

**T.C.
BAHÇEŞEHİR UNIVERSITY**

**BUILDING A FRAMEWORK TO MEASURE GDPR
IMPLEMENTATION IN SOFTWARE DEVELOPMENT**

Master Thesis

BOUSSAHA SOUMAYA

ISTANBUL, 2021

**T.C.
BAHÇEŞEHİR UNIVERSITY**

**GRADUATE SCHOOL
CYBER SECURITY MASTERS PROGRAM**

**BUILDING A FRAMEWORK TO MEASURE GDPR
IMPLEMENTATION IN SOFTWARE
DEVELOPMENT**

Master Thesis

BOUSSAHA SOUMAYA

Thesis Advisor: ASST.PROF.DR. ORHAN GÖKÇÖL

ISTANBUL, 2021



**T.C.
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL**

...../...../.....

MASTER THESIS APPROVAL FORM

Program Name:	Cyber Security (Thesis)
Student's Name and Surname:	BOUSSAHA SOUMAYA
Name Of The Thesis:	BUILDING A FRAMEWORK TO MEASURE GDPR IMPLEMENTATION IN SOFTWARE DEVELOPEMENT
Thesis Defense Date:	16 June 2021

This thesis has been approved by the Graduate School which has fulfilled the necessary conditions as Master thesis.

Assoc. Prof. Dr. Burak K2NTAY
Institute Director

This thesis was read by us, quality and content as a Master's thesis has been seen and accepted as sufficient.

	Title/Name	Signature
Thesis Advisor's	Dr. Orhan Gökçöl	
Member's	Prof. Dr. Alper Tunga	
Member's	Dr. Meriç Merih Aykol	

ACKNOWLEDGMENTS

I would like to thank my parents Abdalwahab Boussaha and Hamaici Hakima for all their sacrifices to ensure I get the best education and pursue my dreams. I could not have aspired to reach this educational level or accomplish this thesis without their encouragement and unconditional support.

I would also like to thank my supervisor, Dr. ORHAN GONKCOL, for all his support, advice, feedback and perspective throughout the research.

I would also like to thank whom-ever is taking the time to go through my work, this thesis topic was inspired by some struggles I had on as a product owner regarding complying with the needs of software development with the GDPR and privacy regulations in the software projects I was part of. As a master student in cybersecurity, I have chosen to put some light on this issue in the scope of my master research with the hope to provide other readers, professionals, or researchers a usable framework.

I have tried to make the language of the work as accessible as possible not only to researchers on Information security but also to people involved in the development of software or privacy enthusiasts. I hope my modest work would be useful to whom-ever comes across this thesis.

ISTANBUL, 2021
BOUSSAHA

SOUMAYA

ABSTRACT

BUILDING A FRAMEWORK TO MEASURE GDPR IMPLEMENTATION IN SOFTWARE DEVELOPEMENT

Soumaya Boussaha

Cyber Security Masters Program

Thesis Supervisor: Asst.Prof.Dr. Orhan Gökçöl

June 2021, 105 pages.

Building GDPR-compliant software is a problem that many companies and organizations face for their software development projects. The aim of this research is to shed some light on this topic in the hopes of developing a useful framework for assessing GDPR compliance levels in relation to the main 6 GDPR principles when developing software by applying a collection of controls as a development practice.

Keywords: GDPR, Software Development, Privacy, Personal Data Protection, Privacy by Design.

ÖZET

YAZILIM GELİŞTİRMEDE GDPR UYGULAMASINI ÖLÇMEK İÇİN BİR ÇERÇEVE OLUŞTURMA

Soumaya Boussaha

Siber Güvenlik Yüksek Lisans Programı

Tez Danışmanı: ASST.PROF.DR. ORHAN GÖKÇÖL

June 2021, 108 Sayfa

GDPR uyumlu yazılım oluşturmak, birçok şirket ve kuruluşun yazılım geliştirme projeleri için karşılaştığı bir sorundur. Bu araştırmanın amacı, bir geliştirme uygulaması olarak bir kontroller koleksiyonunu uygulayarak yazılım geliştirirken temel 6 GDPR ilkesiyle ilişkili olarak GDPR uyum düzeylerini değerlendirmek için yararlı bir çerçeve geliştirme umuduyla bu konuya biraz ışık tutmaktır.

Anahtar Kelimeler: GDPR, Software Development, Privacy, Personal Data Protection, Privacy by design.

TABLE OF CONTENTS

LIST OF TABLES	XIV
LIST OF FIGURES	XV
LIST OF ABBREVIATIONS	XVI
1. INTRODUCTION.....	1
1.1 MOTIVATION.....	1
1.2 GDPR OVERVIEW	4
1.2.1 The Six Main Principles Of G.D.P.R.....	5
1.2.1.1 Lawfulness, fairness, and transparency	5
1.2.1.2 Purpose limitation	8
1.2.1.3 Data minimization.....	9
1.2.1.4 Accuracy.....	9
1.2.1.5 Storage limitation.....	9
1.2.1.6 Integrity and confidentiality	10
1.2.2 The 8th Rights Enforced by GDPR	10
1.2.2.1 The right to information	11
1.2.2.1 The right to information	11
1.2.2.2 The right of access.....	11
1.2.2.3 The right to rectification	12
1.2.2.4 The right to erasure	12
1.2.2.5 The right to restriction of processing	13
1.2.2.6 The right to data portability	13
1.2.2.7 The right to object.....	14
1.2.2.8 The Right to Avoid Automated Decision-Making....	14
1.3 RESEARCH QUESTIONS AND RESEARCH METHODOLOGIES	14
2. LITERATURE REVIEW	17
2.1 GDPR IMPLEMENTATION THESIS CASE STUDY.....	17
2.1.1 Stakeholders Review	17
2.1.2 Architecture Review.....	18
2.1.3 Literature Review	18
2.1.4 Controls Proposed.....	18

2.2 OWASP ASVS CHECKLIST OVERVIEW.....	20
.2.2.1 General Data Protection Controls in ASVS	21
2.2.2 Personal Data Protection Controls in ASVS:.....	23
2.3 OWASP PRIVACY USER CHEAT SHEET OVERVIEW	23
2.3.1 Strong Cryptography	24
2.3.2 Support HTTP Strict Transport Security	24
2.3.3 Digital Certificate Pinning	25
2.3.4 Panic Modes	25
2.3.5 Remote Session Invalidation.....	26
2.3.6 Prevent IP Address Leakage.....	26
2.3.7 Honesty and Transparency	26
2.4 BS 10012:2017 27	
2.4.1 Identifying and Recording Uses of Personal Information	27
2.4.2 Risk Assessment and Treatment.....	28
2.4.3 Training and Awareness	29
2.4.4 Keeping PIMS Up-to-Date.....	29
2.4.5 Fair, Lawful and Transparent Processing	29
2.4.5.1 Records of privacy information	30
2.4.5.2 Accessibility of privacy information.....	30
2.4.5.3 Collection from third parties	30
2.4.6 Processing for Specific Legitimate Purposes.....	30
2.4.7 Data Minimization Principles.....	32
2.4.8 Accuracy	33
2.4.9 Retention and Disposal	33
2.4.10 Security Issues	34
2.5 ISO27701 STANDARD OVERVIEW.....	37
2.6 SOFTWARE DEVELOPMENT OVERVIEW	37
2.6.1 Software Development Life Cycle (SDLC).....	37
2.6.1.1 What are agile methodologies?	40
2.6.1.1.1 Phase number one: requirements.....	41
2.6.1.1.2 Phase number two: design.....	41
2.6.1.1.2.1 Software design	42

2.6.1.1.2.2	<i>Ui/ux design</i>	42
2.6.1.1.3	<i>Phase number three: development and coding</i>	42
2.6.1.1.4	<i>Phase number four: integration and testing.....</i>	42
2.6.1.1.5	<i>Phase number five: implementation and deployment.....</i>	43
2.6.1.1.6	<i>Phase six: review</i>	43
3.	ANALYSIS	44
3.1	PERSONAL DATA PROTECTION CONTROLS THROUGHOUT SOFTWARE DEVELEPEMENT	44
3.1.1	Controls Related To The Requirement And The Design Phases..	44
3.1.1.1	(C 1.1) Providing training and guidelines to the development team	44
3.1.1.1.1	<i>The goal of the security control.....</i>	44
3.1.1.1.2	<i>Explanation of the security control</i>	44
3.1.1.1.3	<i>Implementation of the security control.....</i>	45
3.1.1.2	(C1.2) Assess GDPR compliance gap on the legacy system	45
3.1.1.2.1	<i>The goal of the security control.....</i>	45
3.1.1.2.2	<i>Explanation of the security control</i>	45
3.1.1.2.3	<i>Implementation of the security control.....</i>	46
3.1.1.3	(C 1.3) Personal data identification (mapping)	47
3.1.1.3.1	<i>The goal of the security control.....</i>	47
3.1.1.3.2	<i>Explanation of the security control</i>	47
3.1.1.3.3	<i>Implementation of the security control.....</i>	48
	Tsource.....	49
3.1.1.4	(C 1.4) Design /Re-design access management system	49
3.1.1.4.1	<i>The goal of the security control.....</i>	49
3.1.1.4.2	<i>Explanation of the security control</i>	49
3.1.1.4.3	<i>Implementation of the security control.....</i>	50
3.1.1.5	(C 1.5) Sharing data with 3rd prties.....	50
3.1.1.5.1	<i>Tthe goal of the security control.....</i>	50

3.1.1.5.2	<i>Explanation of the security control</i>	52
3.1.1.5.3	<i>Implementation of the security control.....</i>	52
3.1.2	Controls Related to The Development Phase and The Testing Phase.....	52
3.1.2.1	(C 2.1) Encryption.....	52
3.1.2.1.1	<i>Yhe goal of the security control.....</i>	52
3.1.2.1.2	<i>Explanation of the security control</i>	52
3.1.2.2	(C 2.2) Data minimization by design	54
3.1.2.2.1	<i>The goal of the security control.....</i>	54
3.1.2.2.2	<i>Explanation of the security control</i>	54
3.1.2.2.3	<i>Implementation of the security control.....</i>	56
3.1.2.3	(C 2.3) Logs management	56
3.1.2.3.1	<i>The goal of the security control.....</i>	56
3.1.2.3.2	<i>Explanation of the security control</i>	56
3.1.2.3.3	<i>Implementation of the security control.....</i>	57
3.1.2.4	(C 2.4) Clean databases automatically	57
3.1.2.4.1	<i>The goal of the security control.....</i>	57
3.1.2.4.2	<i>Explanation of the security control</i>	57
3.1.2.4.3	<i>Implementation of the security control.....</i>	58
3.1.2.5	(C 2.5) Automatic file removal service	58
3.1.2.5.1	<i>The goal of the security control.....</i>	58
3.1.2.5.2	<i>Explanation of the security control</i>	58
3.1.2.5.3	<i>Implementation of the security control.....</i>	58
3.1.2.6	(C 2.6) Over-writing data in memory/disk.....	60
3.1.2.6.1	<i>The goal of the security control.....</i>	60
3.1.2.6.2	<i>Explanation of the security control</i>	60
3.1.2.6.3	<i>Implementation of the security control.....</i>	60
3.1.2.7	(C 2.7) Pseudonymization by Default	61
3.1.2.7.1	<i>The goal of the security control.....</i>	61
3.1.2.7.2	<i>Explanation of the security control</i>	61
3.1.2.7.3	<i>Implementation of the security control.....</i>	62

3.1.2.8 (C 2.8) Develop a personal data automatic search method.....	62
3.1.2.8.1 <i>The goal of the security control</i>	62
3.1.2.8.2 <i>Explanation of the security control</i>	62
3.1.2.8.3 <i>implementation of the security control</i>	63
3.1.2.9 (C2.9) Masking data in the test environment	63
3.1.2.9.1 <i>The goal of the security control</i>	63
3.1.2.9.2 <i>Explanation of the security control</i>	63
3.1.2.9.2 <i>Implementation of the controls</i>	64
3.1.2.10 (C 2.10) Back up data.....	64
3.1.2.10.1 <i>The goal of the security control</i>	64
3.1.2.10.2 <i>Explanation of the security control</i>	64
3.1.2.10.2 <i>Implementation of the security control</i>	64
3.1.2.11 (C 2.11) Informed consent.....	65
3.1.2.11.1 <i>The goal of the security control</i>	65
3.1.2.11.2 <i>Explanation of the security control</i>	65
3.1.2.11.3 <i>Implementation of the security control</i>	65
3.1.2.12 (C 2.12) Provide privacy settings features.....	66
3.1.2.12.1 <i>The goal of the security control</i>	66
3.1.2.12.2 <i>Explanation of the security control</i>	66
3.1.2.12.3 <i>Implementation of the security control</i>	66
3.1.2.13 (C 2.13) Refactoring old features.....	67
3.1.2.13.1 <i>The goal of the security control</i>	67
3.1.2.13.2 <i>Explanation of the security control</i>	67
3.1.2.13.2 <i>Implementation of the security control</i>	68
3.1.2.14 (C 2.14) Hashing.....	68
3.1.2.14.1 <i>The goal of the security control</i>	68
3.1.2.14.2 <i>Explanation of the security control</i>	68
3.1.2.14.3 <i>Implementation of the security control</i>	69
3.1.2.15 (C 2.15) Enforce data subject to update data upon expiry.....	69
3.1.2.15.1 <i>The goal of the security control</i>	69

3.1.2.15.2 Explanation of the security control	69
3.1.2.15.3 Implementation of the security control.....	69
3.1.2.16 (C 2.16) Cache controls	70
3.1.2.16.1 The goal of the security control.....	70
3.1.2.16.2 Explanation of the security control	70
3.1.2.16.3 Implementation of the security control.....	70
3.1.3 Controls Prior to Deployment	70
3.1.3.1 (C3.1) Penetration testing	70
3.1.3.1.1 The goal of the security control.....	70
3.1.3.1.2 Explanation of the security control	72
3.1.3.1.3 Implementation of the security control.....	72
3.1.3.2 (C3.2) Personal data - flow diagram.....	72
3.1.3.2.1 The goal of the security control.....	72
3.1.3.2.2 Explanation of the security control	72
3.1.3.2.3 Implementation of the security control.....	73
3.1.3.3 (C3.3) Update/Create privacy notice	73
3.1.3.3.1 The goal of the security control.....	73
3.1.3.3.2 Explanation of the security control	73
3.1.3.3.3 Implementation of the security control.....	74
3.1.3.4 (C3.4) Provide a Method for Data Requests	74
3.1.3.4.1 The goal of the security control.....	74
3.1.3.4.2 Explanation of the security control	75
3.1.3.4.3 Implementation of the security control.....	75
3.1.3.5 (C3.5) Update internal privacy policy and processes	75
3.1.3.5.1 The goal of the security control.....	75
3.1.3.5.2 Explanation of the security control	75
3.1.3.5.2 Implementation of the security control.....	76
3.1.3.6 (C3.6) Update/Create Cookies policy	76
3.1.3.6.1 The goal of the security control.....	76
3.1.3.6.2 Explanation of the security control	76
3.1.3.6.3 Implementation of the security control.....	76

3.1.4 Controls Post Deployment	77
3.1.4.1 (C4.1) Create/Update - incident response program .	77
3.1.4.1.1 <i>The goal of the security control</i>	77
3.1.4.1.2 <i>Explanation of the security control</i>	77
3.1.4.1.3 <i>Implementation of the security control</i>	78
3.1.4.2 (C4.2) Penetration testing	78
3.1.4.2.1 <i>The goal of the security control</i>	78
3.1.4.2.2 <i>Explanation of the security control</i>	78
3.1.4.2.3 <i>Implementation of the security control</i>	78
3.1.4.3 (C4.3) Testing and Auditing	79
3.1.4.3.1 <i>The goal of the security control</i>	79
3.1.4.3.2 <i>Explanation of the security control</i>	79
3.1.4.3.2 <i>Implementation of the security control</i>	79
3.1.4.4 (C4.4) Archiving practices	79
3.1.4.4.1 <i>The goal of the security control</i>	79
3.1.4.4.2 <i>Explanation of the security control</i>	80
3.1.5 Summary of the analysis	80
3.2 ANALYSIS OF THE CONTROLS	83
3.2.1 Principle's definition	83
3.2.2 Controls Mapping	83
3.2.2.1 Fairness, lawfulness, and transparency.....	84
3.2.2.2 Purpose limitation	85
3.2.2.3 Data minimization.....	86
3.2.2.4 Accuracy	86
3.2.2.5 Storage limitation.....	87
3.2.2.6 Integrity and confidentiality	87
3.2.3 Adding labels	88
4.FINDINGS.....	96
4.1 DEFINING THE ASSESMENT MODEL FRAMEWORK.....	96
4.2 CASE STUDY 96	
4.2.1 Company's Presentation	97
4.2.2 Presentation of The Case Study Project.....	97

4.2.2.1 Features to be developed in this project.....	98
4.2.2.2 Development methodology explained	98
4.2.2.3 Assessment.....	99
4.2.2.4 The output of the Analysis of the case study	106
4.2.3 Limitations of The Case Study	109
4.3 MATURITY FRAMEWORK ENHACEMENT	109
5. DISCUSSION AND CONCLUSION.....	113
5.1 LIMITATION OF THE THESIS OUTPUT AND PERSPECTIVE	
RESEARCH	113
5.2 CONCLUSION	114
REFERENCES	115

LIST OF TABLES

Table 1.1: Developers' for GDPR compliance survey	2
Table 1.2: Survey assessment logic for GDPR principles	3
Table 1.3: 5 main issue's software developers encounter to implement GDPR	4
Table 3.1: A personal data mapping example	48
Table 3.2: Principle 1 controls.....	84
Table 3.3: Principle 2 controls.....	85
Table 3.4: Principle 3 controls.....	86
Table 3.5: Principle 4 controls.....	86
Table 3.6: Principle 5 controls.....	87
Table 3.7: Principle 6 controls.....	87
Table 3.8: Must have / good have controls	88
Table 3.9: Must have / good have Controls Mapping to principles.....	93
Table 4.1: Case study Assessment.....	99
Table 4.2: Analysis of the case study output based on the development phases.....	108

LIST OF FIGURES

Figure 1.1: Fines Issued for GDPR Violations , by type of violation.....	1
Figure 1.2: Pyramid chart.....	15
Figure 2.1: An abstract Agile model view	41
Figure 3.1: GDPR GAP Assessment tool example	47
Figure 3.2: Sleuth kit review	61
Figure 3.3: an example of Pseudonymization by default.	62
Figure 3.4: Facebook 2021 privacy settings and tools.....	66
Figure 3.5: Facebook 2021 privacy basics	67
Figure 3.6: Hash function.....	68
Figure 3.7: View controls throughout the development phases	82
Figure 3.8: View mapping logic	84
Figure 4.1: Existing legacy model of case study	97
Figure 4.2: Radar chart Case study assessment.....	107
Figure 4.3: Graphical Chart Case study assessment.	107
Figure 4.4: SPICE process maturity levels (ISO/IEC15504-6).....	111

LIST OF ABBREVIATIONS

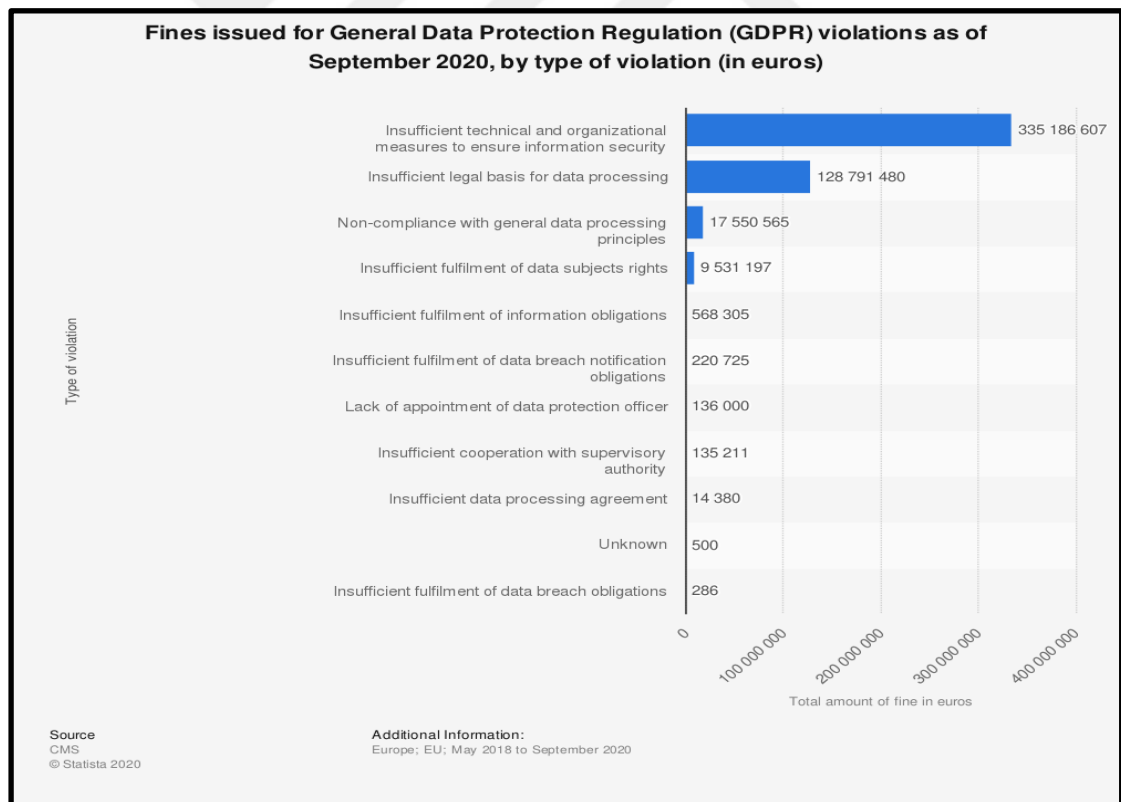
GDPR	:	General Data Protection Regulation
OWASP	:	Open Web Application Security Project
SDLC	:	Software development Life Cycle
PII	:	Personal Identifiable Information
HTTP	:	Hypertext Transfer Protocol
HTTPS	:	Hypertext Transfer Protocol Secure
TLS	:	Transport Layer Security
SSL	:	Secure Socket Layer

1. INTRODUCTION

1.1 MOTIVATION

In the last decade, data privacy has become an enormous concern not only in Europe but around the world. However, a lack of expertise and knowledge may be a major hurdle for organizations aspiring to adopt best practices or at the very least meet the critical needs required by regulations such as the GDPR in Europe. Considering this discussion, a European survey shows the number of fines issued for GDPR violations as of September 2020, segmented by the form of violation (see figure 1.1). According to this recent study, many fines are still issued because of inadequate technological and organizational measures to ensure information security .

Figure 1.1: Fines Issued for GDPR Violations , by type of violation



Source: Statista. 2020.

As a result, it is critical to adopt a system that enables organizations to assess the maturity of their GDPR implementation when managing their processes, including developing software processes being particularly mature in terms of complying with Privacy laws. In many cases, GDPR compliance is the last thing a company examines before releasing software to production, or we are checking the requirements in an Organic non-structured fashion. Unfortunately, this is a clear mistake, as this practice is in direct opposition to the GDPR's privacy by design principle. Considering this thesis, we will tackle this concern.

Firstly, to further support this Issue we would like to start by making a review of a Survey: "I'm all Ears! Listening to Software Developers on Putting GDPR Principles into Software Development Practice" This Study was published by the University of La Trobe by Abdulrahman, A.I & Nalin, A.G.A in 2021; It aims to identify the barriers that prevent software developers from incorporating privacy into their software development practices complying with GDPR principles. Participants were asked during an interview to design an online web health application that allows patients to register/log in and communicate with healthcare experts over the internet for a fee. In that application, clients were supposed to be able to speak with health professionals about their problems. As for health, experts are supposed to gain money by guiding the patient. For every paid service by the clients, the application will gain from a commission.

The segment that was interviewed were developers across the globe and not from one specific region to achieve diversity in total 17 participants were interviewed as follows in the following table.

Table 1.1: Developers' for GDPR compliance survey

	Minimum.	Maximum.	Average.
Age(years)	24	38	29
Duration of interview(minutes)	39	56	46
Experience(years)	4	11	6
Gender information.			
Male	Female		
17 participants	3 participants		
The average rate of coding (per week)	5 days		

Source: Abdulrahman, A.I & Nalin, A.G.A,2021.

Participants were shown a variety of scenarios through UML diagrams and asked how they would incorporate the features mentioned in the diagrams.

The six principles of GDPR law were investigated based on their response; table number 1.2 was extracted from the study, this table explains how they evaluated each GDPR principle during the interview.

Table 1.2: Survey assessment logic for GDPR principles

Principle	Testing Method
Lawfulness, Fairness and Transparency	In this scenario, we collected doctors, nurses, and clients' data. Developers were required to share up-to-date information with the data subjects on how their data will be processed. They were also required to promptly notify data subjects of any data breaches when they occur. The developers were requested to explain how they would achieve these functions in the health application.
Integrity and Confidentiality	This application will require data to be protected from getting into the wrong hands, as this can lead to tampering with the data. The application also has a functional requirement that allows different users to communicate remotely. It was required that security measures are applied so that even if this data gets into the wrong hands, it cannot be edited or be read. Developers were requested to state how they will achieve this and the techniques they will use. In this health application, it was also a requirement for users to authenticate before accessing the application. Data storage and security measures were also implemented in this application to ensure that unauthenticated persons do not access the data. Developers taking part in this research were requested to explain how they can achieve that in the given health application.
Purpose Limitation	In this health application, it was a requirement to ensure that the data subjects know the reason for the collection of each data. It was also required that this data was not to be used for any other purposes other than the stated ones. Participants of this research were required to explain how they could achieve this in this application.
Data Minimization	In this health scenario, there were functional requirements such as login, registration, and payment process. One of the needs of the health scenario was to collect the least amount of user data as possible in each function. For example, the login may only require an email and password. Participants of the research were asked to explain how they would accomplish this need in the health application.
Accuracy	In this health application, one of the requirements was to approve the registration of users. It was also required to prevent unauthorized persons from accessing users' data as they can tamper with the accuracy of the data. Each participant had to explain how he would achieve this for this health application.
Storage Limitation	This health application required that the stored records would not be held by the system for an extended period once they are no longer needed. The participants of this research were requested to explain how they would use different techniques to achieve this.

Source: Abdulrahman, A.I & Nalin, A.G.A,2021.

As an output, the survey pointed out five main issues as shown in the following table:

Table 1.3: 5 main issue's software developers encounter to implement GDPR

Issues	Representative Quotes	Number of responses
Lack of familiarity with GDPR	I am not familiar with GDPR	17
Organizations	To use GDPR effectively depends on organization you're working for.	8
Lack of resources and guidelines	Don't have guidelines on how to do it and where to refer to	16
Not familiar with techniques	I don't have an idea of technique to do that	14
Privacy is not a primary requirement	My greatest concern is functional requirements.	9

Source: Abdulrahman, A.I & Nalin, A.G.A,2021.

Seventeen out of the participants mentioned not being familiar with GDPR, on the other hand a majority of sixteen of the participants addressed the lack of guidelines and resources, furthermore fourteen of participants complained regarding the lack of techniques to implement the GDPR principles. Therefore, we can appoint the lack of knowledge, guidelines, and training as the primary root cause.

It is also important to note that almost half of the participants of this survey also pointed out that the organizational challenges push the developers to act accordingly to the organization culture and If take this assumption and put it side by side with the result of Figure 1.1 of the firstly stated survey we can also assume that usually organizations lack the understanding on how to deliver a system or software while ensuring GDPR compliance. This adds to the value and the purpose of this thesis as we are aiming to build a guideline to address GDPR compliance during software development.

1.2 GDPR OVERVIEW

Before we go any further, it's important to understand what GDPR stands for. GDPR refers to "General Data Protection Regulation," as the names suggest. This law is intended to safeguard people's data (also defined as data subjects) across Europe.

1.2.1 The Six Main Principles Of G.D.P.R

The GDPR article number 5 briefly mentions 6 main GDPR concepts, which are as follows (1):

1.2.1.1 Lawfulness, fairness, and transparency

The first principle is self-evident: organizations must collect data from data subjects in a lawful way, which means they must ensure that their data collection procedures are ethical and transparent to the data subject. This principle is Mostly related to the legal aspect of gathering and processing personal data as these processes must be handled legally and fairly, therefore all the controls falling under the legal aspect of the processing would be contributing to this principle.

The data subject must be notified on how his/her data is being processed and to which purpose, the data subject must be aware of the controller as well as all data processors or Third parties to whom the data might be transmitted before.

Companies and organizations must have a clear understanding of the GDPR, and its data collection rules to remain lawful. The GDPR's concept of lawfulness is detailed in Article 6 of the GDPR (1). The processing is defined as lawful only if 1 out of 6 situations applies:

“

- i. The data subject has given consent to the processing of his or her data for one or more specific purposes.
- ii. Processing is necessary for the performance of a contract to which the data subject is party or to take steps at the request of the data subject prior to entering a contract.
- iii. Processing is necessary for compliance with a legal obligation to which the controller is subject.
- iv. Processing necessary to protect the vital interests of the data subject or another natural person.
- v. Processing is necessary for the performance of a task carried out in the public interest or the exercise of official authority vested in the controller.

- vi. Processing is necessary for the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, where the data subject is a child. (Shall not apply to processing carried out by public authorities in the performance of their tasks.)”

GDPR Article 6.

Companies and organizations must have a clear understanding of the GDPR, and its data collection rules to remain lawful. The GDPR's concept of lawfulness is detailed in Article 6 of the GDPR. (1)

The processing is defined as lawful only if 1 out of 6 situations applies:

- “
- i. *The data subject has given consent to the processing of his or her data for one or more specific purposes.*
 - ii. *Processing is necessary for the performance of a contract to which the data subject is party or to take steps at the request of the data subject prior to entering a contract.*
 - iii. *Processing is necessary for compliance with a legal obligation to which the controller is subject.*
 - iv. *Processing necessary to protect the vital interests of the data subject or another natural person.*
 - v. *Processing is necessary for the performance of a task carried out in the public interest or the exercise of official authority vested in the controller.*
 - vi. *Processing is necessary for the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, where the data subject is a child. (Shall not apply to processing carried out by public authorities in the performance of their tasks.)”*

GDPR Article

6.

It's worth mentioning that the point (A) mentioned in the previous paragraph taken from GDPR article 6 does not apply When our data subject is a child, (when he/she is below the age of 16).

In that case, Article 8, clause 1 (1). states that “such processing shall be lawful only if and to the extent that consent is given or authorized by the holder of parental responsibility for the child.”

To remain transparent with data subjects, on the other hand, Companies and organizations should state in their privacy policies and privacy notice the type of data they collect and the reason they're collecting it from the data subjects in a plain understandable way.

There also exist several cases when consent is not needed to process the personal data. Those cases often reflect the need for homeland security, networks, and systems security, commercial processing, and public interest or duties, or public organizations. Such situations and cases are specified in Article 6 of the GDPR. (1)

The cases when consent is not needed are:

- a. Processing of the data is necessary to prepare or execute a contract to which the data subject is a party (e.g., doing a background check).
- b. The data must be processed to comply with a legal obligation, e.g., when the organization receives an order from the court in their jurisdiction.
- c. It is necessary to process the data to save a person's life.
- d. Processing is needed to perform a task in the public interest or to carry out some official functions.
- e. The organization has a legitimate interest to process a person's data. Although this may be the most flexible lawful basis, the "fundamental rights and freedom of the data subject" will always override the interest of an organization. This is especially true for the data of children.

Of course, there exist specific requirements for situations where the above exceptions are used and implemented, such as the basis must be documented, and the data subject should be notified.

If the organization later decides to change its justification, a good reason is needed. The new reason must then be documented, and the data subject must be notified again.

1.2.1.2 Purpose limitation

Organizations must only collect personal data for a clear and specific purpose and only collect data for as long as necessary to complete that purpose.



Once the processing is done, if archiving is to be considered under the scope of the public interest or for scientific, historical, or statistical purposes then for this matter the processing is given more freedom detailed in Article 89 of GDPR. (1)

Personal data must be processed only for the sole purpose of its collection before its collection, if additional processing is needed then licit consent from the data subject must be given. (2)

1.2.1.3 Data minimization

Companies and organizations must restrain their processing to only the information that's fundamental for accomplishing the purpose of the process.

This decreases the impact of a data breach as the unauthorized person has access to minimal information. This moreover helps to keep the information accurate and up to date because it is much simpler updating a littler parcel of information hence ensuring information integrity. Before processing or collecting data, it is necessary to determine how much and what is the data that is needed to fulfill the purpose of the processing data controllers should not collect more than the needed data to be processed, techniques such as anonymization are used to reduce the amount of data that is shared that's an example of data reduction and data economy. (1)

1.2.1.4 Accuracy

The Accuracy principle in GDPR states that “every reasonable step must be taken” to delete or correct data that are wrong or incomplete. Individuals have the right that inaccurate or inadequate data be deleted or corrected within 30 days . (1)

1.2.1.5 Storage limitation

Additionally, organizations must be compelled to erase personal data when it's no longer essential to fulfill the purpose.

(1): General Data Protection Regulation, 2016. 9

(2) *British Standards Institution*, BS 10012 Personal Information Management System, 2017, <https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/> [Accessed 01 December 2020] , pp1-42.

Any doubt that may arise regarding how long an organization should keep personal data should be consulted by a legal professional as this would vary between industries and needs. Personal data that no longer serves a purpose must be deleted after the expiration of the legal business-related obligations, data is usually attributed to a retention period, Once reached the data should be deleted.. (1)

1.2.1.6 Integrity and confidentiality

Personal data is a very sensitive asset that should be subject to data secrecy. It must be secured with suitable measures against unauthorized access, illegal processing or disclosure, loss or accidental destruction, or corruption. Personal data must keep its integrity by being correct and complete.

The GDPR states that “personal data must be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and accidental loss, destruction or damage, using appropriate technical or organizational measures”.

The GDPR is intentionally unclear around what measures organizations ought to take since organizational best practices are continually changing. Right now, organizations ought to encrypt and/or pseudonymize personal data wherever conceivable, but they should consider whatever other alternatives are possible and feasible. (1)

1.2.2 The 8th Rights Enforced by GDPR

Aside from the six principles GDPR introduced, GDPR also granted the data subjects 8 main rights, in the scope of this thesis we would be focusing on the principles, but we want to give a brief introduction on these 8 rights to clear any misconceptions readers may have:

1.2.2.1 The right to information

These right states that a data subject has the right to inquire about the type of data that a data controller processes and why the data controller requires it. Articles 13 and 14 of the GDPR grant data subject to the right to information.

When personal data is collected directly from data subjects, the information that needs to be communicated to the data subject by the controller is detailed under Article 13
When you collect data about the data subject from a third party or indirectly, you are accountable under Article 14. (1)

1.2.2.1 The right to information

These right states that a data subject has the right to inquire about the type of data that a data controller processes and why the data controller requires it. Articles 13 and 14 of the GDPR grant data subject to the right to information.

When personal data is collected directly from data subjects, the information that needs to be communicated to the data subject by the controller is detailed under Article 13
When you collect data about the data subject from a third party or indirectly, you are accountable under Article 14. (1)

1.2.2.2 The right of access

The data subject has the right to access the personal data that you process about them. described in article 15 of GDPR. (1)

Aside from the data itself, data subjects have the right to request this additional information :

- i. The goal of processing the data and how it is done?
- ii. Personal data categories that are involved
- iii. Who has access to the data? Especially if the data is shared outside Europe.
How long is the data intended to be kept

- iv. If data was not collected from the data subject, what was the source that shared the data?
- v. Is there any use of profiling?

1.2.2.3 The right to rectification

The right to rectification is also tied to one of the six GDPR Privacy Principles, Data Accuracy since it provides a basis for maintaining accurate data. This right is important and must be encouraged to be used by the data subject by the organization because the Other Data subject's privacy is put at risk by keeping inaccurate data.

For example, if an organization is keeping old phone numbers that don't belong to the data subject who gave consent anymore then they risk calling a different data subject who now owns the number but didn't give consent. GDPR violations include keeping data on data subjects and contacting them without their consent as stated in article 16 of the GDPR . (1)

1.2.2.4 The right to erasure

The data subject's right to erasure, also known as the right, to be forgotten, is defined in Article 17. It states that every Data subject has "the right to be forgotten," meaning that, upon request, companies are required to discard any personal data related to a particular individual. Therefore, software, systems, or databases should include tools that let organizations isolate and delete personal data as needed to be able to delete it.

In certain contexts, a request to access the right to erasure does not require the processor or data controller to respond. These circumstances are outlined in the GDPR in article 17 as follows (1) :

- i. When there is a right to freedom of speech and information at stake.
- ii. When there is a legal duty to comply with and the public interest to be served.
- iii. When there are reasons for, public concern in the field of public health.
- iv. When the processing adheres to the Article 89 clause 1 guidelines (or the public interest, historical or scientific purposes, or statistics purposes).

- vi. When the purpose of the processing is to "establish, exercise, or defend legal claims".

1.2.2.5 The right to restriction of processing

Article 18 of GDPR describes the right of the data subject to seek the restriction of processing of their data in certain conditions. (1)

That means the organization in charge of the personal data must temporarily stop processing it if one of the following conditions are met :

- i. The data subject doubts the data's accuracy.
- ii. The data subject objects to unlawful processing and chooses to have the processing limited rather than have their data erased.

The data controller does not need the information for processing, but it is required for the "establishment, exercise, or defense of a legal claim. "Article 18 clause 3 states that if the organization is willing to lift the restriction the data subject must be informed.(1)

1.2.2.6 The right to data portability

GDPR grants the right to data subjects to receive their data from a data processor or data controller in a structured way and an understandable format, this also grants the data subject the right to request a data controller to share the data they have on the data subject to another controller.

This right however applies only to personal data; it is important to understand that this also affects pseudonymous data as the data can be linked back to the individual risks databases by providing the correspondent Identifier therefore this data is in scope for this right.

For instance, if you provide mobile phone services, you'll need to configure your software, so it allows users to take their phone numbers to another service provider .

1.2.2.7 The right to object

The GDPR states that if a data subject expresses their right to object then Unless the controller can assert convincing valid reasons for the processing that outweigh the data subject's interests, privileges, and freedoms, or for the establishment, exercise, or defense of legal claims, the controller must no longer process the personal data.

When the processing relates to direct marketing, the data subject has the right to object at any time, according to Article 21 clause 2, and the organization must comply. If a data subject unsubscribes from marketing emails, for example, the organization must not send them any more emails unless and until they provide consent. (1)

1.2.2.8 The Right to Avoid Automated Decision-Making

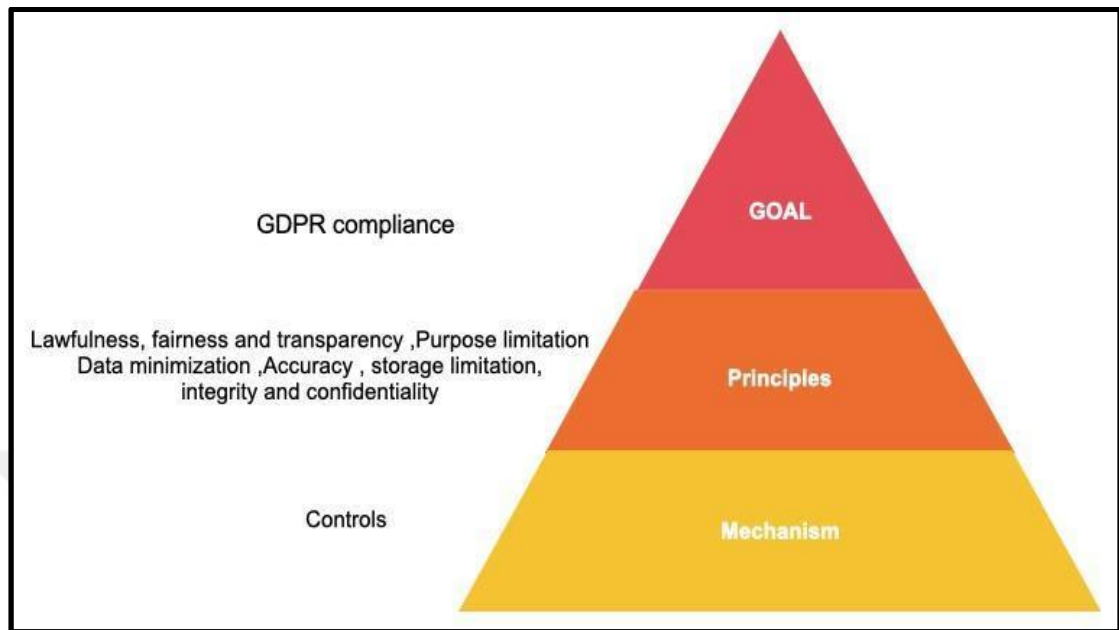
GDPR article 22 states that the data subject has the right not to be subjected to a decision based solely on automatic processing, including profiling, which produces legal effects concerning him/ her or another similar effect. (1)

1.3 RESEARCH QUESTIONS AND RESEARCH METHODOLOGIES

The goal is to understand how we can have a usable framework for organizations to comply with GDPR's principles while developing software and being able to evaluate the level of compliance they achieved.

We would like to share this pyramid chart that we drafted (see figure 1.2) to give an idea about how we intend to approach the compliance with GDPR principles dilemma within this thesis. "GDPR compliance" is strongly connected with adhering to the GDPR Principles therefore, a set of controls must be adopted as a mechanism to ensure compliance. So, on an operational level, a set of security measures and controls are to be implemented to ensure that the principles of GDPR are met and therefore ensure better compliance with the regulation.

Figure 1.2: Pyramid chart



We intend to address the following research questions as a part of this investigation:

Research question 1 -What are the controls and the best practices that exist to ensure the implementation of GDPR regarding software development?

To respond to this research question (RSQ1), we will conduct a literature review to put together a list of best practices and controls that might be used during software development to comply with GDPR. We'd refer to a few of the related research that has been conducted in this area. We'd also look at some standard documents published: including the OWASP ASVS, ISO27701, and BS10012:2017 standards.

Research question 2 -What are the high-level principles That need to be implemented to ensure GDPR compliance?

To respond to this question, we would examine high-level needs based on GDPR articles, then map those needs to the controls and practices we identified as a response to Research Question 1.

Research question 3- What could be a framework to assess the maturity level for software GDPR compliance?

We will introduce a framework as a solution to assess the implementation success level based on the answers given for research question 1 and research question 2 as well as the literature review that we will conduct. We would End the work with a Case study to test our framework.



2. LITERATURE REVIEW

2.1 GDPR IMPLEMENTATION THESIS CASE STUDY

The authors of this thesis: “head power Oy Master’s thesis March 2019 Author(s) Reini, Pasi /JAMK university of applied science” (Reini & Pasi ,March 2019) have focused on GDPR implementation practices for a case study of a Finnish company Named: Headpower Oy that is a data controller. published in March 2019. in this Thesis the author addressed a real case study for an application developed by the company that has 15 years of history to which additional development is still happening. During its lifetime, it was created by many software developers. Most of the original developers have moved on, and as a result, some fragment of the application's design elements has been forgotten. This is what attracted my attention to this thesis as building compliant software is easier when we start from scratch, but the main challenge is having a legacy system and dealing with the functional constraints it brings up. Some obsolete features were deleted, while others have been introduced. Databases may contain information about features that no longer exist which means this increases the vector of risk of non-compliant personal data that is still there on the system.

2.1.1 Stakeholders Review

During the study, the authors of the thesis did a stakeholder analysis to understand the roles of these stakeholders and what type of access they need for the huge amount of personal data that is stored in the application’s databases. The main stakeholders that the author mentioned were:

- i. help desk personal: The help desk is classified into tiers of service, with each level having access to the system as well as customer information. The help desk’s main task is to respond to customers’ requests, when addressing customer issues, the help desk staff interact with the software development team regularly.
- ii. Development Team: Including developers, testers. their main task is to build the application, solve the bugs, enhance the application and the user experience. The

development methodology they use is Scrum with a sprint of 10 days, each release has normally 4 sprints.

- iii. Other roles were presented: Product Owners, Salespersons, Corporation management in some use cases of these roles also process personal data.

2.1.2 Architecture Review

Here the author of the thesis explained on a high level what are the components of the system concerning the functioning application, since the application is a web application the author's thesis presented what are the servers and APIs involved and which ones are storing processing or transmitting PII.

2.1.3 Literature Review

The author conducted a literature review to see if there was any framework or guideline that was available to direct users to implement GDPR for software development see reference for detailed review (Reini & Pasi ,March 2019) the researcher concluded that according to previous surveys, there are currently no common guidelines on how it should be done. Instead, tiny bits of knowledge about how the GDPR criteria have been applied in various organizations can be gathered from previous surveys.

2.1.4 Controls Proposed

The author of this thesis gathered some controls and best practices and applied the ones that made sense to the scope of the case study, these controls are summarized here (Reini & Pasi ,March 2019):

- i. Personal data mapping: Data mapping assists a company in determining what PIIS it has, how these PIIS are transferred, and where they are stored, and who has access to them.
- ii. Who has access to the data, review the access management rules?

- iii. Consent: Data controllers must prove that consent was given by data subjects therefore they must keep track of how and when an individual consented to the processing of his or her data..
- iv. Lawfulness, fairness, and transparency, for this control in the case study they reviewed and updated the Users terms and conditions of usage to include a report providing transparency on the registry owner's contact information, the purpose for personal data processing, the origin of personal data, the legal grounds for the processing, and the type of data that will be collected.
- v. Data minimization and Storage limitations: here the company checked for the gaps in the legacy system for these 2 principles.
- vi. Software's GDPR features here the author grouped a set of technical controls that support a better compliant software:
 - a. Change, access, and error logs: here the author explained the importance of a good logs and errors management system and offered a checklist on how to implement such a system.
 - b. Do not archive unnecessary data.
 - c. Clean databases automatically: the company built a service to remove obsolete data records in the database containing personal data after there is no need to store them anymore to comply with the storage limitation principle.
 - d. Person search: the company built a tool to isolate and check the personal data related to one data subject for the customers directly without having to open a request to the helpdesk.
 - e. Automatic file removal service: the company built a service to remove obsolete files containing personal data after there is no need to store them anymore to comply with the storage limitation principle.
 - f. The use of HTTPS over HTTP.
 - g. Right to data portability
 - h. Review of existing features of the application (refactoring of the features).

- vii. Notification of data breaches: here the author addressed some concerns about how to notify the data subject and authorities and the importance of having a policy for incident response
- viii. How to ensure GDPR compliance in the future: the author offered some recommendations such as auditing the system and software, complying with training the employees regularly, monitoring the new GDPR requirements, and the Use of privacy by design as an approach for future developments.

Comments

These controls are useful because they were issued from a real case study and not solely based on the theory that proves that these controls are to some degree feasible. since in our thesis we would be trying to have a more generic framework we would be taking some of these controls to inspire the creation of our controls set we would be formulating them to a more generic control (see chapter 3) for example instead of saying “the use of HTTPS over HTTP” as a control since not all applications are web applications the control can be rephrased to the use of proven solid protocols that supports encryption of personal data while been transferred and we can extend this encryption to data at rest in some case so a generic control that might come up from this constatation is the use of encryption as a principle when possible for PIIS.

2.2 OWASP ASVS CHECKLIST OVERVIEW

The OWASP ASVS stands for OWASP Application Security Verification Standard standards published by the OWASP organization. OWASP Explains that:” The ASVS checklist is a community-driven effort to establish a framework of security requirements and controls that focus on defining the functional and non-functional security controls required when designing, developing and testing modern web applications and web services”. In the scope of this thesis research, we would be going through this standard to extract the controls that are significantly important to GDPR compliance. (3)

OWASP ASVS checklist has a set of controls under the section V8.3 of the documentation dedicated to sensitive personal data protection controls we would focus mainly on this chapter to extract some controls to answer our first research question in

(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf, [Accessed 20 January 2021].

the analysis part, however, we would be going through the whole document to take any other control that may have an impact on GDPR compliance, references to the documentation would be added to the respected controls in chapter 3 of Analysis .

2.2.1 General Data Protection Controls in ASVS

Ensuring General Data protection is handled is a broad topic that can vary a lot based on the software application type we want to develop. The development environment and the company's or organization's infrastructure.

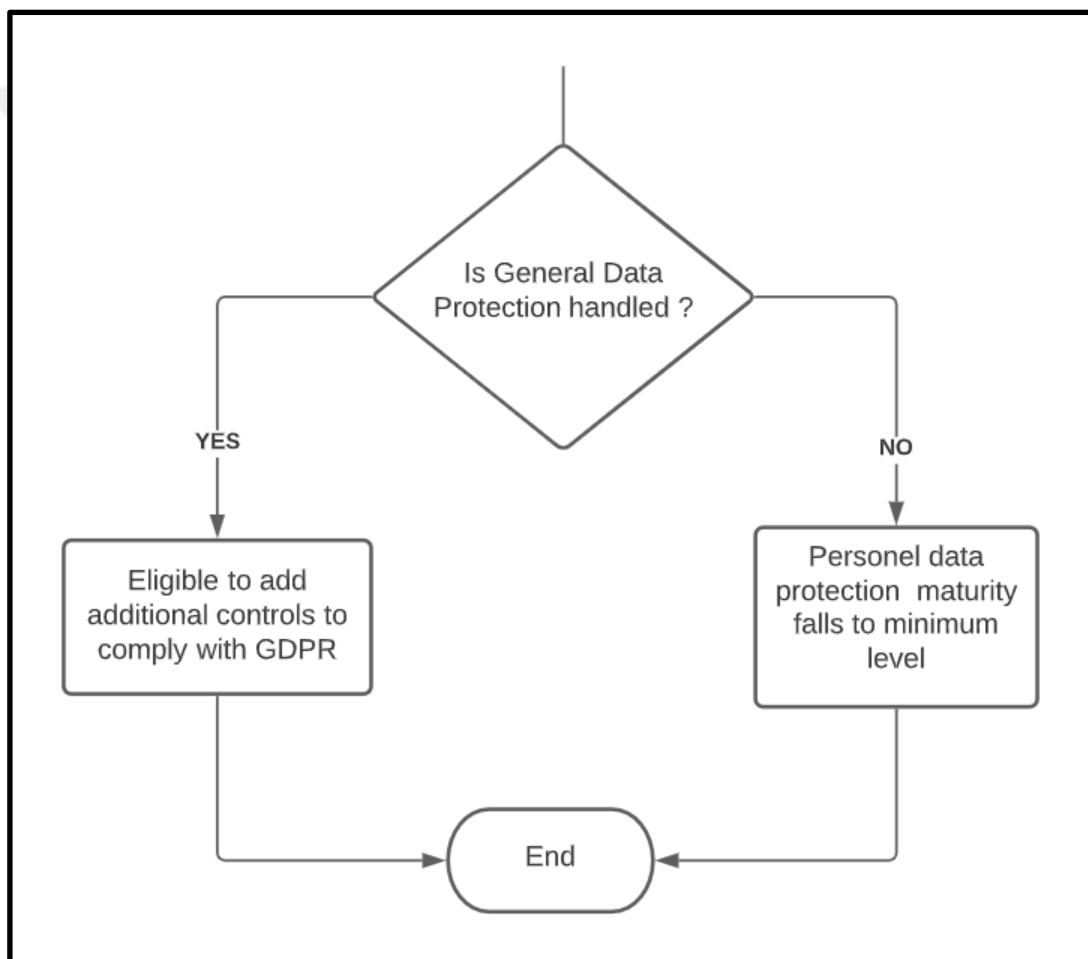
Many standards exist to ensure application security and General data protection in an organization: OWASP and NIST are examples of organizations providing a set of standards that are serving that purpose. However, a question that might come to mind is "Why is General data protection Important to mention and how does it impact GDPR compliance?" "The basic reason is that having general data protection implemented, sets a healthy foundation for Personal data protection. Therefore, we want to emphasize that personal data protection is an additional security layer that cannot be achieved if General data security is not achieved there is no such thing as simply protecting personal data without implementing any general data protection controls. Therefore, before moving forward with the security controls that are directly related to personal data, we want to give some general guidelines for general data protection that we found interesting and well explained in the ASVS documentation. General data protection is detailed in chapter V8 and V4 of the ASVS documentation some of these controls are as follows (3):

- i. Verify that the principle of least privilege is implemented, and end-users should have access to only the resources they possess a specific need for (Resource's count: functions, data files, URLs, controllers, services. Etc.).
- ii. Implement the principle of denying by default when a new user of the system is created, meaning a new user starts with minimal or no access/rights unless explicitly granted.
- iii. Verify that Access control Fails securely including when an exception occurs.
- iv. Verify administrative interfaces use appropriate multi-factor authentication to prevent unauthorized use.

(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf, [Accessed 20 January 2021].

- v. Verify that the application makes regular backups of important data and that a restoration test from backups is performed.
- vi. Verify that backups are stored safely (using encryption for example) to prevent the backups from being corrupted or stolen.
- vii. Verify the application or system can detect and alert abnormal frequency of requests a measure could be by reporting abnormal requests from an IP address per hour or per day or by user ID per hour or per day.(3)

Figure 2.1: General data protection Impact on personal data protection



(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf, [Accessed 20 January 2021].

2.2.2 Personal Data Protection Controls in ASVS:

The V8.3 paragraph of the ASVS checklist addresses a set of controls for web applications to ensure the protection of Sensitive Private Data. These controls are the following (3):

- i. verify those query parameters of HTTP are sensitive data free and that sensitive data is sent to the server in the HTTP message body or headers.
- ii. Make sure users can delete or export their data on demand.
- iii. Check that users are given explicit guidance about how their personal information will be collected and used and that they have given their consent to the use of that data before it's being used in any way.
- iv. Verify that all sensitive data generated and processed by the application has been established and that a policy on how to handle sensitive data has been implemented.
- v. If the data is collected under applicable data protection directives or where access logging is necessary, verify that accessing sensitive data is audited (without logging the sensitive data itself).
- vi. To prevent memory dumping attacks, make sure the sensitive information in memory is overwritten with zeros or random data as soon as it is no longer needed.
- vii. Check that any sensitive or private data that must be encrypted is encrypted using certified algorithms that provide both confidentiality and integrity.
- viii. Check that sensitive personal data is subject to data retention classification, such that old or out-of-date data is removed automatically, on a schedule, or as required.

2.3 OWASP PRIVACY USER CHEAT SHEET OVERVIEW

OWASP has released a project.(4) Where public documentation is available named the OWASP Cheat Sheet Series. To provide Organizations with guidelines to implement web applications security gather all the best practices, among these, they released a Cheat sheet solely dedicated to user's privacy.

(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf , [Accessed 20 January 2021].

(4): OWASP, Cheat-Sheets project ,2021, <https://cheatsheetseries.owasp.org/> [Accessed February 2021].

This OWASP cheat sheet provides security techniques that web developers may use to defend their users from a wide range of threats and aggressions that could compromise their privacy and anonymity. This cheat sheet focuses on the privacy and anonymity risks that user can encounter when using online services.

2.3.1 Strong Cryptography

The OWASP privacy cheat sheet states that strong cryptography must be used to encrypt any online platform that manages user identities, private information, or communications. Speaking of cryptography both encryption and hashing would come to mind as controls. For Encryption: In transit and at rest, user messages must be encrypted.

Developers need to use and conform to TLS/SSL guidelines such as verified certificates, properly secure private keys, use of robust ciphers exclusively informative, and consistent user alerts, and adequate key lengths to protect data in transit. Furthermore, private data must be encrypted at rest using keys of appropriate length. For hashing: To greatly minimize the risks of exposed credentials as well as proper integrity control and regardless of whether user credentials are encrypted in storage, they must be hashed by using powerful, collision-resistant hashing algorithms.

2.3.2 Support HTTP Strict Transport Security

HTTP Strict Transport Security (HSTS) is an HTTP header set by the server that instructs the user agent (Browser) to accept only protected (HTTPS) connections, obligating the user agent (Browser) to convert all vulnerable HTTP links to HTTPS and enforcing the user agent to fail-safely by rejecting any TLS/SSL connection that the client does not trust. If enforcing HSTS on all users is impractical, web developers should at the very least give users the option to activate it if they want to (4).

.

(4): OWASP, Cheat-Sheets project ,2021, <https://cheatsheetseries.owasp.org/> [Accessed February 2021].

2.3.3 Digital Certificate Pinning

Hardcoding or storing a predefined collection of information that usually hashes, for digital certificates/public keys in the user agent (browser, mobile application, or browser plugin) is referred to as certificate pinning. The following are some of the benefits of pinning. (4)

- i. When the target audience is unlikely to comprehend certificate alerts and is more likely to accept an invalid certificate.
- ii. In the case where a compromised CA trusted by a user will issue certificates for any domain, allowing malicious actors to eavesdrop on the user's communication.
- iii. In circumstances where users must accept a potentially malicious root CA, such as corporate environments or national PKI schemes.

2.3.4 Panic Modes

When threatened users are under direct risk of disclosing their credentials, here is where a panic mode feature can be helpful. Giving users the ability to activate panic mode can help them survive these threats, particularly in risky areas of the world. Unfortunately, many users all over the world are vulnerable to threats that most organizations are unaware of or fail to consider. Based on the software, a specific panic mode may be used that would differ based on the context

For example, be Email providers may enforce a panic mode that masks predefined sensitive emails or contacts, permitting only reading of innocent email messages, (sensitive content: emails and contacts as specified by the user).

Another example could be in case of disk encryption software If the user has entered their distressed password, the disk encryption software can execute a panic mode that launches up a fake device partition. It is primarily important to notice that a panic mode should be as real as possible and hardly recognized for the safety of the user, for instance, emails and notifications should be triggered normally.

Implementing a panic mode should always be done to confuse attackers and prevent them from accessing the victim's actual accounts/sensitive data, as well as preventing the detection of any existing panic modes for that user. (5) (4)

2.3.5 Remote Session Invalidation

If the user equipment is missing, stolen, or seized, or if cookie theft is suspected, it can be very helpful for users to be able to access their current online sessions and deactivate any doubtful ongoing sessions, especially those associated with stolen or confiscated devices. (4)

2.3.6 Prevent IP Address Leakage

When user security is at stake, preventing the leakage of user IP addresses is critical. Any application that hosts external third-party content, such as animations, images, or file attachments, should consider the advantages of allowing the user to prevent third-party content from loading on the application page. Many web applications need useful content to function, which is perfectly appropriate as a business process; however, web developers should consider allowing users to restrict external content as a safeguard therefore it is about giving the end-user the choice. A very similar problem occurs in HTML-formatted emails that include third-party images; nevertheless, most email clients and providers block third-party content loading by default, providing users with greater privacy. (4)

2.3.7 Honesty and Transparency

If the web application is unable to provide the user with adequate legal or political protection, or if the web application is unable to prevent the misuse or exposure of sensitive information such as logs, the truth must be communicated to users clearly and understandably so that users can make an informed decision about whether to use this website.

(4): OWASP, Cheat-Sheets project ,2021, <https://cheatsheetseries.owasp.org/> [Accessed February 2021].,

(5): VeraCrypt , veraCrypt documentation
, 2013, <https://www.veracrypt.fr/en/Documentation.html> [Accessed February 2021].

Honesty goes a long way toward promoting a culture of trust between a web service and its customers, and it allows users all over the world to think carefully about their decisions, preventing harm to users in a variety of ways (4).

Comments: OWASP is an organization that focuses on sharing the best regularly security practices for web- development, in our selected controls for thesis work we won't be going through the minor details as we are trying to address the issue on a larger scale, we would re-use the controls in a way that is scalable for different software development types.

2.4 BS 10012:2017

The BS10012:2017 is a British standard that aims to enable organizations to adopt a personal information management system (PIMS) as part of a larger information governance system that provides a framework for maintaining and improving data security standards and best practices. This British Standard Directions to businesses of all sizes and sectors. It is designed for those in charge of preparing, developing, implementing, and sustaining a PIMS within a company. Although it is not directed to software development some of the controls are also applicable to software development as it is one of the processes that an organization has so we would be trying to focus on taking the controls that are relevant to that purpose. In chapter 8 the controls to implement the framework are detailed, some of these controls are:

2.4.1 Identifying and Recording Uses of Personal Information

Goal: Ensure that the company is aware of the different types of personal information it collects, and the level of risk associated with that processing .

Based on the Paragraph 8.2.2 of BS 10012 framework this could be achieved by firstly keeping up an Inventory The company must keep track of the different types of personal information it processes. This inventory should also provide a list of the purposes for which each type of personal data is used. The company must keep track of where personal information flows in its processes .(2)

(2) *British Standards Institution*, BS 10012 Personal Information Management System, 2017, <https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/> [Accessed 01 December 2020] , pp1-42.

The inventory should allow for the explicit detection of the company's high-risk categories of personal data that might include as described in the standard the following categories:

- i. Special category personal information according to paragraph 3.1.30 of the standard (2):
 - a. Origins in terms of race or ethnicity.
 - b. Political perspectives.
 - c. Beliefs in religion or philosophy
 - d. Membership in a labor union.
 - e. Genetic data processing.
 - f. Biometric data to uniquely identify a human being.
 - g. Health-related facts or information about a natural person's sex life or sexual orientation.
- ii. Personal financial records, such as a bank account.
- iii. Identifiers unique to a country, such as national insurance numbers.
- iv. Personal details related to vulnerable children and adults.
- v. Natural person profiles (including children).

2.4.2 Risk Assessment and Treatment

Goal: Verify that the company is aware of any risks associated with the processing of specific types of personal data.

The organization will be able to determine the level of risk to natural persons involved in the processing of their personal information by adopting the PIA. Other organizations' (3rd parties) processing should be included in such evaluations

According to paragraph 8.2.4 ,to minimize the probability of noncompliance with the PIMS policy, the company must adopt a risk treatment strategy to mitigate any threats found by the risk assessment. (2)

2.4.3 Training and Awareness

Goal: is to ensure that employees are aware of their obligations when handling personal data.

According to Paragraph 8.2.2, the Organization must ensure that the workers in charge of the day-to-day operations can demonstrate compliance to security standards and best practices. Employees should be able to demonstrate expertise in their knowledge of data security standards and best practices, as well as how they can be applied within the organization. Employees should be kept informed on issues relating to the handling of personal data, as necessary, by communication with external bodies. Workers must receive training to process personal information in compliance with the guidelines in place. This training must be personalized to each employee's specific position within the organization. Staff should be reminded of the importance of following all relevant data security policies. (2)

2.4.4 Keeping PIMS Up-to-Date

Goal: This assessment aims to see if the PIMS continues to provide an infrastructure for sustaining and enhancing compliance with data security standards and best practices.

According to Paragraph 8.2.5, the workers with day-to-day responsibility for PIMS policy implementation shall decide on a regular basis whether the PIMS enables and will continue to monitor compliance with data security requirements and good practice and make any necessary adjustments. (2)

2.4.5 Fair, Lawful and Transparent Processing

Goal: To ensure that personal information is handled fairly, lawfully, and transparently and that the legitimate reasons for processing are established before processing begins . (2)

(2) British Standards Institution, BS 10012 Personal Information Management System, 2017, <https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/> [Accessed 01 December 2020], pp1-42.

2.4.5.1 Records of privacy information

The PIMS is responsible for keeping track of all privacy information given to individuals (such as privacy notices and online privacy statements). These documents must be kept for at least if the personal information they refer to is kept. If a specific privacy notice (or variant of a privacy notice) was in place, information about it must be kept.

2.4.5.2 Accessibility of privacy information

The PIMS shall ensure that any knowledge provided to natural persons is presented in a manner that enables the intended audience to easily access and understand it.

2.4.5.3 Collection from third parties

When personal data is acquired from third parties, the PIMS must ensure that it is collected equally and lawfully. If personal information is obtained from third parties, the PIMS must ensure that the identified natural persons are provided with information within one month of the collection, unless the natural person already has the information, and/or providing it would be disproportionately difficult.⁽²⁾

2.4.6 Processing for Specific Legitimate Purposes

Goal: The aim is to ensure that personal information is only collected for one or more specific reasons and that it is not further handled in any way that is inconsistent with those goals.

The PIMS must ensure that personal information is only collected for one or more specific purposes and that it is not further manipulated in any way that is conflicting with those purposes. The PIMS must ensure that personal information obtained for one purpose will not be used for another inappropriate purpose unless the following conditions are met:

- i. There is a relevant exception from the law.

⁽²⁾ *British Standards Institution*, BS 10012 Personal Information Management System, 2017, <https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/> [Accessed 01 December 2020], pp1-42.

- ii. The natural persons whose personal data is to be processed for the new purpose had given their consent to that processing.



Any processing must be in keeping with the original aim. If personal information is used for a purpose other than that for which it was originally collected, the new user must be reasonable and not unexpected.

According to Paragraph 8.2.7 of BS 10012 , any consent for incompatible processing must be freely given and consulted, according to the PIMS. If new consent documents are collected for a different processing purpose, they must be maintained. (2)

According to Paragraph 8.2.7.4 of BS 10012, when an entity exchanges personal information with another organization, the PIMS must ensure that all parties' obligations for the personal information are formally documented in a written contract, as necessary .(2)

2.4.7 Data Minimization Principles

Goal: The goal is to ensure that personal data is appropriate, significant, and not unjustified or excessive.

The PIMS must ensure that the personal data obtained by the organization is appropriate for the organization.

The PIMS shall ensure that technology and procedures concerning the processing of personal information are reviewed on a regular contractual basis for example to ensure that the personal information continues to be appropriate for such purposes.

The PIMS must make certain of that:

- i. The company only processes the bare minimum of personal data to achieve its legitimate objectives.
- ii. Any additional personal information that is not necessary or excessive for the specified purposes is not processed unless it is voluntary and only processed with the natural person's consent.
- iii. New systems and procedures involving the processing of personal data are investigated to ensure that the data being processed is both appropriate and not excessive for the processing goal.

2.4.8 Accuracy

Goal: The goal is to ensure that personal data is correct and up to date, as required.

The PIMS is responsible for maintaining the integrity and accuracy of personal data that is processed. The PIMS guarantees that natural persons can question the accuracy of their details and, where possible, have it corrected. If this process determines that the alleged inaccuracy is false and the data is correct, PIMS will keep the relevant evidence. According to the PIMS, employees must be educated on the importance of properly recording personal information and using only up-to-date personal information when making critical decisions about natural persons. (2)

When inaccurate information has been shared with 3rd parties it is the organization responsibility to:

- i. Warn any third party with whom the agency has exchanged inaccurate or out-of-date personal information that the data is inaccurate and/or out-of-date and should not be used to make decisions about the natural persons involved.
- ii. Where necessary, inform the third party of any changes to personal information.

2.4.9 Retention and Disposal

Goal: The aim is to make sure that personal information isn't held any longer than it needs to be.

To identify decisions on personal information retention times, the PIMS will refer to retention schedules. These schedules would include the following:

- i. Include any legal minimum retention periods as well as any retention periods established by the organization.
- ii. Make the reason and basis for the retention periods transparent and recorded.

(2) *British Standards Institution*, BS 10012 Personal Information Management System, 2017, <https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/> [Accessed 01 December 2020], pp1-42.

By referring to disposal procedures that are handled by the PIMS at the expiration of the retention period, all records of personal information that are no further needed by the company will be disposed of employing proven processes with a degree of protection proportional with the sensitive nature of the personal data and in accordance with the organization's risk assessment for information security.

2.4.10 Security Issues

Goal: Using appropriate technological and organizational measures and controls, ensuring that personal information is secured from unauthorized or unlawful processing, as well as external loss, degradation, or harm.

The PIMS must identify effective security measures based on the current technology, implementation costs, and the type, design, meaning, and purposes of personal data processing based on unique the system needs.

Some of these measures that are mentioned in the standard in paragraph 8.2.11.1 are :

- i. Personal information is pseudonymized and/or encrypted.
- ii. Ability to maintain system confidentiality, integrity, availability, and durability on an ongoing basis.
- iii. The capacity to recover personal information in the event of a physical or technological incident.
- iv. The capacity for testing and assessment of the efficacy of the security measures that have been established.

Other controls the standards mentioned from paragraph 8.2.11.3 till paragraph 8.2.11.6 are:

- i. Storage and handling :The PIMS must store and manage personal information safely, with precautions proportional to the degree of confidentiality and sensitivity related to the data category. Cautious care must be taken for third-party storage services like cloud storage service to make sure they comply with the requirements.
- ii. Transfer :When personal information is transmitted electronically or manually within the organization or to other organizations, the PIMS must ensure that the information is protected during transmission using the appropriate means specified by the organization. If electronic transfers are made, encryption should be used.
- iii. Access controls :Where employees have access to personal information, the PIMS must ensure that this access is limited to those workers who need it as part of their job. Employees must recognize that if access is legally granted, it is only for Job purposes and that information would only be used/processed for legitimate purposes to the data collection. The PIMS must ensure that access to personal information is controlled and evaluated in accordance with the organization's risk assessment for information security. When high-risk personal information is processed, the PIMS must ensure that access controls are sufficient for the information's sensitivity.
- iv. Security assessments :The PIMS will ensure that security assessments are carried out regularly. These evaluations will determine if current security measures are sufficient and, if so, make suggestions for improvements. In the case a security breach occurs, these assessments must consider the chances of damage, loss, and/or inconvenience to natural persons.
- v. Managing security breaches: in this section the standard presents a set of guidelines to ensure effective incident response management as follows:
 - a) Assess, manage, and record personal information security breaches, providing methods for mitigating the harm caused by any security incident.

- b) Any breaches that constitute a violation that is likely to cause damage to the rights of any natural persons must be reported to the supervisory authority (with the necessary information within 72 hours of being aware of the incident). The following information must be included in such notifications :
 - i. A summary of the personal data that is being affected by the incident.
 - ii. information on the types of personal data and the estimated number of records involved.
 - iii. contact information for the DPO or another point of contact within the organization.
 - iv. an overview of the breach's potential repercussions.
 - v. a summary of the actions taken (or proposed) to mitigate the violation and minimize any potential negative consequences.
- c) Where the breach is likely to pose a high risk to natural persons' rights and freedoms, inform the involved natural persons as soon as possible about the security breach and its nature, as well as share advice for mitigating risks.
- d) Each security breach should be documented, with an assessment of how it happened, what steps were taken to fix it, and what lessons can be learned from it.
- e) Act on whether a security incident should be reported to any relevant regulators and keep track of any alerts sent out.

Comments:

We cannot cover all the standard here rather we chose to cover some parts that are relevant to our research, although this standard is building a PIMS (personal Information management system) for an organization if developing software is concerned as part of the organization processes that it would highly make sense that some controls are to be applied for the software development scope.

2.5 ISO27701 STANDARD OVERVIEW

The ISO27701 standard is an enhancement of the ISO 27001 standard for information security, but it adds more specific privacy controls. (6)

This standard aims to offer organizations a functional structure for extending their existing ISMS (Information Security Management System) into a PIMS (Privacy Information Management System).

At the end of the ISO27701 standard, a list of annexes is present to provide more clarity of this privacy information management system and how it would map to other regulations and standards.

Annex D is dedicated to mapping the ISO27701 controls with the related GDPR articles from article 5 till article 49 except for article 43. This mapping list will also serve as one of our resources to fetch some useful controls to serve our purpose.

Unfortunately, due to the copyright regulations regarding of the ISO27701 standards, we cannot be sharing this mapped list in our thesis document report , as this standard is not freely available but companies have to pay to get their copy.

2.6 SOFTWARE DEVELOPMENT OVERVIEW

2.6.1 Software Development Life Cycle (SDLC)

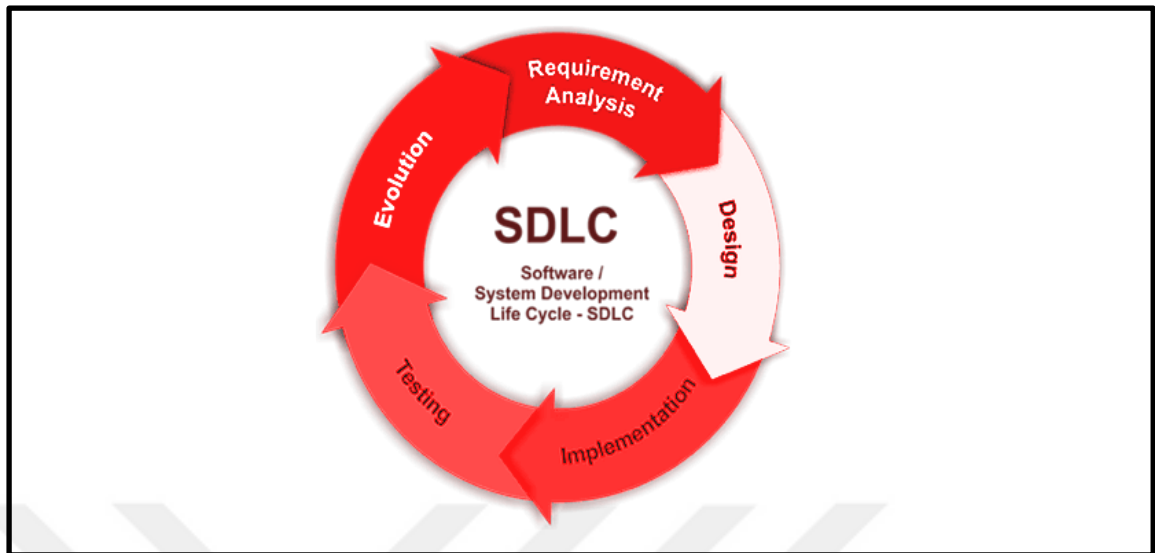
When we look at the Software development lifecycle, there are many ways that the life cycle is implemented or maintained. Many variations exist based on the needs and requirements and the nature of the software product we are building. (7)

But we generally should have 5 different phases regardless of the methodology of the implementation (see figure 2.2)

(6) :*International Organization for Standards*, ISO/IEC 27701:2019 Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines, 2019, <https://www.iso.org/standard/71670.html> [Accessed 20 December 2020] , pp.1-66.

(7): Scrum , Scrum Standard Guideline: "The Definitive Guide to Scrum The Rules of the Game ",2017, scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf , [March 2020.] ,pp.3-18.

Figure 2.2: An abstract SDLC view



First, we have a requirement Analysis phase. During that phase, we would be gathering up all the requirements from the clients or relevant stakeholders on what they wish to build as a solution to their problems throughout this product.

The following phase is the design phase where we would be making some architectural decisions on the methodologies and technologies that would be used in developing the software the client wishes for.

As of the implementation phase that is where the actual code writing and development happens. The following step is conducting a test to the written code to make sure that the code is replying to the functional needs of the client, this is also the phase where we are conducting some security testing usually.

Then the Evolution phase comes once the code is pushed to the production environment, we will have some monitoring going on to prevent defects and vulnerabilities, Any Issues that may occur are addressed in the future Iterations any new client or business requirements also would be Identified and pushed to the next Iterations of the life cycle. Security controls must be implemented concerning the different phases to reduce the risks of vulnerabilities and mitigate the risks. These controls usually are addressing several layers of Application security, in the scope of this thesis we want to understand

how these controls should be handled so that the produced software product is GDPR compliant



2.6.1.1 What are agile methodologies?

The main principle of Agile development is Iterations. In each Iteration, there is a set of actions that are repeated till the goal of the iteration is reached. Different Agile methodologies implement iterations differently in Scrum methodology for example we use the terminology of sprints, a sprint is a fixed time interval generally a week or two durations. (7)

During a sprint, the project team develops a part of the functionalities in the scope of the project. In a regular sprint Usually, the product owner checks the product backlog items (backlog items are tasks that the team needs to achieve to finish the product they are developing) then prioritize which Items would be handled in the new sprint.

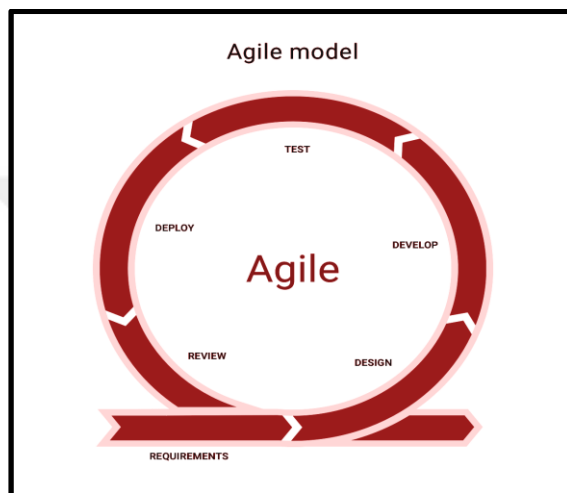
Therefor the Product Owner determines the scope and goals of the sprint, in collaboration with the development team and the product owner they create the core-pendant user stories. different ceremonies are played during the stream to communicate context and check progress: daily scrum stand-ups, retrospective feedback meetings, and huddles are examples. one of the main advantages of this methodology is that the revision and the prioritizing of the backlog items mean that the team could start delivering functional parts of the software at an early stage that can be showcased to the stakeholders and the customer. (7)

In the scope of this thesis, we are interested to know how we can incorporate GDPR compliance as part of the tasks played along with an agile development methodology. For that need first to give some clarity on what are the main phases of an agile approach in further chapters we would distribute our controls on each of the phases they should be handled in.

(7): Scrum , Scrum Standard Guideline: "The Definitive Guide to Scrum The Rules of the Game ",2017, scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf , [March 2020.] ,pp.3-18.

In the scope of this thesis, we are interested to know how we can incorporate GDPR compliance as part of the tasks played along with an agile development methodology. For that need first to give some clarity on what are the main phases of an agile approach in further chapters we would distribute our controls on each of the phases they should be handled in.

Figure 2.1: An abstract Agile model view



2.6.1.1.1 Phase number one: requirements

The Product owner needs to create initial documentation for the initial requirements before being able to proceed with the development or the design. The initial documentation should provide clarity on the result the project is willing to achieve (the value to the customer). (7)

The Scope of the project and the features to be addressed are also to be defined and the features that won't be part of the scope also are to be documented. Generally, in agile development, we can use the terminology MVP (Minimum viable product) as a principle in that case in each iteration of the project we are defining the only necessary crucial features in scope to that iteration other features are pushed as part of future iterations.

2.6.1.1.2 Phase number two: design

For the design phase, 2 types of design are generally to be addressed

(7): Scrum , Scrum Standard Guideline: "The Definitive Guide to Scrum The Rules of the Game ",2017, scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf , [March 2020.] ,pp.3-18.

2.6.1.1.2.1 Software design

In this phase the PO introduces the requirements to the development team then discussions happen regarding the implementation of these requirements like development tools, programming languages, Environment Issues, Frameworks involved, and libraries etc. and all technical needs that the project would require.

2.6.1.1.2.2 Ui/ux design

Here the Design team creates a rough mock-up based on the requirements. If the product is going to be customer-grade, then the User interface and User Experience are Crucial to the project, in the further iterations refining is done to the initial design to align with the updated needs.

2.6.1.1.3 Phase number three: development and coding

In this phase, we are doing the actual coding to respond to our design and requirements and convert it to actual working software.

2.6.1.1.4 Phase number four: integration and testing

In this stage, we try to assure the quality insurance of the code by checking if the code is bug-free and is compatible with previous code written also that the code respects the acceptance criteria from the requirements.

During future iterations, the testing becomes much more involved not just from a functional point of view but also regarding the system integration, user acceptance testing, and hardening.

2.6.1.1.5 Phase number five: implementation and deployment

At this stage the application is pushed to production and is made available to the customers, future iterations are used to make updates or launch new releases.

2.6.1.1.6 Phase six: review

Once all previous phases are completed the product owner along with the development team collaborates once again and reviews the work done in comparison with the requirements. After that, the software development life cycle starts a new iteration



3. ANALYSIS

3.1 PERSONAL DATA PROTECTION CONTROLS THROUGHOUT SOFTWARE DEVELOPEMENT

In the previous chapters we have explained the different phases an agile software development methodology follows, we also discussed the state of the art and set the ground for the analysis phase. In this chapter we would give Specific controls for each development phase based on controls and best practices we found in the literature review, then we would analyze and group these controls under the umbrella of the principles they contribute Finally, we would propose a framework to assess the implementation level of the principle concerning the assessed development Iteration.

3.1.1 Controls Related To The Requirement And The Design Phases

3.1.1.1 (C 1.1) Providing training and guidelines to the development team

3.1.1.1.1 The goal of the security control

Mitigate security risks and compliance risks on personal data by having a trained development team.

3.1.1.1.2 Explanation of the security control

One of the controls that may get neglected is the importance of having an aware development team that would be involved in the development of the product regarding to the personal data protection requirements making training to the development team

In paragraph 1.1 of this thesis regarding the motivation of this study we talked about a survey that was conducted among developers and one of the main concerns they had was the lack of training and guidelines to implement GDPR as developers. therefore, this control is important, furthermore, this best practice was specified as a control in mostly all the literature we reviewed (Reini & Pasi ,March 2019) .

3.1.1.1.3 Implementation of the security control

The manner this control is implemented varies depending on the organization and development team. Some businesses can undergo face-to-face or online training. Others may hold technical meetings "on the best practices for GDPR compliance" regularly to share information among the development team.

3.1.1.2 (C1.2) Assess GDPR compliance gap on the legacy system

3.1.1.2.1 The goal of the security control

Assess the gap in the legacy system's GDPR implementation (previous releases) so that it can be mitigated or patched in the new release.

3.1.1.2.2 Explanation of the security control

If the application for which we want to provide development has a legacy system that has been in place for a long time, perhaps even before GDPR was adopted. There's a good chance it'll fall short of regulatory enforcement, chances are it would have a gap in compliance with the regulation.

Software development projects are tightened up to a definite scope, deadlines, business needs therefore they cannot cover refactoring all the features of a legacy system of an application, distributing such ownership is a management task among the different projects.

Still doing a gap assessment to the whole legacy system if possible or to the involved old features in the scope of the new development is the necessary thing to do.

The gap feedback that would come out of this auditing should be handled during the development phase of the project by implementing the controls we would address in the paragraph (3.1.2).

3.1.1.2.3 Implementation of the security control

Assessing the current state of the GDPR implementation could be tricky but with the growing concern on the topic of privacy and personal data protection, more and more solutions are out there for organizations to use.

Organizations can choose to hire consultancy services from a security firm or conduct such an assessment.

Alternative solutions exist such as the roles that are quite usable, one of the tools we have found while doing research is a simple assessment sheet proposed by DeFRadar, a privacy training and consulting company owned by Roland Costea, a privacy Expert.

They offered an Easy assessment as you can see in the following figure (see figure 3.1) the tool goes through all the 99 Articles of GDPR and retrieve the requirements of understandable actionable questions to which the person conducting the assessment answers by yes or no, the person conducting the auditing can add actions to be met to comply and finally assign an owner.

Comment:

If multiple development projects are running in an organization Ownership of these GDPR requirements could be assigned under the scope of each project accordingly for example.

Figure 3.1: GDPR GAP Assessment tool example

DEF RADAR

GDPR Gap Assessment Tool
Note: this gap assessment must be conducted with reference to a copy of the GDPR

Security Classification: [Insert classification]
Version: 1
Dated: dd/mm/yy
Approval: [Name of approver]

Chapter	Section	Article	Paragraph and Point	Requirements	Compliant?	Action required to achieve compliance	Action owner
CHAPTER I - General provisions							
		Article 1 Subject-matter and objectives	All	None - informational only			
		Article 2 Material scope	All	Has it been established that the GDPR applies to the personal data processing activities that the organization undertakes?	Yes	✓	
		Article 3 Territorial scope	All	Has it been established that the GDPR applies, based on the data subjects whose personal data we process?	Yes	✓	
		Article 4 Definitions	All	None - informational only			
Total:					2		
CHAPTER II - Principles							
		Article 5 - Principles relating to processing of personal data	1a	Are personal data processed lawfully, fairly and transparently?	Yes	✓	
			1b	Are personal data collected for specified, explicit and legitimate purposes?	Yes	✓	
			1c	Are personal data collected adequate, relevant and limited to what is necessary?	Yes	✓	
			1d	Are personal data accurate and, where necessary, kept up to date?	Yes	✓	
			1e	Are personal data kept for no longer than is necessary?	Yes	✓	
			1f	Are personal data processed in a manner that ensures its appropriate security?	Yes	✓	
			2	As the controller, can we demonstrate compliance with all principles?	Yes	✓	
		Article 6 - Lawfulness of processing	1	Has the lawful basis for processing of all personal data been established?	Yes	✓	
			2	None - informational only			

+ ≡ Gap Assessment ▾ Compliance Results ▾ Compliance Chart ▾ % Compliance Chart ▾ % Compliance Radar ▾

Source: DEF RADAR GDPR GAP Assessment Tool.

3.1.1.3 (C 1.3) Personal data identification (mapping)

3.1.1.3.1 The goal of the security control

Understand the personal data that is currently being collected and will be processed by the time the project is completed. Dividing the data into groups and assigning each one a retention period.

3.1.1.3.2 Explanation of the security control

Personal data identification refers to identifying the personal data that we would need to collect, process, or transfer when developing the system. It is important to have this clarity as it is not possible to protect personal data if we do not know if it is there.

3.1.1.3.3 Implementation of the security control

The best practice is to create a mapping for the data involved by testing the following points:

- i. Identify all the personal data: collected, processed, or created by the application.
- ii. Sort/segment the personnel data types into categories.
- iii. For each category attribute a retention period after which this data should be deleted.
- iv. Identify for which purpose the data is collected and processed.
- v. Identify where personal data is going to be stored.
- vi. Identify the endpoints that would be transferring the Data.
- vii. If this Control is to be implemented on an application that has a legacy system, then an investigation of the previous Data on the legacy system must be conducted. If gaps are to be found these would be identified and work should be conducted to fix them.

A high-level example of what a mapping would look like:

Table 3.1: A personal data mapping example

Asset Name / storage name	Description	Purpose	Owner	Location Stored	Data Retention Period	Date of acquisition	Business Criticality	Information Classification
Customer Database	Address, name, surname	Serving the customers with the product of the company	Sales	X Server	5 years	D:M: Y	High	Confidential
Employees Database	The role, age, name surname	working contract	HR	Y Server	6 years	D:M: Y	High	Confidential

Source: DEFRA DAR GDPR Trainings

3.1.1.4 (C 1.4) Design /Re-design access management system

3.1.1.4.1 The goal of the security control

Mitigate unauthorized access and ensure data confidentiality and integrity.

3.1.1.4.2 Explanation of the security control

Based on the GDPR principle data collection and access must be restricted to the minimum needs to fulfill the collection and processing goal, under this light: employees, users, stakeholders, as well as development team members interacting with the software, should not have the same access rights to the personal data in the system whether it is while developing the system or when using it once in production. Every human interacting with the system either as an end-user or as part of the development team should have access (as personal data) to only what is strictly needed to fulfill his role.

All different Roles interacting with the system must be identified and their access level must be cleared for each data collected Also access types have to be restricted to what is only necessary as a type of access using Least privilege by default then adding the different writes to the data access: Read Only, Write and Read or write only for Example.

If we are still in the early phases of software development, we will still be developing the system's foundation, therefore access management should be provided with loads of effort to do a detailed investigation and create a stable foundation. However, if we are at a later stage of the product life cycle, such as the design phase of a new Iteration for a new update, we must ensure that the new features we are implementing do not violate our Access management policies. Furthermore, updates on the Users Access rights in the system can happen due to the introduction of the new features, so conducting a re-analysis could avoid the situation where a certain user of the system has access to personal data Assets that he should not access.

If changes need to be made, then once these new changes have been designed the policies need to be updated See control: C3.5.

3.1.1.4.3 Implementation of the security control

Many different access management methodologies exist (Charles E.P,(2004)):

- i. MAC (Mandatory access control)
- ii. DAC (Discretionary access control)
- iii. RBAC (Role based access control)
- iv. UBAC (User based access control)
- v. ABAC (Attribute based access control)

Each of these methodologies has its pros and cons, while designing a new system cautious thought and analysis must be done to choose which methodology or combination of methodologies is going to be used.

The choice must be justified based on the asset types, the level of protection, and the type of protection we want to achieve (Charles E.P,2004).

Some Real-life examples of the need of having a good access management system could be:

For many applications, there could be support Executives Role that handles Customers requests in that case they need to have access to some of the customer's data, therefore for the design phase the tasks that would be conducted by helpdesk executives should be clarified and mapped to the minimum needed information the Helpdesk Executive needs.

For example, if the Application is an E-commerce delivering products to customers the helpdesk Executive may need the ability to check and update the delivery address of the customer's order if the customer calls for such a request. but he might not need to have access to other customers' details such as the customer's gender, nationality, or Age.

3.1.1.5 (C 1.5) Sharing data with 3rd parties

3.1.1.5.1 The goal of the security control

Mitigate risk on personal data that must be shared with 3rd parties. By making the sharing of data on a legal basis.



3.1.1.5.2 Explanation of the security control

In some cases while designing the system we uncover the need to use 3rd parties to provide services (It could be for storage purposes or a payment provider.). To comply with such need's steps must be taken before sharing personal data with any 3rd party .
(2)

3.1.1.5.3 Implementation of the security control

Here's an overview of the checklist the product owner, project manager, business analyst, or whoever is the stakeholder that has the authority to facilitate this process is to check before choosing the service provider:

- a. The entity that would store the data must be in a country that has data protection laws that are just as strong as GDPR (as determined by the EU Commission).
- b. A contract should be in place between the organization and the hosting service provider company.
- c. The entity to whom the data is passed agrees by legally binding the contract to follow GDPR principles of data protection.

3.1.2 Controls Related to The Development Phase and The Testing Phase

3.1.2.1 (C 2.1) Encryption

3.1.2.1.1 The goal of the security control

Protect the confidentiality of personal data during transfer and at rest.

3.1.2.1.2 Explanation of the security control

Personal data should be protected from unauthorized access therefore Encryption is a mandatory control, personal data should be encrypted both at rest and during transit.

(2) *British Standards Institution*, BS 10012 Personal Information Management System, 2017, <https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/> [Accessed 01 December 2020] , pp1-42.



Encryption should be done by approved methods or algorithm that provides both confidentiality and integrity, security through using encryption methods that are not approved by the large researcher's public and have not proven to be efficient is called security through obscurity as the security relies on the algorithm to be unknown such practices are considered unsafe and should not be used to encrypt sensitive data such as personal data.

3.1.2.1.3 Implementation of the security control

- a) At rest: Encrypting data at rest should be done by strong encryption methods such as AES or RSA. (3)
- b) In transfer: If the software development involves calling APIs or building APIs that sends personal data then the personal data should be encrypted and transmitted in a secured manner. SSL, TLS are examples of networking protocols that are widely used to encrypt data in transmission.

3.1.2.2 (C 2.2) Data minimization by design

3.1.2.2.1 The goal of the security control

Achieve Data minimization by ensuring that personal data is relevant, acceptable, and not unnecessary or unjustified.

3.1.2.2.2 Explanation of the security control

While developing new features the development team must make sure they are not using personal data in an excessive way that is not needed for the feature, the question that needs to be asked each time is what is the bare minimum data that needs to be involved.

(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf , [Accessed 20 January 2021].



3.1.2.2.3 Implementation of the security control

An example of this control could be: the data should be minimized before being shared (through an API or different services for example) to the minimum needed data by the concerned service while developing this need must be identified and developed. As another example: For web applications, we should verify the application minimizes the number of parameters in a request (parameters counts header values, cookies, and hidden fields... etc.). The data should be minimized while being collected concerning the legal basis of the data collection and in accordance with the privacy notice and agreements.

3.1.2.3 (C 2.3) Logs management

3.1.2.3.1 The goal of the security control

Apply best practices to manage logs to ensure that personal data can be properly audited.

3.1.2.3.2 Explanation of the security control

Keeping logs is very important for auditing, incident response ... etc., However, when auditing Access to personal data we should make sure we are keeping the logs without logging the sensitive data itself.

A life cycle of log processing contains e.g., the following phases: collecting logs, analyzing logs, storing logs, giving logs to third parties, and removing them. Handling logs is one of the most important practices for maintaining information systems and monitoring their security (Reini & Pasi, March 2019).

3.1.2.3.3 Implementation of the security control

Checklist for legal log processing (page 36 of Reini & Pasi, March 2019) as follows:

- i. Identify why and what is the purpose and legal base for each log processing.
- ii. Evaluate the need for each log data entity that is being stored. Identify what personal data might be stored in logs.
- iii. Find out how each log should be protected.
- iv. Pay attention to legal aspects of log monitoring.
- v. Notify users and other stakeholders of log processing.
- vi. Pay attention to requirements on the person register that is given by authorities if the log is to be a person register.
- vii. Design and document a purpose for log storing and ensure its implementation.

3.1.2.4 (C 2.4) Clean databases automatically

3.1.2.4.1 The goal of the security control

Remove personal data records after their retention period is met to ensure the database is not storing any outdated non-compliant personal data.

3.1.2.4.2 Explanation of the security control

One of the GDPR principles is Data minimization therefore companies must get rid of any unnecessary personal data after the purpose of the data has been achieved or is no more needed.

One of the controls could be enforcing a mechanism to automatically clean the records that are obsolete or no longer needed containing personal data.

. Many tools and solutions exist for that as an example: scheduling cronjobs to delete the unnecessary personal data automatically, once the retention period of the data is over, non-complaint previous controls we discussed that retention period can be defined to each data category and therefore once this retention period is over the data can be deleted.

3.1.2.4.3 Implementation of the security control

Check that sensitive personal data is classified for data retention, so that old or out-of-date data is deleted automatically, on a schedule, or as required. Automated cron jobs on the database management system can be used to guarantee such control, one of the implementation cases is (Reini & Pasi, March 2019) that we review in paragraph 2.1.

3.1.2.5 (C 2.5) Automatic file removal service

3.1.2.5.1 The goal of the security control

Remove Files containing personal data after their retention period is met to ensure the system is not storing any outdated non-compliant personal data.

3.1.2.5.2 Explanation of the security control

The same thing as the automatic cleaning of the database, files containing personal data should also be deleted once their retention period is reached, such mechanisms can be implemented automatically while developing controls to automate such a process could be added if the software stores for its functioning goal files that have personal data.

3.1.2.5.3 Implementation of the security control

One of the implementation cases is (Reini & Pasi, March 2019) that we review in paragraph 2.1 where the company-built software to automatically delete files from the system for achieving this purpose.



3.1.2.6 (C 2.6) Over-writing data in memory/disk

3.1.2.6.1 The goal of the security control

Ensure the confidentiality of data from unauthorized access.

3.1.2.6.2 Explanation of the security control

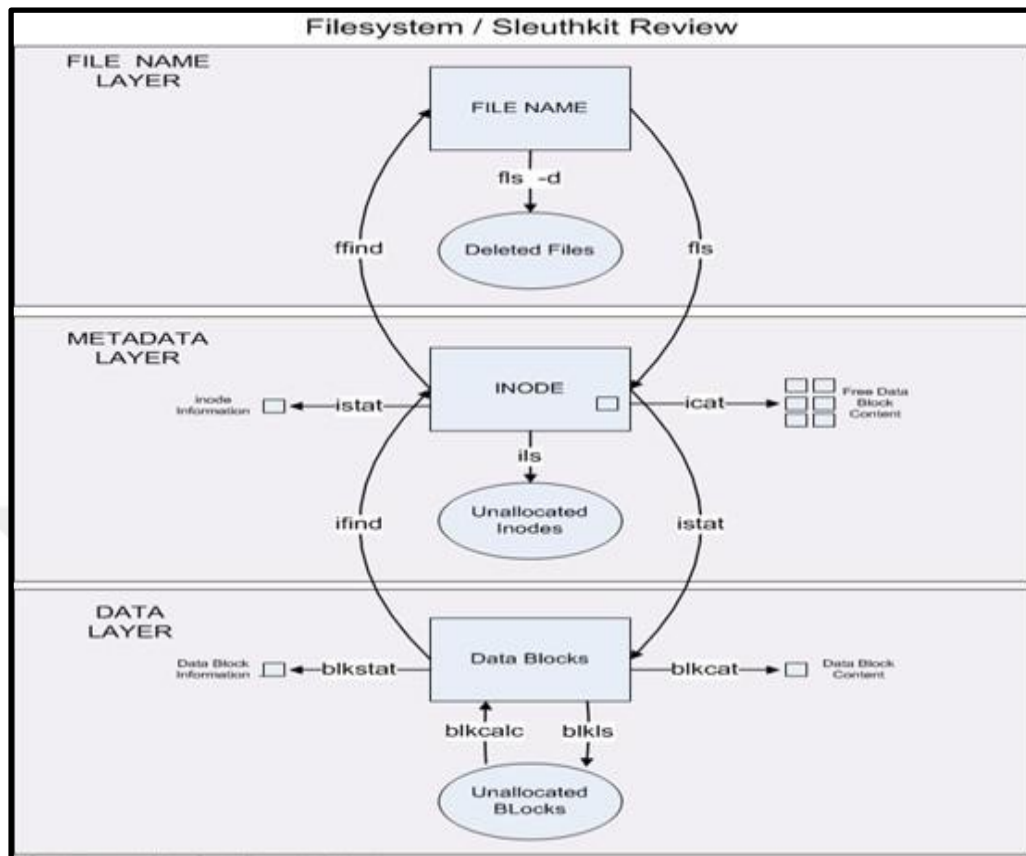
One of the principles of GDPR is to be able to delete the non-necessary personal data, unfortunately, deletion is often not conducted in an un-proper way. Data should be overwritten or wiped properly so that it cannot be retrieved.

3.1.2.6.3 Implementation of the security control

For example, while deleting files generally, companies delete only the meta-data of the file, the file content is therefore no longer accessible through that file structure, but it can be retrieved easily with proper tools as the data is still there on the disk.

One of the tools to conduct such an attack on a disk is “The sleuth kit” TSK (see figure 3.2), the figure explains the set of commands used by the tool to navigate through 3 layers filename, metadata layer, data layer based on the operating system file system structure this tool is a digital forensic tool that helps investigate in disk to retrieve data. If data is not wiped from the disk properly then such tools can be used to retrieve personal data.

Figure 3.2: Sleuth kit review



Source: Computer forensics File system “layers & the sleuthkit”.

3.1.2.7 (C 2.7) Pseudonymization by Default

3.1.2.7.1 The goal of the security control

Ensure the confidentiality of sensitive personal data.

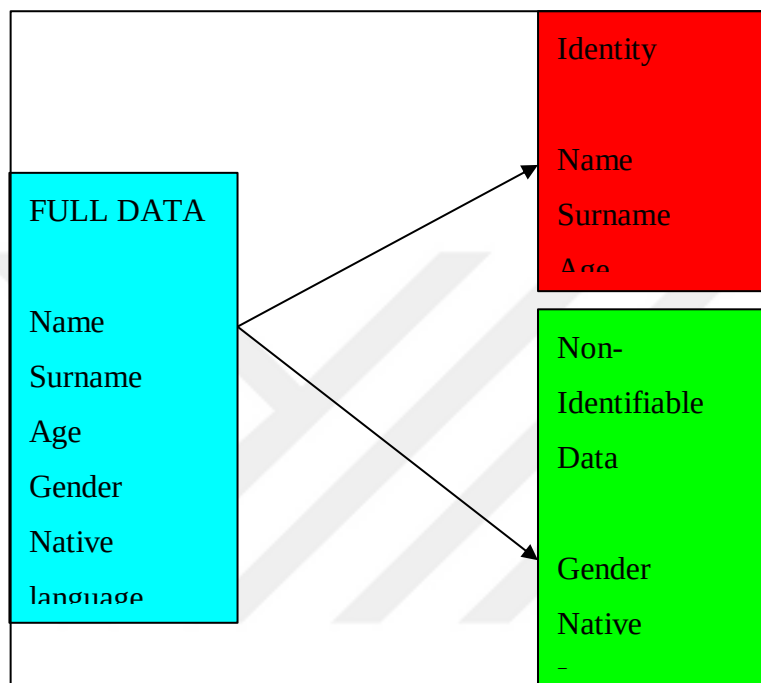
3.1.2.7.2 Explanation of the security control

Pseudonyms must be created for each individual, and data about the person’s identity should be stored in an area that is fully partitioned and separate from other user data, such as information on the individual’s account within an app or software platform .

3.1.2.7.3 Implementation of the security control

Here is an example in the figure number 3.3 on how personal data can be partitioned to achieve Pseudonymization by default.

Figure 3.3: an example of Pseudonymization by default.



3.1.2.8 (C 2.8) Develop a personal data automatic search method

3.1.2.8.1 The goal of the security control

Isolate and retrieve personal data easily to achieve transparency with data subjects.

3.1.2.8.2 Explanation of the security control

With respect to GDPR, the data subject has the right to request a report of all the data a company or an organization has on them. Furthermore, GDPR obliges companies to notify Data subjects or authorities of a data breach affecting personal data within 72hours.

This report should include a detailed explanation of the personal data that is impacted. To comply with such requirements Organization must implement a method to Isolate and retrieve all personal data they hold on to a data subject automatically.

Building such an ability while building the product during development is a feature that should be considered to make compliance easier. This control can also help the data portability mentioned on ART 20 of GDPR (1), as this Article gives the right to data subjects to retrieve their data or request a data processor to share their data with a different data processor in an understandable readable format

3.1.2.8.3 implementation of the security control

An example of such a tool could be the tool they developed in (Reini & Pasi ,March 2019) where data subjects can download directly their personal data from the system. that we discussed in paragraph 2.1 of this thesis.

3.1.2.9 (C2.9) Masking data in the test environment

3.1.2.9.1 The goal of the security control

Ensure data confidentiality of personal data.

3.1.2.9.2 Explanation of the security control

During developing software, it might be attempting to test our software with real customer data thinking our tests would be much more accurate and would provide more accurate results. This practice is not recommended especially without the consent of the data subject.

3.1.2.9.2 Implementation of the controls

An implementation for this control could be by masking the personal data or using mock data instead of real personal data to test the software needs. In case consent is not provided, personal data not only should not be used for tests but also should not be accessible in the test environment.

3.1.2.10 (C 2.10) Back up data

3.1.2.10.1 The goal of the security control

Ensure data availability of personal data.

3.1.2.10.2 Explanation of the security control

During development companies should provide a solution to back up personal data in a safe way so the data is not lost. Purchase orders, Customer contact lists, Invoices, medical records are examples of personal data that companies should keep in their systems for the system or software they are building to function, backups should be planned (6).

3.1.2.10.2 Implementation of the security control

Many solutions exist to create automatic back up, among which using cloud storage services is common however careful care should be made to choose a solution that is GDPR compliant if back-up is stored in a 3rd party storage service as specified in control: (C1.5).

.

3.1.2.11 (C 2.11) Informed consent

3.1.2.11.1 The goal of the security control

Ensure the Consent on the processing is handled legally and is up to date to the processing conducted on personal data.

3.1.2.11.2 Explanation of the security control

Users must be allowed to provide informed consent for the collection and processing of their data. Therefore, so many privacy-related disclaimer panels have popped up on websites, software platforms, and mobile apps in recent months. Another example of informed consent applies to tick boxes when registering for an account. In most cases, tick boxes should not be ticked by default; the user must click them manually. If personal data is used for a purpose other than that for which it was collected in the new development, the new use must be reasonable and not unexpected, and any contradictory processing must be explicitly communicated and consented to by the data subject (6) (3).

3.1.2.11.3 Implementation of the security control

Consent for data processing and collection is mandatory by the GDPR therefore: The Privacy Policy and Agreements should be accessible and provided in clear language regarding the collection and use of supplied personal information. Also, the consent should be Explicitly given by the User. when the application that is developed is a web application, Cookie's policy must be in scope. The cookies policy must be accessible and the person surfing the webpage must be informed and give Explicit consent to accept the policy. It is important to note that in the case of the data subject a child consent Must be Explicitly provided by the holder of parental responsibility in all cases based on the GDPR Article 8 regarding Conditions applicable to child's consent in relation to information society services.

(6): ISO27701 Standard, 2019.

(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf , [Accessed 20 January 2021].

3.1.2.12 (C 2.12) Provide privacy settings features

3.1.2.12.1 The goal of the security control

Provide more privacy solutions and alternatives to data subjects who are using the product that is being developed, so that they have more control over their privacy.

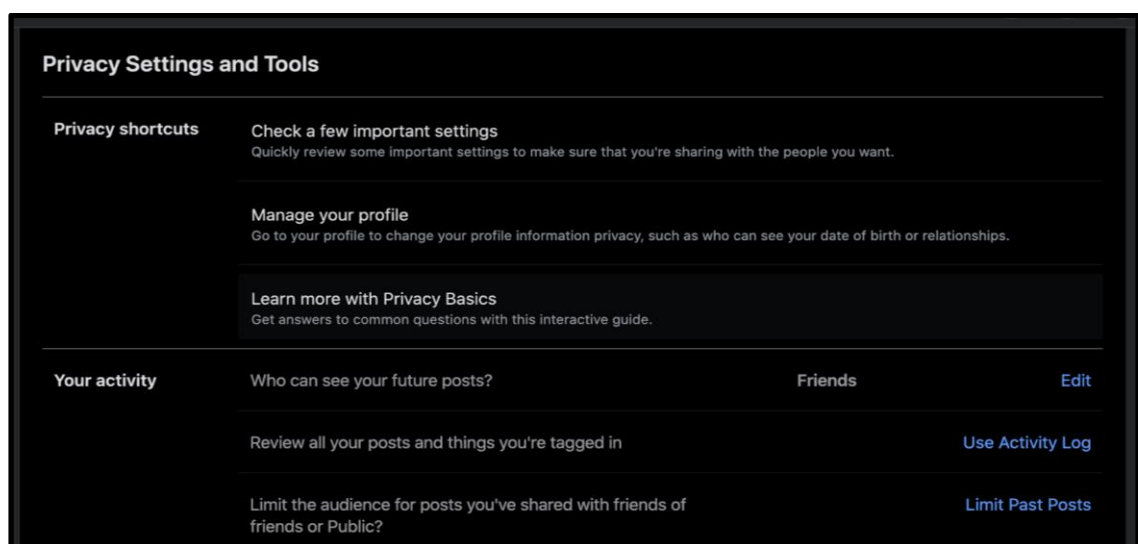
3.1.2.12.2 Explanation of the security control

Software products that manipulate personal data such as social media platforms should give a level of control over the gathering and processing of personal data to the final user.

3.1.2.12.3 Implementation of the security control

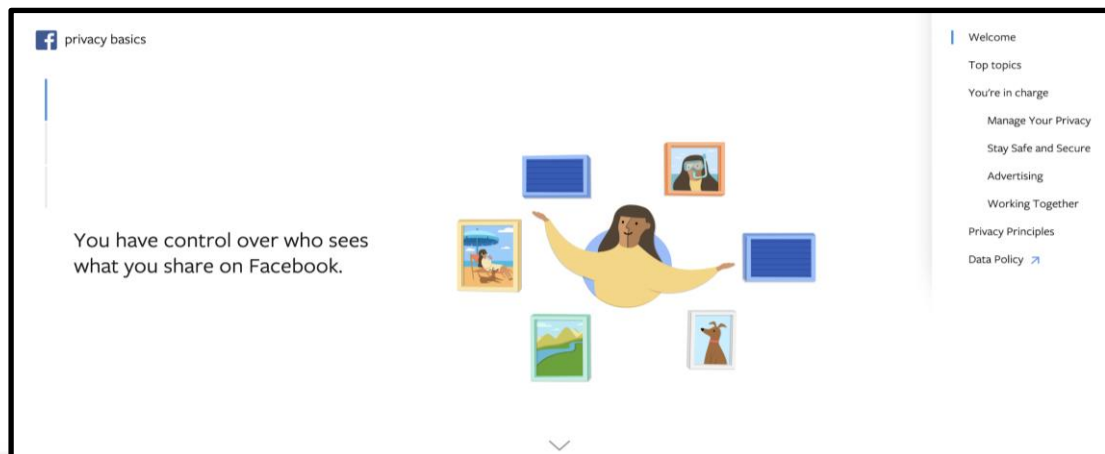
If we take Facebook, for example, Facebook has a very detailed section for privacy settings where customers can choose what is shared and not shared. Privacy settings should provide choices to the customer in an accessible manageable way (See figures 3.4 and 3.5)

Figure 3.4: Facebook 2021 privacy settings and tools



Source: facebook.com

Figure 3.5: Facebook 2021 privacy basics



Source:facebook.com

3.1.2.13 (C 2.13) Refactoring old features

3.1.2.13.1 The goal of the security control

Refactoring benefits are (Reini & Pasi ,March 2019):

- a) Getting rid of stored unneeded data that could be GDPR un-compliant.
- b) The code becomes much more well documented and readable.
- c) Better performance and fewer storage costs.
- d) Decrease attack vectors by getting rid of obsolete endpoints, databases, or code.

3.1.2.13.2 Explanation of the security control

When an application has years of history and a legacy system that is still permanently different development Teams: developers, testers, business analysts, Product Owners have worked on it and over the years some features became obsolete. In most cases, there would be GDPR un-compliant data or code in the application. It could be data that is still in the database and has not been cleaned yet for example. A technique to avoid building such systems is refactoring.

3.1.2.13.2 Implementation of the security control

Refactoring is the act of changing the code or the database content or structure without bringing change to the application behavior. The refactoring could affect the code, the databases, or the user interface (Reini & Pasi ,March 2019).

3.1.2.14 (C 2.14) Hashing

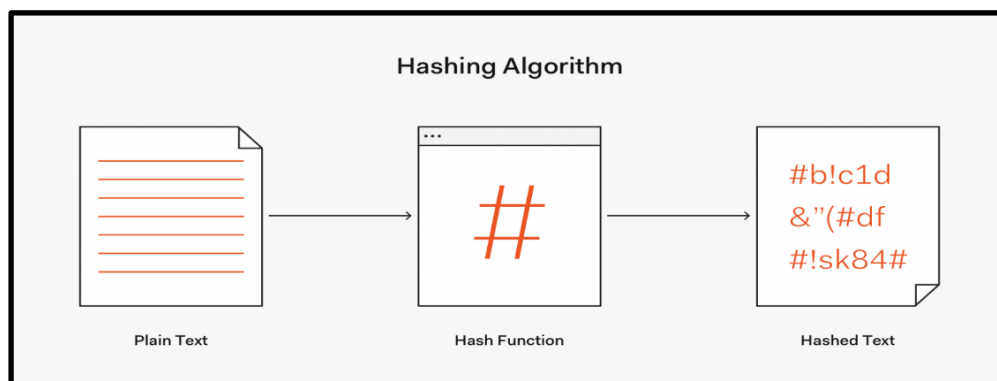
3.1.2.14.1 The goal of the security control

The goal of this control is to ensure integrity and accuracy of the data (not been altered) because the hash output text is unique therefore it is used to ensure that the original output was not altered.

3.1.2.14.2 Explanation of the security control

Hashing Algorithms are one-way functions that turn any form of data into a unique and fixed-size output hashed text.

Figure 3.6: Hash function



Source: [computersciencewiki.org](https://www.computersciencewiki.org/)

3.1.2.14.3 Implementation of the security control

Some famous hashing functions already exist (MD5, sha1, sha2... etc.). We can see multiple use cases for hashing functions such as hashing a file before transfer and after it was received to ensure the received version is Identical to the sent one by comparing the hashes, this can be useful to ensure integrity of backups Now if we think that the backup data contain personal data then we can see how this practice is important to ensure the integrity and accuracy of personal data. (3)

3.1.2.15 (C 2.15) Enforce data subject to update data upon expiry

3.1.2.15.1 The goal of the security control

Mitigate the risk of having inaccurate data, by enforcing users to update the data when possible.

3.1.2.15.2 Explanation of the security control

Accuracy of the data must be insured based on the GDPR regulation one way to do this is to set control points for personal data expiry. Some personal data has an expiry date and this can be used in the favor of the development team that can enforce the costumer to input Up-To-Date data after the old one expired.

3.1.2.15.3 Implementation of the security control

For example if software processes personal data that has a clear expiration date like the identifiers of passport numbers "as the passport has an expiry date that can be provided when collecting the passport ID", a process can be developed to ensure the customer is forced to enter the new passport details after the current one has expired. such logical control points can exist for other personal data and should be supported by the software to ensure the users are forced to update data after it expires.

(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf , [Accessed 20 January 2021].

3.1.2.16 (C 2.16) Cache controls

3.1.2.16.1 The goal of the security control

Ensure Personal Data confidentiality, by protecting it from unauthorized access.

3.1.2.16.2 Explanation of the security control

This control is applicable for web-applications that uses caches, personal data should not be cached and in case it must be cached for the application to run then proper measures to safeguard the data has to be made, we added a checklist of some points to check for this in the how to implement section of this control (3).

3.1.2.16.3 Implementation of the security control

- Verify the application is protecting sensitive data from being cached when unnecessary.
- Verify all the necessary cached data is protected from unauthorized access.
- Verify that the application from the client-side sets anti-caching headers so that the sensitive data does not get cached in the browsers. For example, a web application should verify that authenticated data is cleared from the client storage side, for example from the browser after the client or session is terminated.

3.1.3 Controls Prior to Deployment

3.1.3.1 (C3.1) Penetration testing

3.1.3.1.1 The goal of the security control

Ensure personal data security: confidentiality, integrity, availability is properly maintained.

(3):Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf , [Accessed 20 January 2021].



3.1.3.1.2 Explanation of the security control

Penetration testing could seem at first unrelated to GDPR compliance however it is clearly stated in Article 32 of GDPR: “(d) a process for regularly testing, assessing, and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing.” (1). Therefore in the scope of software development before deploying the product to production conducting pen-testing is required to lower the risk of an incident after deployment.

3.1.3.1.3 Implementation of the security control

During a pen-test, we can verify for example that sensitive data and APIs are protected against direct object attacks targeting creation, reading, updating, and deletion of records, such as creating or updating someone else's record, viewing everyone's records, or deleting all records.

If a pen-test brings to the table vulnerabilities that affect personal data, these vulnerabilities must be fixed. they should be created as elements of the backlog of the software development and prioritized to be fixed prior to deployment.

3.1.3.2 (C3.2) Personal data - flow diagram

3.1.3.2.1 The goal of the security control

Understand the personal data flow to ensure it is in compliance with the organizational policies, privacy notices, and agreements.

3.1.3.2.2 Explanation of the security control

One of the most critical points is to know where the data is and how it is flowing across the system. We cannot protect something if we are not aware of its existence or if we are not clear on where it is located.

3.1.3.2.3 Implementation of the security control

A best practice is to have a personal Data flow Diagram, Data diagrams are graphical representations that illustrate the flow of the data across the system, a DFD shows the original provider of the data: Process, internal group, external entity, a data subject, or a user of the system. It also shows which processes are processing the data, which endpoints are transferring or receiving data, where data might flow even across countries and geographical borders.

Once a DFD is complete for personnel data, it can be aligned with the organizational policies, privacy notices, to be sure it complies with the internal policies if a change needs to be made then C3.5 should be applied, or a refactoring of the noncompliant feature must be made.

3.1.3.3 (C3.3) Update/Create privacy notice

3.1.3.3.1 The goal of the security control

Ensure the privacy notice is in accordance with the new development.

3.1.3.3.2 Explanation of the security control

A privacy notice is a publicly accessible document that informs customers, authorities, lawyers, or consumers on the privacy implementation of the software.

In case we are developing a new product then the privacy notice should be written and analyzed from scratch. If on the other hand, we are bringing updates to existing software that is already in production then we should make sure that the updates are in accordance with the privacy notice, any processing must adhere to the original purpose of the data collection.

If personal information is used for a purpose other than what it was gathered for in the new development, then the new use must be fair and not unexpected and any conflicting processing must be openly communicated and contested by the data subject. Furthermore, the privacy notice must be updated accordingly before going to production.

3.1.3.3.3 Implementation of the security control

A privacy notice should contain Information about (1) :

- i. What personal information is being collected on the data subject?
- ii. What controls the data subject or User has over the collection?
- iii. With which mechanism the data is being collected?
- iv. Who has access to the data?
- v. To whom the data is shared?
- vi. For how long the data is kept before being deleted?
- vii. How can data subjects' access and export or request to have the data the organization has on them?
- viii. What is the point of contact the users may use to get in touch with the company to respond to their privacy inquiries?

3.1.3.4 (C3.4) Provide a Method for Data Requests

3.1.3.4.1 The goal of the security control

Provide a touchpoint for data subjects to exercise their rights. and to authorities to connect with in case it is needed.

3.1.3.4.2 Explanation of the security control

According to the GDPR articles 14 and 15 the data subjects have the right to reach out to the organization hosting their data for requesting the data to be deleted or downloaded or transmitted to a different processor or controller this contact should be explicitly mentioned and be able to respond on short notice to customer's requests. (1)

3.1.3.4.3 Implementation of the security control

Processors and controllers are required by the GDPR to provide a system for dealing with data access requests from subjects. They must, however, provide in their Privacy notice specific and comprehensive instructions for the process.

An Example could be directing the request to the DPO (data protection officer) or to a trained customer service center that would handle the request, Organizations can also choose to provide a request form that data subjects can fill to send their request. 3.1.4 controls post to deployment.

3.1.3.5 (C3.5) Update internal privacy policy and processes

3.1.3.5.1 The goal of the security control

This control aims to align the new development with the current processes and internal privacy policies.

3.1.3.5.2 Explanation of the security control

The development team with the help of relevant stakeholders must assess if any newly developed feature is affecting the internal processes , any mis-alignment must be reviewed and mitigated.

3.1.3.5.2 Implementation of the security control

For example, new roles in the company may need to be accessing personal data they did not need to access before due to the new development. there would be a need to update internal privacy policy or to sign a specific agreement to ensure accountability of the involved parties. whatever the effect these touchpoints must be reviewed and assessed then action Items should be taken accordingly.

3.1.3.6 (C3.6) Update/Create Cookies policy

3.1.3.6.1 The goal of the security control

In case of a web application development, this control aims to assess if the new development has caused any changes to the cookie policy that the data subjects need to be informed about.

3.1.3.6.2 Explanation of the security control

If our software is a web applications chances are the development of new features would impact the cookies behavior therefor, the cookies policy must be Up-To-Date to the new processing.

3.1.3.6.3 Implementation of the security control

The e-privacy (cookie law) (8) introduced and explains the legal and mandatory practices concerning the cookies management, this directive can be thought of as “complementing” the GDPR rather than being replaced by it. it states that:

- i. Before processing cookies on a user's device or monitoring them, cookie law requires explicit, informed consent.

- ii. Consent to cookies must be notified and based on an explicit affirmative action; these actions can include continuing to browse, clicking, scrolling the page, or some other method that allows the user to actively proceed, depending on the local authority.

3.1.4 Controls Post Deployment

3.1.4.1 (C4.1) Create/Update - incident response program

3.1.4.1.1 The goal of the security control

Mitigate risks on personal data when an incident occurs.

3.1.4.1.2 Explanation of the security control

In GDPR the article 33 clearly states that organizations must report a breach within 72 hours of time frame to discovery (1). This is a very tight time in cybersecurity. If the Organization does not have a predefined incident response plan, meeting GDPR Art33 requirements would be almost impossible. A greater issue is that the report should clearly state the impact of the breach:

- i. How many data subjects were affected?
- ii. How many personal data records are affected?
- iii. What is the nature of the breach?
- iv. What are the possible consequences?
- v. What are the measures that have been taken to mitigate the damage?

3.1.4.1.3 Implementation of the security control

When developing software, it's necessary to maximize security measures and include security breach detection and reporting methods that can send notifications to the incident response team. The detection and alert systems must be implemented based on every new development that is affecting anything on the structure of the system in production. The incident response must be reviewed based on the new features that have been developed.

3.1.4.2 (C4.2) Penetration testing

3.1.4.2.1 The goal of the security control

Ensure personal data security: confidentiality, integrity, availability is properly maintained.

3.1.4.2.2 Explanation of the security control

Pen-testing is to be executed in this phase too as pen-testing should be also conducted after the development is pushed in production, if we think about web applications for example the security issues that may arise in testing or staging environments could be very different from the ones on production, therefore, re-conducting pen-testing could be an additional layer of security.

3.1.4.2.3 Implementation of the security control

Same as discussed in the control number (C3.1) , we have to check during this control the security vulnerabilities that may arise in the production environment as mentioned in (C3.1) “.. we can verify that sensitive data and APIs are protected against direct object attacks targeting creation, reading, updating, and deletion of records, such as creating or updating someone else's record, viewing everyone's records, or deleting all records..”.

If the feedback that comes out from this control finds vulnerabilities that are a potential risk to personal data, then these vulnerabilities should be prioritized and addressed as soon as possible in the new releases.

3.1.4.3 (C4.3) Testing and Auditing

3.1.4.3.1 The goal of the security control

Ensure ongoing compliance with GDPR.

3.1.4.3.2 Explanation of the security control

Ensuring compliance with GDPR is not a once-in-a-time task but rather continuous monitoring, auditing is a mandatory control that organizations use as a tool to make sure their processes or software are running properly.

3.1.4.3.2 Implementation of the security control

It is important to audit the output of the logs, the automated scripts such as the ones discussed in controls: C2.3, C2.4, and C2.5 to make sure that they are working properly to keep GDPR compliance. Furthermore, Laws, as well as business needs, will change and it is important to navigate these changes so that the application's stays compliant with GDPR.

3.1.4.4 (C4.4) Archiving practices

3.1.4.4.1 The goal of the security control

Ensure that any archived data is going to be handled in compliance with GDPR.

3.1.4.4.2 Explanation of the security control

Personal Data should be minimized and deleted from the system once it does not serve its collection purpose.

However under some circumstances and with the data subject approval we may need to archive some personal data if such a case if we must make sure that the software can (Reini & Pasi ,March 2019):

- i. Isolate and minimize the data to what is needed to be archived.
- ii. Analyzing data is encrypted before being archived.
- iii. Ensure encryption keys are stored safely.

3.1.5 Summary of the analysis

By analyzing the literature, we were able to group 30 controls, that can be performed during software development process as a best case practices to ensure better GDPR compliance while designing , creating and developing the software product.

These controls that have been detailed in paragraph 3.1 are summarized here ss follows:

- i) Controls related to the Requirement and the Design phases:(5 controls)
 - (a) (C 1.1) providing Training and guidelines to development team
 - (b) (C1.2) Assess GDPR compliance Gap on Legacy system.
 - (c) (C 1.3) Personal data Identification (Mapping).
 - (d) (C 1.4) Design /re-Design access management System.
 - (e) (C 1.5) Sharing data with 3rd parties.

ii) Controls related to the development phase and the testing phase: (16 controls)

- (a) (C 2.1) Encryption.
- (b) (C 2.2) Data minimization by design.
- (c) (C 2.3) Logs management.
- (d) (C 2.4) Clean databases automatically.
- (e) (C 2.5) Automatic file removal service.
- (f) (C 2.6) Over-writing data in memory/Disk.
- (g) (C 2.7) Pseudonymization by Default.
- (h) (C 2.8) Develop a Personal data automatic search method.
- (i) (C 2.9) Masking Data in the testing environment.
- (j) (C 2.10) Back up data.
- (k) (C 2.11) Informed Consent.
- (l) (C 2.12) Provide privacy settings Features.
- (m)(C 2.13) Refactoring Old features.
- (n) (C 2.14) Hashing.
- (o) (C 2.15) Enforce Data subject to update Data upon Expiry.
- (p) (C 2.16) Cache controls.

iii) Controls prior to deployment (5 controls).

- (a) (C3.1) Penetration testing.
- (b) (C3.2) Personal Data - Data flow diagram.
- (c) (C3.3) Update/Create Privacy notice.
- (d) (C3.4) Provide a Method for Data Requests.
- (e) (C3.5) Update internal privacy policy and processes.

iv) Post deployment controls (4 controls).

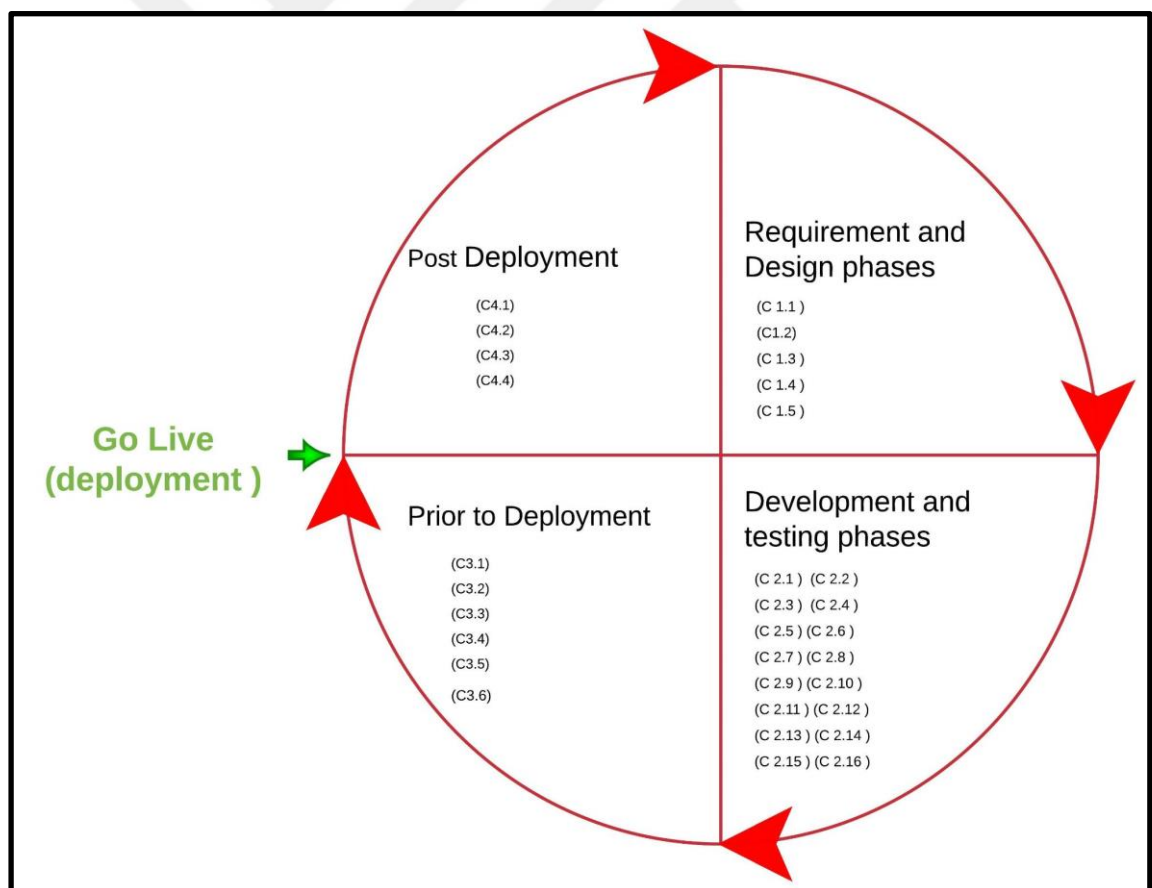
- (a) (C4.1) Create / Update - Incident response program.

- (b) (C4.2) Penetration testing.
- (c) (C4.3) Testing and Auditing.
- (d) (C4.4) Archiving practices.

The additional thing we added here is we have proposed a timeline on when the control can be played during the phases of the software development. See the following figure (see figure 3.7).

As a recommendation, these are controlled to be checked based on the needs of every new release. some controls may need to be built from scratch, but others need to be just double-checked if the new development is still valid for them.

Figure 3.7: View controls throughout the development phases



3.2 ANALYSIS OF THE CONTROLS

3.2.1 Principle's definition

To answer the second research question, we would focus on the GDPR 6 principles as the high-level principles to which we would be mapping our controls and to which we would, later, try to measure the success level of their implementation.

In future work, the 8 data subject rights can be addressed also using the same approach, but it won't be our focus for this thesis.

It is interesting to note though that the 6 principles of the GDPR and the 8 rights of the data subject are interrelated and contribute to each other.

By referring to paragraph (2.1.2) of this thesis we can see that the GDP principles are:

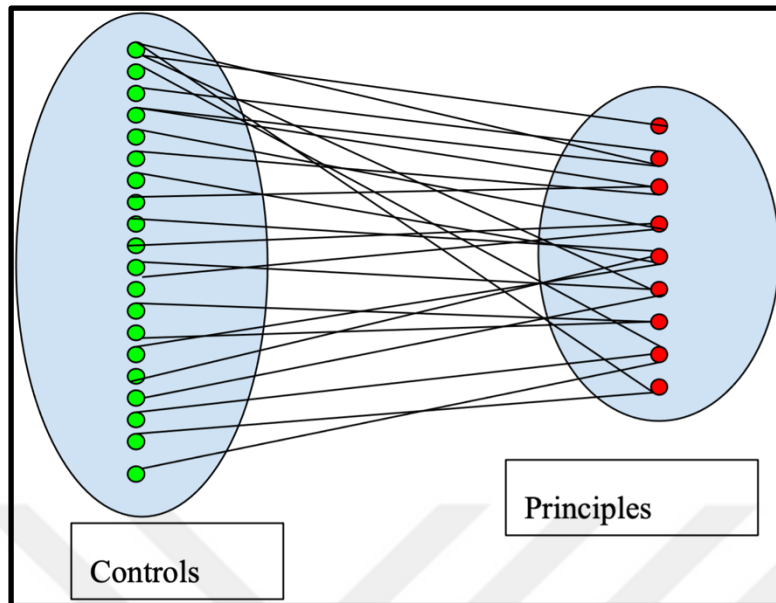
- i. lawfulness, fairness, and transparency.
- ii. purpose limitation.
- iii. Data minimization.
- iv. Accuracy.
- v. Storage limitation.
- vi. Integrity and confidentiality.

3.2.2 Controls Mapping

Before we move forward with mapping the controls to the respective principles they contribute to, It is important to understand that each control could contribute to multiple principles at the same time.

However, each control must contribute to at least one principle, If we picture this as an application, we may present this in such a Diagram (see figure 3.8):

Figure 3.8: View mapping logic



- Each Control has at least one or multiple principles to which it contributes too.
- Each principle has at least one control that contributes to it.

3.2.2.1 Fairness, lawfulness, and transparency

Table 3.2: Principle 1 controls

	P1 lawfulness, fairness, and transparency.
(C 1.1) providing Training and guidelines to the development team.	x
(C1.2) Assess GDPR compliance Gap on the Legacy system.	x
(C 1.4) Design /re-Design access management System.	x
(C 1.5) Sharing data with 3rd parties.	x
(C 2.3) Logs management.	x
(C 2.4) Clean databases automatically.	x
(C 2.5) Automatic file removal service.	x
(C 2.6) Over-writing data in memory/Disk.	x
(C 2.8) Develop a Personal data automatic search method.	x

(C 2.9) Masking Data in the testing environment.	x
(C 2.11) Informed Consent.	x
(C 2.12) Provide privacy settings Features.	x
(C 2.13) Refactoring Old features.	x
(C3.2) Personal Data - Data flow diagram.	x
(C3.3) Update/Create Privacy notice.	x
(C3.4) Provide a Method for Data Requests.	x
(C3.5) Update internal privacy policy and processes.	x
(C3.6) Update/Create Cookies policy.	x
(C4.1) Create / Update - Incident response program.	x
(C4.3) Testing and Auditing.	x
(C4.4) Archiving practices.	x

3.2.2.2 Purpose limitation

Table 3.3: Principle 2 controls

	P2 purpose limitation.
(C 1.1) providing Training and guidelines to the development team.	x
(C1.2) Assess GDPR compliance Gap on the Legacy system.	x
(C 1.3) Personal data Identification (Mapping).	x
(C 1.4) Design /re-Design access management System.	x
(C 1.5) Sharing data with 3rd parties.	x
(C 2.1) Encryption.	x
(C 2.7) Pseudonymization by Default.	x
(C 2.9) Masking Data in the testing environment.	x
(C 2.11) Informed Consent.	x
(C 2.13) Refactoring Old features.	x
(C3.2) Personal Data - Data flow diagram.	x
(C3.5) Update internal privacy policy and processes.	x
(C4.3) Testing and Auditing.	x

3.2.2.3 Data minimization

Table 3.4: Principle 3 controls

	P3 Data minimization.
(C 1.1) providing Training and guidelines to the development team.	x
(C1.2) Assess GDPR compliance Gap on the Legacy system.	x
(C 1.3) Personal data Identification (Mapping).	x
(C 1.5) Sharing data with 3rd parties.	x
(C 2.2) Data minimization by design.	x
(C 2.7) Pseudonymization by Default.	x
(C 2.11) Informed Consent.	x
(C 2.13) Refactoring Old features.	x
(C3.2) Personal Data - Data flow diagram.	x
(C3.5) Update internal privacy policy and processes.	x
(C4.3) Testing and Auditing.	x
(C4.4) Archiving practices.	x

3.2.2.4 Accuracy

Table 3.5: Principle 4 controls

	P4 Accuracy.
(C 1.1) providing Training and guidelines to the development team.	x
(C1.2) Assess GDPR compliance Gap on the Legacy system.	x
(C 1.4) Design /re-Design access management System.	x
(C 2.3) Logs management.	x
(C 2.7) Pseudonymization by Default.	x
(C 2.10) Back up data.	x
(C 2.13) Refactoring Old features.	x
(C 2.14) Hashing.	x
(C 2.15) Enforce Data subject to update Data upon Expiry.	x
(C3.1) Penetration testing.	x
(C3.5) Update internal privacy policy and processes.	x

(C4.1) Create / Update - Incident response program.	x
(C4.2) Penetration testing (C4.2).	x
(C4.3) Testing and Auditing.	x
(C4.4) Archiving practices.	x

3.2.2.5 Storage limitation

Table 3.6: Principle 5 controls

	P5 Storage limitation.
(C 1.1) providing Training and guidelines to the development team.	x
(C1.2) Assess GDPR compliance Gap on the Legacy system.	x
(C 1.3) Personal data Identification (Mapping).	x
(C 2.4) Clean databases automatically.	x
(C 2.5) Automatic file removal service.	x
(C 2.6) Over-writing data in memory/Disk.	x
(C 2.13) Refactoring Old features.	x
(C3.5) Update internal privacy policy and processes.	x
(C4.3) Testing and Auditing.	x

3.2.2.6 Integrity and confidentiality

Table 3.7: Principle 6 controls

	P6 Integrity and confidentiality.
(C 1.1) providing Training and guidelines to the development team.	x
(C1.2) Assess GDPR compliance Gap on the legacy system.	x
(C 1.3) Personal data Identification (Mapping).	x
(C 1.4) Design /re-Design access management System.	x
(C 2.1) Encryption.	x
(C 2.3) Logs management.	x

(C 2.7) Pseudonymization by Default.	x
(C 2.10) Back up data.	x
(C 2.13) Refactoring Old features.	x
(C 2.14) Hashing.	x
(C 2.16) Cache controls.	x
(C3.1) Penetration testing.	x
(C3.5) Update internal privacy policy and processes.	x
(C4.1) Create / Update - Incident response program.	x
(C4.2) Penetration testing (C4.2).	x
(C4.3) Testing and Auditing.	x
(C4.4) Archiving practices.	x

3.2.3 Adding labels

Another point to highlight is that the level of contribution is different from one control to another concerning the same principle, for this we wanted to keep our framework as simple and usable as possible therefore we would make labels for controls as follows:

- a) *Must have (MH)*: We would label this as MH must have means that if the control is not implemented this will cause a violation of the related principle.
- b) *Good to have (GH)*: We would label this as GH This means that this control is more of an enhancement that ensures a better implementation.

Note: for good to have label, the control is considered as an additional layer of safety, however, if this control is not met then the principle is still able to meet the minimum needed maturity level to be compliant to GDPR and won't necessarily directly lead to a violation. Using the above-mentioned labels, here is the output:

Table 3.8: Must have / good have controls

	label	Explanation
(C1.1) providing Training and guidelines to the	GH	Training can ensure a better implementation for all the principles but if training is not ensured there is no direct violation of the GDPR mentioned as a sanction.

development team.		
(C1.2) Assess GDPR compliance Gap on the Legacy system.	MH	Companies and organizations are sanctioned if they fail to show the GDPR compliance effort they made in case of an incident, if no assessment is done and documented organization will risk sanctions.
(C1.3) Personal data Identification (Mapping).	MH	It is necessary to have clarity on the personal data and where they are stored, which categories are collected, and what retention period is attributed to them, If Organizations fail to comply with this then it can lead to violations when responding to Data subject requests or authorities.
(C1.4) Design /re-Design access management System.	MH	If personal data is subjected to unauthorized access that is unlawful and does not respect the data collection's purpose, this is considered a violation. therefore, this control is a must-have.
(C1.5) Sharing data with 3rd parties.	MH	Data must be shared with 3rd parties lawfully otherwise it would lead to a violation.
(C2.1) Encryption.	GH	While the GDPR does not require organizations to encrypt data, it does require them to implement security measures and safeguards. Encryption and pseudonymization are frequently listed in the GDPR as “adequate technological and organizational measures” for personal data protection. therefore, encryption would be marked as good to have.
(C2.2) Data minimization by design.	MH	Minimization is one of the GDPR principles therefore implementing this control is a must-have.
(C2.3) Logs management.	MH	Data must be logged properly in case of incident response companies must provide a report, if logs are not there it would be hard to comply with incident response requirements or to perform auditing which also leads to a violation.

(C 2.4) Clean databases automatically.	GH	The need for this control can be handled manually therefore the automatic removal is good to have practice
(C2.5) Automatic file removal service.	GH	The need for this control can be handled manually therefore the automatic removal is good to have practice
(C2.6) Over-writing data in memory/Disk.	GH	
(C2.7) Pseudonymization by Default.	GH	Pseudonymization is frequently listed in the GDPR as “adequate technological and organizational measures” for personal data protection. but there is no direct enforcement by the regulation or direct violations related to the non-implementation, therefore we would mark this as good to have.
(C2.8) Develop a Personal data automatic search method.	GH	
(C 2.9) Masking Data in the testing environment.	MH	If personal data is used in tests without data subject consent this is a direct violation, therefore Personal data must be masked in test environments.
(C 2.10) Back up data.	MH	
(C2.11) Informed Consent.	MH	If processing happens without data subject consent, and processing does not fall under (exceptional permitted cases based on GDPR) then this processing is unlawful and would be considered as a violation, therefore informed consent must be handled in case of releasing new futures.



(C2.12) Provide privacy settings Features.	GH	
(C2.13) Refactoring Old features.	GH	
(C2.14) Hashing.	GH	
(C2.15) Enforce Data subject to update Data upon Expiry.	GH	A company can choose to handle this need manually or differently rather than handling this on the software part.
(C2.16) Cache controls.	GH	
(C3.1) Penetration testing.	GH	
(C3.2) Personal Data - Data flow diagram.	GH	
(C3.3) Update/Create Privacy notice.	MH	The privacy notice must be up to date with the current processing practices that the software wants to deliver, otherwise, it may lead to a violation in case of a complaint.
(C3.4) Provide a Method for Data Requests.	MH	This control is mostly related to the first principle and Data subject rights like performing the right to data portability, the right to be forgotten, the right to transparency... etc. Therefore not implementing this control is a direct violation of multiple GDPR requirements
(C3.5) Update internal privacy policy and	MH	If the privacy policy and internal agreements are not up to date based on the development that has been conducted, a huge risk of having inconsistencies, therefore, a risk of committing a

processes.		violation.
(C3.6) Update/Create Cookies policy.	MH	Same as the Privacy notice the cookies policy for web applications must be up to date with the current processing practices that the software wants to deliver, otherwise it may lead to a violation in case of a complaint.
(C4.1) Create / Update - Incident response program.	MH	GDPR has strict requirements for incident response therefore this control is a must to avoid violations.
(C4.2) Penetration testing .	GH	
(C4.3) Testing and Auditing.	MH	Companies and organizations are sanctioned if they fail to show the GDPR compliance effort, if no assessment is done and documented organization will risk sanctions.
(C4.4) Archiving practices.	MH	Archiving personal data has strict rules in case archiving is eligible and needed therefore this control is mandatory to avoid violations

Finally, here is a summary:

Table 3.9: Must have / good have Controls Mapping to principles

	P1 lawfulness , fairness and transparency.	P2 purpose limitation .	P3 Data minimisation.	P4 Accuracy.	P5 Storage limitation.	P6 Integrity and confidentiality.	label
(C 1.1) providing Training and guidelines to development team.	x	x	x	x	x	x	GH
(C1.2) Assess GDPR compliance Gap on Legacy system.	x	x	x	x	x	x	MH
(C 1.3) Personal data Identification (Mapping) .		x	x		x	x	MH
(C 1.4) Design /re-Design access management System .	x	x		x		x	MH

(C 1.5) Sharing data with 3rd parties .	x	x	x				MH
(C 2.1) Encryption .		x				x	GH
(C 2.2) Data minimization by design .			x				MH
(C 2.3) Logs management .	x			x		x	MH
(C 2.4) Clean databases automatically .	x				x		GH
(C 2.5) Automatic file removal service .	x				x		GH
(C 2.6) Over-writing data in memory/Disk.	x				x		GH
(C 2.7) Pseudonymization by Default .		x	x	x		x	GH
(C 2.8) Develop a Personal data automatic search method .	x						GH
(C 2.9) Masking Data in the testing environment .	x	x					MH
(C 2.10) Back up data .				x		x	MH
(C 2.11) Informed Consent.	x	x	x				MH
(C 2.12) Provide privacy settings Features .	x						GH
(C 2.13) Refactoring Old features .	x	x	x	x	x	x	GH
(C 2.14) Hashing .				x		x	GH
(C 2.15) Enforce Data subject to update Data upon Expiry .				x			GH
(C 2.16) Cache controls .						x	GH
(C3.1) Penetration testing.				x		x	GH
(C3.2) Personal Data - Data flow diagram .	x	x	x				GH
(C3.3) Update/Create Privacy notice .	x						MH
(C3.4) Provide a Method for Data Requests .	x						MH
(C3.5) Update	x	x	x	x	x	x	MH

internal privacy policy and processes .							
(C3.6) Update/Create Cookies policy.	x						MH
(C4.1) Create / Update - Incident response program .	x			x		x	MH
(C4.2) Penetration testing				x		x	GH
(C4.3) Testing and Auditing .	x	x	x	x	x	x	MH
(C4.4) Archiving practices .	x		x	x		x	MH



4. FINDINGS

4.1 DEFINING THE ASSESMENT MODEL FRAMEWORK

Now that we have all our Principles Identified and Mapped to respected controls (see paragraph 3.1 and 3.2), we are finally able to Propose our Maturity Framework concerning our controls Labeling as “GH” good to have and “MH” must-have, The Maturity level assessment we are offering an answer to research question 3, Goes like this in 7 levels from 0 to 6:

- i. level 0: The principle is not defined or not implemented at all.
- ii. Level 1: Only Good to have controls are implemented (either partially or fully), no must-have controls were implemented.
- iii. Level 2: The “Must have labeled controls “were implemented partially. good to have controls that were not implemented at all.
- iv. Level 3: The “Must have labeled controls “were implemented partially. good to have controls were implemented either partially or fully.
- v. Level 4: The “Must have labeled controls “were implemented Fully, good to have controls that were not implemented.
- vi. Level 5: The “Must Have labeled controls” were fully implemented, Good to have controls were partially implemented.
- vii. Level 6: The “Must Have labeled controls” were fully implemented, Good to have controls were Fully implemented.

4.2 CASE STUDY

We will assess a case study on the first release of a software project built by an E-commerce platform business that operates in the retail sector, but we will keep the company's identity secret for the sake of confidentiality.

4.2.1 Company's Presentation

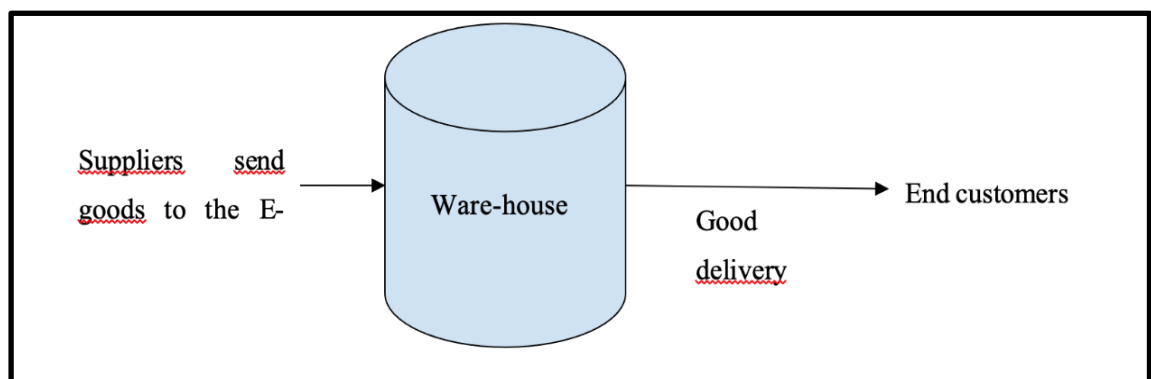
The company is a fashion E-commerce platform that operates in the retail sector in multiple countries. They have multiple suppliers of the products they sell on their website and mobile app, and they provide various distribution models and services to ensure the goods reach their customers.

The company is well-established, with 501-1 000 employees. It operates on three platforms: an Android mobile application, an iOS mobile application, and a website. Consumers should be able to surf the E-commerce platform, search for goods, and update their accounts on all three platforms.

4.2.2 Presentation of The Case Study Project

By encouraging manufacturers to send their goods directly to consumers, the company hopes to become a marketplace and to reduce the cost of receiving the goods to their warehouse and then resending to the end customer in the current fulfillment models. To achieve this end, they launched a software marketplace project to enhance their application and website to support the fulfillment option.

Figure 4.1: Existing legacy model of case study



4.2.2.1 Features to be developed in this project

The aim of this project is for the business to be able to deliver products directly from suppliers to end customers by integrating with suppliers and related 3rd parties such as distribution companies and payment providers through APIs, reducing warehouse costs and allowing customers to receive their goods more quickly from the suppliers.

4.2.2.2 Development methodology explained

The company's technology team uses an Agile development approach as follows:

To begin, an Inception is held with various stakeholders and representatives from all development positions. The aim is to identify the needs and analyze the touchpoints. As a result, the team created high-level User story narratives and a roadmap.

Iterations: Multiple Iteration happens where the Product owners would be analyzing the different user stories and detailing them and prioritizing them along with the different product backlog Items, the developers would be working on these tasks simultaneously and Quality assurance would be testing each Item the developers finish to ensure its quality.

Other positions involved in the production include the UI/UX specialist, who will be responsible for planning the designs and ensuring that the stories created by the developers and detailed by the product owners are consistent with a good user experience.

DevOps engineers, with the help of various members of the development team, attempt to make the necessary environment and structure infrastructures preparations for the project to run on test environments as well as to go live on production.

Other parties are involved like the security Team to conduct regular pentest and the legal team to confirm the legality of the features developed ..etc.

Impact of the project on the personal data

The legacy system of the running website and application of the online store already: collects, stores, and processes customer data to fulfill the services they provide. the additional change point that his product would be adding is sharing this data with:

- a. Suppliers that would integrate to the marketplace model to fulfill the order: create invoices to the customers for example.
- b. Delivery companies: some suppliers would use integrated delivery companies with the E-commerce platform to send their orders to the clients directly therefore personal information like the address, the name, the surname of the clients must be shared through the integration with these companies.

4.2.2.3 Assessment

For our case, we would be checking the GDPR maturity they have achieved on their first release of this marketplace project.

During our assessment we assessed the controls as either: Implemented, not implemented and not in scope, a control that is labeled as not in scope means that it does not correspond to the reality of the project for example some controls are related to specific platforms or special situations that might not apply to the context of the development therefor the control can be skipped.

While conducting a study by interviewing some of the development team members (A product owner and A devops-engineer volunteered to help us to conduct the assessment on the project) we achieved the following assessment on the first release of the project:

Table 4.1: Case study Assessment

Control	State	Comments	P 1	P 2	P 3	P 4	P 5	P 6	Labels
(C1.1) providing Training and guidelines to the development team.	Not Implemented	There was no special training for GDPR compliance offered by the company or a set of guidelines shared with the development team.	x	x	x	x	x	x	GH

(C1.2) Assess GDPR compliance Gap on the Legacy system.	Not Implemente d	No special assessment was conducted on the touchpoint the project would have in the legacy system and if any gaps are there that could be addressed.	x	x	x	x	x	x	MH
(C 1.3) Personal data Identificatio n (Mapping).	Not Implemente d	A personal data mapping was not prepared		x	x		x	x	MH
(C 1.4) Design /re- Design access managemen t System.	Not Implemente d	This has been handled from a functional perspective and in an organic way, the personal data protection was not part of the design focus therefore we chose to label this control as not implemented.	x	x		x		x	MH
(C 1.5) Sharing data with 3rd parties.	Implemente d	Our E-commerce company made sure to make the necessary legal agreements and mandatory checks with the supplier and cargo companies prior to sharing the personal data and starting the integration process	x	x	x				MH

(C 2.1) Encryption.	Implemented	The personal data that is transmitted through Internal and external APIs was properly encrypted.		x				x	GH
(C 2.2) Data minimization by design.	Not Implemented	This have been marked as not implemented because some roles in the company have access to real customer's personal data. A product owner stated been able to check personal data that does not late to their job directly, A product owner may be interested to see the age of the clients to discuss maybe some business needs with stakeholders and create the personas for the clients, but information like name, surname. etc. should not be visible to theProduct owner.			x				MH
(C 2.3) Logs	Implemented		x			x		x	MH

managem nt.									
(C 2.4) Clean databases automatic ly.	Not Implemented	This process is handled now manually	x				x		GH
(C 2.5) Automatic file removal service.	Not Implemented	This process is handled now manually	x				x		GH
(C 2.6) Over- writing data in memory/Di sk.	Not in scope	This control does not apply to the touchpoints of this project.	x				x		GH
(C 2.7) Pseudonym ization by Default.	Not Implemented			x	x	x		x	GH
(C 2.8) Develop a Personal data automatic search method.	Not in scope	This control does not apply to the touchpoints of this project.	x						GH

(C 2.9) Masking Data in the testing environmen t.	Not Implemented	As mentioned on the minimization by design control some roles have access to irrelevant personal data in test environments that belong to data subjects. Even though this data has not been used in the tests without the customer's consent, if such access is available, we cannot entrust this control to the ethics of the employees, this should be enforced by not providing such access in the first place.	x	x					MH
(C 2.10) Back up data.	Implemented					x		x	MH
(C 2.11) Informed Consent.	Implemented		x	x	x				MH
(C 2.12) Provide privacy settings Features.	Not in scope	This control does not apply to the touchpoints of this project.	x						GH
(C 2.13) Refactoring Old features.	Not Implemented	As there was no proper gap assessment done for the personal data on legacy systems, the	x	x	x	x	x	x	GH

		refactoring needs were not clear, we are marking this control as not implemented.							
(C 2.14) Hashing.	Not Implemented					x		x	GH
(C 2.15) Enforce Data subject to update Data upon Expiry.	Not in scope	This control does not apply to the touchpoints of this project.				x			GH
(C 2.16) Cache controls.	Implemented							x	GH
(C3.1) Penetration testing.	Implemented	2 Rounds of pen-testing were conducted prior to launching the new development to the production.				x		x	GH
(C3.2) Personal Data - Data flow diagram.	Not Implemented		x	x	x				GH
(C3.3) Update/Create Privacy notice.	Implemented		x						MH
(C3.4) Provide a Method for Data Requests.	Implemented		x						MH
(C3.5)	Implemented		x	x	x	x	x	x	MH

Update internal privacy policy and processes.									
(C3.6) Update/Create Cookies policy.	Implemented		x						MH
(C4.1) Create / Update - Incident response program.	Implemented		x			x		x	MH
(C4.2) Penetration testing.	Implemented					x		x	GH
(C4.3) Testing and Auditing.	Implemented		x	x	x	x	x	x	MH
(C4.4) Archiving practices.	Not in scope	This control does not apply to the touchpoints of this project.	x		x	x		x	MH
	Assessment out of 6		3	3	3	3	2	3	

The six principles assessment output is as follows:

- i. The maturity level for Fairness and lawfulness and transparency is 3, based on the model we introduced. The “Must have labeled controls “were implemented partially, the good to have controls were implemented either partially or fully.
- ii. The maturity level for Restriction to a specific purpose is 3, based on the model we introduced. The “Must have labeled controls “were implemented partially, the good to have controls were implemented either partially or fully

- iii. The maturity level for Data reduction and data economy is 3, based on the model we introduced. The “Must have labeled controls “were implemented partially, the good to have controls were implemented either partially or fully.
- iv. The maturity level for Accuracy is on level 3, based on the model we introduced. As the “Must have labeled controls “were implemented partially, good to have controls were implemented either partially or fully
- v. The maturity level for Storage limitation’s principle on level 2, based on the model we introduced. The “Must have labeled controls “were implemented partially and good to have controls that were not implemented.
- vi. The maturity level for the Integrity and confidentiality principle is on level 3, based on the model we introduced. The “Must have labeled controls “were implemented partially. good to have controls were implemented either partially or fully.

4.2.2.4 The output of the Analysis of the case study

If we group the previous results on a graph, we can realize that the project has a poor to average maturity level when it comes to the implementation of the GDPR principles, however, all principles are fortunately implemented to a certain level and none of the principles is completely missed or overlooked.

We can see here in figures (figure 4.2 and figure 4.3) the output of the case study assessment. A deeper explanation of this analysis is shown in the following table , number 4.2 .

Figure 4.2: Radar chart Case study assessment.

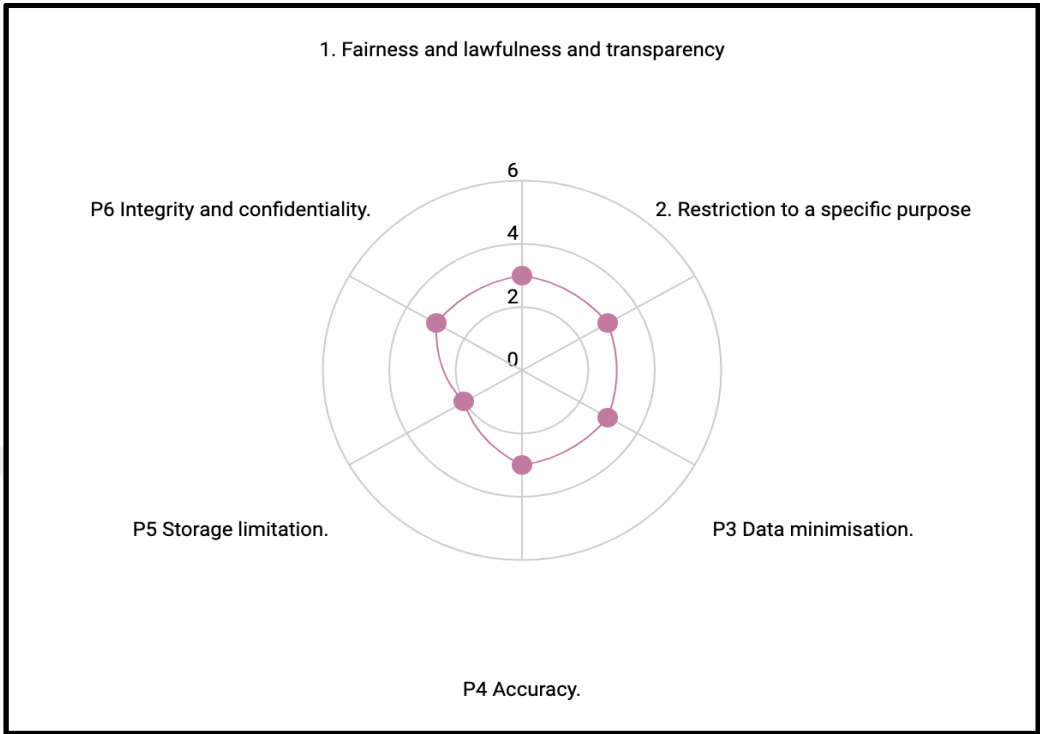
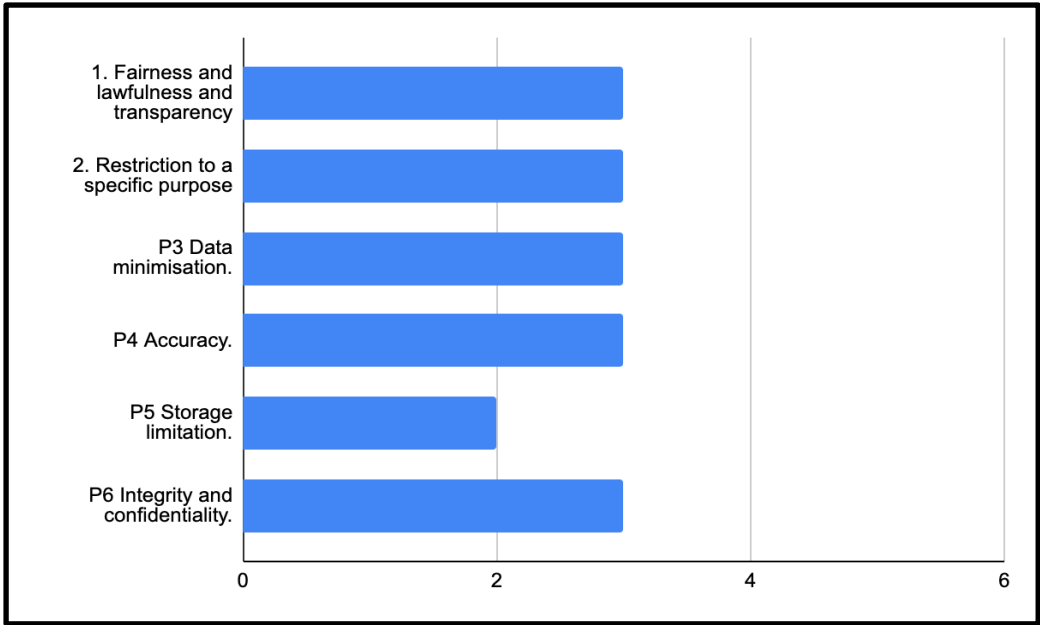


Figure 4.3: Graphical Chart Case study assessment.



If we group back the controls based on the related development phases, as we can see in the following table 4.2.

Table 4.2: Analysis of the case study output based on the development phases

Development phases	Implemented	Not Implemented	Not in scope	Total
Controls related to the Requirement and the Design phases	1	4	0	5
Controls related to the Development Phase and the Testing phase	6	6	4	16
Controls prior to Deployment	5	1		6
Post Deployment	3		1	4
Total	15	11	5	31

It is interesting to notice that among the controls that are in scope three out of three were implemented for the post-deployment phase, five out of six controls from the controls to check prior to deployment were also implemented. This means that during the first release of our case study project the GDPR compliance Issue has been given a good deal of attention.

However, we can see that has not been the case in the other initial phases as about half of the controls related to the development and testing phase were implemented, more precisely only six controls out of twelve controls. As for the design and requirement analysis phase, only one control was implemented. the non-proper-implementation of some of the controls of the analysis of the requirements and the system design phase lead automatically to missing some controls while developing and testing for instance if there was no-legacy gap system assessment conducted while refactoring features it would be hard to understand what parts of the system GDPR non-compliant are and requires change, therefore, the refactoring control is automatically missed. As an output of the analysis of this case study we can affirm that:

- i. In the analysis of requirement and design phases, more time should be dedicated to analyzing the gaps on the legacy system so that the refactoring needs would be much more defined,

- ii. There is a big room for improvement in the automation of tasks such as the deletion of the personal data records automatically in the database and the files in the system.
- iii. There is room for improvement by being Oriented toward a development that supports privacy by design.

4.2.3 Limitations of The Case Study

For making the case study we faced a few obstacles as this thesis was conducted during the pandemic of covid-19 which can cause a bias as it was conducted online. Another risk of biased information can come from the development team members that were interviewed that could have provided non-accurate information.

4.3 MATURITY FRAMEWORK ENHACEMENT

The work we have achieved in this thesis can be enhanced by the organization wishing to implement it using a maturity framework of their choice that complies with the organization strategy to measure the maturity of the implementation of these controls across multiple projects and not only one project.

We propose here that a business process can be define for each control and this process can be assessed by the maturity framework that the organization prefers to adopt, we would give brief example of an organization that already adopt SPICE or CMMI to assess the maturity of their organization processes.

The business processes we offer would be:

Processes related to the software-development requirement and the design phases:

(P 1.1) provide for providing Training and guidelines to development team

(P1.2) Assess GDPR compliance Gap on Legacy system.

(P 1.3) Personal data Identification (Mapping).

(P 1.4) Design /re-Design access management System.

(P 1.5) Sharing data with 3rd parties.

Processes related to the software-development development phase and the testing phase:

(P 2.1) Encryption.

(P 2.2) Data minimization by design.

- (P 2.3) Logs management.
- (P 2.4) Clean databases automatically.
- (P 2.5) Automatic file removal service.
- (P 2.6) Over-writing data in memory/Disk.
- (P 2.7) Pseudonymization by Default.
- (P 2.8) Develop a Personal data automatic search method.
- (P 2.9) Masking Data in the testing environment.
- (P 2.10) Back up data.
- (P 2.11) Informed Consent.
- (P 2.12) Provide privacy settings Features.
- (P 2.13) Refactoring Old features.
- (P 2.14) Hashing.
- (P 2.15) Enforce Data subject to update Data upon Expiry.
- (P 2.16) Cache controls.

Processes related to the software-development to the prior to deployment phase:

- (P3.1) Penetration testing.
- (P3.2) Personal Data - Data flow diagram.
- (P3.3) Update/Create Privacy notice.
- (P3.4) Provide a Method for Data Requests.
- (P3.5) Update internal privacy policy and processes.

Post related to the software-development to after the deployment:

- (P4.1) Create / Update - Incident response program.
- (P4.2) Penetration testing.
- (P4.3) Testing and Auditing.
- (P4.4) Archiving practices.

Where each processes' goal would be to ensure that all development projects in the organization are implementing the respective controls. Each process can be defined, monitored, document and audited. On top of this a maturity framework can be applied based on the organization's preference.

We want to make a high-level review of one maturity framework to provide more clarity. The organization can choose to implement for example ISO/IEC 15504 Process Assessment Model also known as SPICE framework.

ISO/IEC 15504 (Software Process Improvement and Capability Determination) is an ISO-Institute international standard for assessing and evaluating software and electronic development business processes.

SPICE stood for "Software Process Improvement and Capability Evaluation" at first, but after French concerns about the idea of "evaluation," it was updated "Software Process Improvement and Capability Determination." The acronym SPICE is mostly used by the community to refer to this model. ISO/IEC 15504 addresses the assessment of software development processes and associated assessment criteria, as well as the adoption of a clear standard. This standard incorporates the process assessment model (PAM), which specifies how the capability of an organization's processes should be assessed. And that is the part that can be used to assess the previously defined processes as an example. In the process assessment model, we can find that assess the process maturity level in 6 levels as follows in figure 4.4.

Figure 4.4: SPICE process maturity levels (ISO/IEC15504-6).

Maturity Level		Process Attributes / Indicators	
5	"Optimising"	PA 5.1 Process Innovation PA 5.2 Continuous Optimisation	
4	"Predictable"	PA 4.1 Process Measurement PA 4.2 Process Control	
3	"Established"	PA 3.1 Process Definition PA 3.2 Process Deployment	
2	"Managed"	PA 2.1 Performance Management PA 2.2 Work Product Management	
1	"Performed"	PA 1.1 Process Performance	
0	"Incomplete"		

Source :ISO/IEC15504-6

- i. Level 0: Incomplete
- ii. Level 1: Perfumed
- iii. Level 2: Managed
- iv. Level 3: Established
- v. Level 4: Predictable
- vi. Level 5: Optimizing

A maturity level is achieved by a process if it achieves the attributes of that maturity level that are defined in the standard for instance, the attributes that influence the level 2 are the attribute "Performance Management" meaning, that responsibilities and resources are planned, assigned, and monitored and the attribute "Work Product Management" referring to ensuring that work product specifications and goals are met (9).

To consider an attribute as achieved or not the standard provides an assessment of the attribute as 4 measurement scales:

- i. N — Not achieved
- ii. P — Partially achieved
- iii. L — Largely achieved
- iv. F — Fully achieved

If the attribute is largely achieved or fully achieved, it is realized and is mapped accordingly to the maturity levels defined in figure 4.5.

Comments:

In scope of this thesis we focused on assessing GDPR for one given project case, but we have added this paragraph here to show possibilities of extending this work to a larger scale on an organizational level, however it is important to note that have not been the primary focus of our research.

(9): *International Organization for Standardization*, ISO/IEC 15504-6:2013 Information technology — Process assessment — Part 6: An exemplar system life cycle process assessment model, 2013, <https://www.iso.org/fr/standard/61492.html> , [Accessed 01 January 2021] , pp.1-135

5. DISCUSSION AND CONCLUSION

5.1 LIMITATION OF THE THESIS OUTPUT AND PERSPECTIVE RESEARCH

1. This thesis research focused on having a general framework regardless of the type of software that has been developed: web development, embedded system, offline/online video games running in consoles or desktop computers, computer services... etc. More specific research can be conducted on a specific software type, this can set the light on more specific technical controls and methodologies concerning the context of development and could be an interesting research area.
2. We reflected on the 6 GDPR principles in this thesis, but we did not address the 8 GDPR rights to data subjects in a detailed mapping with the controls, this can be covered in a future work.
3. Our framework focuses on the quality of each GDPR principle implementation for a single release; A maturity framework can be used to assess the improvement between releases in regards to the software development methodology used and a case study can be analyzed in a future research for that purpose.
4. Future work may include specific frameworks for development methodologies such as scrum, waterfall, and v-software development models, among others. This can also reflect on controls that are specific to the development model. Ownership of the various controls can also be mapped to the various roles involved in the process, giving development teams more guidance about who should be doing what as a best practice.
5. Building an evaluation system for each control and rating it, rather than only grouping them as must-have controls (for those that will trigger violations if not implemented) and should have controls, will add a layer of accuracy to the importance of the controls to the principal's implementation's quality level.

5.2 CONCLUSION

This research aimed to identify a usable framework to measure GDPR compliance in the scope of software development with respect to the 6 main GDPR principles.

Based on the analysis of the literature review we were able to answer the first research question by gathering a set of controls and map them to the codependent software development phase.

Later, we mapped these controls to the correspondent principles they contribute to, and then we divided the controls based on whether their absence would lead to a direct violation of the GDPR, on this perspective we labeled the controls either as mandatory or as an enhancement to the GDPR compliance this approach helped us draft the answer to our second research question.

Furthermore, we proposed an assessment scale to be able to use this framework on a real case study, we tested our framework assessment scale on a case study and discussed the outputs.

Lastly, we criticized the framework proposed solution and proposed further research perspectives that can complete the work we have achieved in this research.

REFERENCES

Periodicals

Abdulrahman, A.I & Nalin, A.G.A,2021. I'm All Ears! Listening to Software Developers on Putting GDPR Principles into Software Development Practice , pp. 2-18.

Aysem D.V,2018. "The Right to Data Portability in the GDPR: What Lessons Can Be Learned from the EU Experience?" , pp. 3-19.

Dataguise, Inc. ,(2010) .Why Add Data Masking To Your Best Practices For Securing Sensitive Data , pp. 2-8.

Other publication

BACS 371 Computer Forensics. 2016. <https://slideplayer.com/slide/6396157/> [Accessed 20 March 2021]

British Standards Institution, BS 10012 Personal Information Management System, 2017, <https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/> [Accessed 01 December 2020] , pp1-42.

Charles, E.P., 2004 . Security Assurance For A Resource-Based RBAC/DAC/MAC security Model. *PHD thesis* , University of connecticut.

Computer science wiki , Hashing article , 2021, <https://computersciencewiki.org/index.php/Hashing> [Accessed March 2021], p.1.

ePrivacy Directive Law, 2002.

Facebook, privacy basics page, 2021, <https://www.facebook.com/about/basics> [Accessed January 2021]

Facebook, settings page, 2021, <https://www.facebook.com/settings?tab=privacy> [Accessed January 2021]

GDPR Enforcement Tracker , statista.com , September 2020 . Fines issued for General Data Protection Regulation (GDPR) violations as of September 2020, by type of violation. *Region , Region : Europe.*

General Data Protection Regulation, 2016.

International Organization for Standards, ISO/IEC 15504-6:2013 Information technology — Process assessment — Part 6: An exemplar system life cycle process assessment model, 2013, <https://www.iso.org/fr/standard/61492.html> , [Accessed 01 January 2021] , pp.1-135

International Organization for Standards, ISO/IEC 27001:2005 Information technology — Security techniques — Information security management systems — Requirements ,2005, <https://www.iso.org/standard/42103.html>, [Accessed 01 December 2020] , pp.1-34.

International Organization for Standards, ISO/IEC 27001:2013 Information technology — Security techniques — Information security management systems — Requirements ,2013, <https://www.iso.org/standard/54534.html>, [Accessed 01 December 2020] , pp.1-23.

International Organization for Standards, ISO/IEC 27002:2005 Information technology — Security techniques — Code of practice for information security

- management, 2005, <https://www.iso.org/standard/50297.html> [Accessed 01 December 2020] , pp.1-115.
- International Organization for Standards*, ISO/IEC 27002:2013 Information technology — Security techniques — Code of practice for information security management, 2005, <https://www.iso.org/standard/54533.html> [Accessed 15 December 2020] , pp.1-80.
- International Organization for Standards*, ISO/IEC 27701:2019 Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines, 2019, <https://www.iso.org/standard/71670.html> [Accessed 20 December 2020] , pp.1-66.
- Owasp, Application Security Verification Standard version 4.0, 2019, https://owasp.org/www-pdf-archive/OWASP_Application_Security_Verification_Standard_4.0-en.pdf , [Accessed 20 January 2021]
- OWASP, Cheat-Sheets project ,2021, <https://cheatsheetseries.owasp.org/> [Accessed February 2021]
- OWASP, Top ten project , 2013 , <https://owasp.org/www-project-top-ten/> [Accessed February 2021]
- Reini & Pasi .,(March2019). head power Oy. *Master's Thesis* .JAMK university of applied science.
- Ronald Costea , “Privacy & Data Protection Online Academy”, 2020 : <https://www.defradar.com/> [Accessed December 2020]
- Scrum , Scrum Standard Guideline: ”The Definitive Guide to Scrum The Rules of the Game ”,2017, scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf , [March 2020.] ,pp.3-18.
- VeraCrypt , veraCrypt documentation , 2013, <https://www.veracrypt.fr/en/Documentation.html> [Accessed February 2021]