and Perjons, 2012), considering that there is no extra cost involved and the focus group and group interview can be efficiently organised. Also, the semi-structured interview was chosen to ensure the successful interviews as the issues and questions were prepared beforehand but still leave enough room for the interviewer to elaborate the point of interest (Denscombe, 2014)

A literature review is regarded as a demonstration of author's knowledge because *"a researcher cannot perform significant research without first understanding the literature in the field"* (Randolph, 2009) (Boote and Beile, 2005). The literature reviews were performed throughout the research process, and the discussion on how the researchers conducted them can be found in section 3.

7.5. Alternative Data Collection Method

In this section, the researchers present and discuss alternate data collection method to the Interviews and Literature Reviews that can be used to answer the research questions of this research paper.

#	Data Collection Method	Applicability to the Research
1	Questionnaire	A questionnaire can be used to substitute the interview method when developing an artefact. The main difference of the questions in questionnaires to those in the interview is that it is possible to have a question with predefined answers (Johannesson and Perjons, 2012). The questionnaire method could provide a better result if the system is developed per the common development standard with all documents intact. Hence questions can be tailored accordingly to receive an answer within the scope. However, the system is still in the beta stage. Therefore, the interview method is more suitable to the goal to explore and understand how the system works.
2	Observations	According to (Denscombe, 2014), there are two types of observational research: a *systematic observation* which usually used with the quantitative data and *participant observation* which typically associated with qualitative data (Denscombe, 2014). While this research also collect qualitative data, the participant observation is not applicable because this research aims to answer a research question by produce, create, or suggest a solution. If the observation were to be implemented, the researchers can only study the behaviour of the participant and make a conclusion without being able to interact with the participant. Therefore, the interview would provide a more direct answer on the issue.
3	Documents	If the purpose of this research were to explore what approaches others have been done to implement privacy requirement into IT system, the documents method would potentially be a primary data collection method. However, since the GDPR is a new regulation, the related papers and sources were still limited at the time when this research was conducted. Moreover, the researchers aimed to contribute to solving the issue, so the study on previous existing approaches became part of literature reviews which used to lay groundwork for the created framework.

7.6. Comparison of Privacy Impact Assessment and Other Approaches

The Information Commissioner's Office (ICO)'s PIA emphasis on stakeholders engagement at early development stage (Wright et al., 2011). This PIA uses a decision tree to determine the level of PIA to be conducted in the initial stage (ICO, 2009), similar to the European Commission (EC)'s PIA which has the processes separated into two phases: 1) Initial Analysis Phase, and 2) Risk Assessment Phase. Moreover, (Oetzel and Spiekermann, 2014) claimed that their step-by-step PIA could help achieving Privacy by Design. With their PIA, the created artefact consists of *'formal problem representation structure'* and *'simpler privacy regulation landscape'* will help practitioners in the analysis of privacy requirements and in making decision concerning privacy management respectively. Also, according to (Smart Grid Task Force, 2014), the DPIA helps entities to identify and anticipate risks concerning data protection, privacy and security, as well as to ensure rights to the protection of personal data and privacy. It can imply that the PIA has a wider scope as DPIA only focus on data protection, but not on other types of privacy (Wright and Raab, 2014). The following table shows the different activity of each approach categorised in phase.

Table 12 Comparison of PIA and Similar Approaches

Phase	ICO's PIA	EC's PIA Framework for RFID	Marie Caroline Oetzel and Sarah Spiekermann's PIA	European Commission's DPIA
Initialization	• Initial Assessment - Full/Small scale PIA, Privacy & another legal compliance check, DP compliance check	• Initial Analysis (decision tree)	• N/A	• Pre-assessment and criteria determining
Characterization	• Preliminary phase – project plan, project background paper, discussion with stakeholders • Preparation phase – stakeholder analysis, consultation strategy and plan, PIA Consultative Group	• Characterization of Application	• Characterization of the system	• Initiation • Identification, characterization and description of systems/applications processing personal data, including data flows
Assessment	• Consultation and analysis phase – Implement consultation plan issue register, privacy design features paper	• Identification of Risks • Identification and Recommendation of Measures	• Definition of privacy targets • Evaluation of degree of protection demand for each privacy target • Identification of threats for each privacy target • Identification and recommendation of	• Identification of relevant risks • Data protection risk assessment • Identification and Recommendation of measures and residual risks

			measures suited to protect against threats	
Documentation	• Documentation phase – consolidate decisions, produce PIA report • Review and audit phase – Undertake review, create review report	• Documentation of Resolution and Residual Risks	• Assessment and documentation of residual risks • Documentation of PIA process	• Documentation and drafting of the DPIA Report

7.7. Comparison of Privacy by Design Approaches

Table 13 compares the previous attempts to implement Privacy by Design in both system and organisation level. The data present the differences in the approach and whether it is risk-based or not. GDPR provision 25 stated that data protection by Default implementation should be based on risk, which in almost all the previous attempts, the risk-based approach was not the case.

Table 13 Comparison of literature regarding Privacy by Design

#	Author	Domain	Type (Mgmt./Eng.)	Risk Based	Notes
1.	A. Monreale (Monreale et al., 2014)	Big Data analytics and social mining	Eng. Technological Framework	n/a	Privacy-by-design in big data analytics and social mining
2.	European Commission	RFID Applications	MGMT	Yes	Privacy and Data Protection Impact Assessment Framework
3.	P. Caire	Ambient Intelligence Systems		n/a	Privacy challenges in Ambient Intelligence systems
4.	J. Ikonen	Gamecloud	Eng.	No	Development of game cloud with Privacy by Design
5.	J. Laakkonen	Gamecloud	Eng.	No	Continuous Development of Gamecloud with Privacy by Design
6.	Dawn N. Jultla	Heterogeneous Big Data	Eng.	No	PAUSE: A Privacy Architecture for Heterogeneous Big data environment
7.	M. Chibba	Big Data	Discussion	No	Privacy, consumer trust and big data: Privacy by Design and 3 C's (Consultation, Co-operation, and collaboration)
8.	A. Cavoukian	Smart Pricing program	High-level MGMT, small discussion about risk	n/a	Applying Privacy by Design best practices to SDG&E's Smart Pricing Program

9.	R. Hörbe	Federated Identity Management	More towards engineering.	No	Privacy by design in Federated Identity Management
10.	B. Vanderose	E-Governance	Mgmt/ high level	n/a	Privacy by design and Administrative Efficiency in E-Governance
11.	J. Bringer	Biometric System Architectures	Eng. Formal methods.	n/a	Privacy by Design in Practice: Reasoning about Privacy Properties of Biometric System Architectures
12.	G. Drosatos	e-Health Systems	High-Level Model – Towards engineering,	No	Towards Privacy by design in personal e-Health Systems
14.	J. Laakkonen	Smart Grid	Eng.	No	Abstract architecture for Smart Grid privacy analysis
15.	J. Laakkonen	Digital Game Platform	Engineering	Yes	Incorporating privacy into digital game platform design