PReparing Industry to Privacy-by-design by supporting its Application in REsearch

Deliverable D1.1
Privacy and Security, Concepts and Principles Report

Project: PRIPARE
Project Number: ICT-610613
Deliverable: D1.1
Title: Privacy and Security, Concepts and Principles Report
Version: v1.00
Date: 14/03/2014
Confidentiality: Public
Author/s: Antonio Kung (Trialog)
Alberto Crespo Garcia (ATOS)
Nicolás Notario McDonnell (ATOS)
Inga Kroener (Trilateral)
Daniel Le Métayer (Inria)
Carmela Troncoso (Gradiant)
José María del Álamo (UPM)
Yod Samuel Martín (UPM)

Funded by the European Union's Seventh Framework Programme
# Table of Contents

**DOCUMENT HISTORY** .................................................................................................................. 4

**LIST OF FIGURES** ......................................................................................................................... 5

**ABBREVIATIONS AND DEFINITIONS** ............................................................................................ 6

**EXECUTIVE SUMMARY** .................................................................................................................. 8

1 **INTRODUCTION** ............................................................................................................................. 10

2 **PRIVACY BY DESIGN AND SECURITY BY DESIGN WITHIN PRIPARE** ...................... 12
   2.1 **PRIVACY BY DESIGN AND SECURITY BY DESIGN PRINCIPLES** ................................................. 12
   2.2 **PRIPARE’S METHODOLOGY** ......................................................................................................... 13

3 **TERMINOLOGY** ............................................................................................................................... 15
   3.1 **TERMS AND DEFINITIONS** ........................................................................................................... 15
      3.1.1 Security, risk management and other terms .................................................................................. 15
      3.1.2 Privacy terms ............................................................................................................................... 19

4 **PRINCIPLES** ..................................................................................................................................... 28
   4.1 **SECURITY PRINCIPLES AND BEST PRACTICES** ............................................................................ 28
   4.2 **PRIVACY PRINCIPLES** .................................................................................................................. 30
      4.2.1 Data protection principles ...................................................................................................... 31

5 **COMMON GROUNDS NECESSARY TO ENGINEER TRULY TRUSTWORTHY SYSTEMS** .... 36
   5.1 **ETHICAL IMPACT ASSESSMENTS (EIAs)** ...................................................................................... 36
   5.2 **BEST PRACTICES ON PIAS** ........................................................................................................ 37
   5.3 **RISK MANAGEMENT PROCESS** .................................................................................................. 39
      5.3.1 Common notions and principles ................................................................................................. 40
      5.3.2 CNIL ........................................................................................................................................... 42
   5.4 **USER EMPOWERMENT** ................................................................................................................. 43
      5.4.1 Identity management .................................................................................................................. 44
      5.4.2 User-centricity ........................................................................................................................... 45
   5.5 **PRIVACY INSTRUMENTS** ............................................................................................................. 46
   5.6 **ARCHITECTURAL SELECTION** .................................................................................................... 47

6 **THE CHALLENGES OF INTEGRATING INTO EXISTING METHODOLOGIES** ........ 48
   6.1 **SOFTWARE AND SYSTEM ENGINEERING METHODOLOGIES** ............................................. 48
      6.1.1 Waterfall ..................................................................................................................................... 48
      6.1.2 Iterative or incremental .............................................................................................................. 49
      6.1.3 Prototype .................................................................................................................................... 50
      6.1.4 Agile .......................................................................................................................................... 51
   6.2 **PROJECT MANAGEMENT METHODOLOGIES** ............................................................................ 52
      6.2.1 PMBOK ...................................................................................................................................... 52
      6.2.2 PRINCE2 .................................................................................................................................... 53
7 CONCLUSIONS ................................................................. 54
8 REFERENCES ................................................................. 56
Document History

<table>
<thead>
<tr>
<th>Version</th>
<th>Status</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>Initial draft with contributions from Atos</td>
<td>10/01/2014</td>
</tr>
<tr>
<td>0.20</td>
<td>Update sections with contributions from Inria</td>
<td>19/01/2014</td>
</tr>
<tr>
<td>0.30</td>
<td>Update sections with contributions from Trilateral</td>
<td>20/01/2014</td>
</tr>
<tr>
<td>0.40</td>
<td>Update sections with contributions from UPM</td>
<td>24/01/2014</td>
</tr>
<tr>
<td>0.50</td>
<td>Update sections with contributions from Gradiant</td>
<td>26/01/2014</td>
</tr>
<tr>
<td>0.60</td>
<td>Update sections with contributions from Trialog</td>
<td>29/01/2014</td>
</tr>
<tr>
<td>0.70</td>
<td>Update sections with contributions from Trilateral</td>
<td>30/01/2014</td>
</tr>
<tr>
<td>0.80</td>
<td>Merge all sections aligning styles and content</td>
<td>05/02/2014</td>
</tr>
<tr>
<td>0.90</td>
<td>Creation of Executive Summary and Conclusions</td>
<td>12/02/2014</td>
</tr>
<tr>
<td>1.00-draft</td>
<td>Final draft sent for Quality Review</td>
<td>19/02/2014</td>
</tr>
<tr>
<td>1.00</td>
<td>Final version</td>
<td>14/03/2014</td>
</tr>
</tbody>
</table>

Approval

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepared</td>
<td>Person Name</td>
</tr>
<tr>
<td>Reviewed</td>
<td>All Project Partners</td>
</tr>
<tr>
<td>Authorised</td>
<td>Antonio Kung</td>
</tr>
</tbody>
</table>

Circulation

<table>
<thead>
<tr>
<th>Recipient</th>
<th>Date of submission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project partners</td>
<td>day/month/year</td>
</tr>
<tr>
<td>European Commission</td>
<td>day/month/year</td>
</tr>
</tbody>
</table>
# List of Figures

Figure 1: Risk management strategies  
Figure 2: CNIL's five phases  
Figure 3: Waterfall methodology phases  
Figure 4: Iterative methodology phases  
Figure 5: Prototype methodology phases [33]  
Figure 6: Scrum methodology phases  
Figure 7: PMBOK process and knowledge matrix  
Figure 8: PRINCE2 seven processes
# Abbreviations and Definitions

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASTRA</td>
<td>Aegis Strategic Risk Assessment</td>
</tr>
<tr>
<td>CNIL</td>
<td>Commission Nationale de l'Informatique et des Libertés</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>DoS</td>
<td>Denial of Service</td>
</tr>
<tr>
<td>DPIA</td>
<td>Data Protection Impact Assessment</td>
</tr>
<tr>
<td>DPO</td>
<td>Data Protection Officer</td>
</tr>
<tr>
<td>EBIOS</td>
<td>Expression des Besoins et Identification des Objectifs de Sécurité</td>
</tr>
<tr>
<td>EC</td>
<td>European Commission</td>
</tr>
<tr>
<td>ECHR</td>
<td>European Commission of Human Rights</td>
</tr>
<tr>
<td>ECI</td>
<td>European Court of Justice</td>
</tr>
<tr>
<td>EDPS</td>
<td>European Data Protection Supervisor</td>
</tr>
<tr>
<td>EIA</td>
<td>Ethical Impact Assessment</td>
</tr>
<tr>
<td>ESRIF</td>
<td>European Security Research and Innovation Forum</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FIPPs</td>
<td>Fair Information Practice Principles</td>
</tr>
<tr>
<td>FOSAD</td>
<td>Foundations of Security Analysis and Design</td>
</tr>
<tr>
<td>FRAAP</td>
<td>Facilitated Risk Analysis and Assessment Process</td>
</tr>
<tr>
<td>FTC</td>
<td>Federal Trade Commission</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-computer interaction</td>
</tr>
<tr>
<td>HIPAA</td>
<td>Health Insurance Portability and Accountability Act</td>
</tr>
<tr>
<td>ICO</td>
<td>Information Commissioner's Office</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technologies</td>
</tr>
<tr>
<td>IMS</td>
<td>Identity Management Systems</td>
</tr>
<tr>
<td>IOI</td>
<td>Items of Interest</td>
</tr>
<tr>
<td>IEEE</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>ISO</td>
<td>International Standards Organization</td>
</tr>
<tr>
<td>ISTPA</td>
<td>International Security, Trust and Privacy Alliance</td>
</tr>
<tr>
<td>LIBE</td>
<td>Civil Liberties, Justice and Home Affairs</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
<tr>
<td>NSA</td>
<td>National Security Agency</td>
</tr>
<tr>
<td>NTIA</td>
<td>National Telecommunications &amp; Information Administration</td>
</tr>
<tr>
<td>OASIS</td>
<td>Organization for the Advancement of Structured Information Standards</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>OWASP</td>
<td>Open Web Application Security Project</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PbD</td>
<td>Privacy by Design</td>
</tr>
<tr>
<td>PbD-SE</td>
<td>Privacy by Design Documentation for Software Engineers</td>
</tr>
<tr>
<td>PET</td>
<td>Privacy Enhancing Technology</td>
</tr>
<tr>
<td>PIA</td>
<td>Privacy Impact Assessment</td>
</tr>
<tr>
<td>PII</td>
<td>Personal Identifiable Information</td>
</tr>
<tr>
<td>PIR</td>
<td>Private Information Retrieval</td>
</tr>
<tr>
<td>PMBOK</td>
<td>Project Management Body of Knowledge</td>
</tr>
<tr>
<td>PMI</td>
<td>Project Management Institute</td>
</tr>
<tr>
<td>PMRM</td>
<td>Privacy Management Reference Model</td>
</tr>
<tr>
<td>PRINCE2</td>
<td>Projects in Controlled Environments, version 2</td>
</tr>
<tr>
<td>PRIPARE</td>
<td>PReparing Industry to Privacy-by-design by supporting its Application in REsearch</td>
</tr>
<tr>
<td>PSbD</td>
<td>Privacy and Security by Design</td>
</tr>
<tr>
<td>PwC</td>
<td>PricewaterhouseCoopers</td>
</tr>
<tr>
<td>RFID</td>
<td>Radio Frequency IDentification</td>
</tr>
<tr>
<td>SbD</td>
<td>Security by Design</td>
</tr>
<tr>
<td>SEI</td>
<td>Software Engineering Institute</td>
</tr>
<tr>
<td>STRIDE</td>
<td>Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, Elevation of Privilege</td>
</tr>
<tr>
<td>TC</td>
<td>Technical Committee</td>
</tr>
<tr>
<td>TOE</td>
<td>Target Of Evaluation</td>
</tr>
<tr>
<td>TPM</td>
<td>Trusted Platform Modules</td>
</tr>
<tr>
<td>TSF</td>
<td>TOE Security Functionality</td>
</tr>
<tr>
<td>OGC</td>
<td>Office of Government Commerce</td>
</tr>
<tr>
<td>UCA</td>
<td>User-Centered Design</td>
</tr>
<tr>
<td>UMA</td>
<td>User Managed Access</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>W3C</td>
<td>World Wide Web Consortium</td>
</tr>
<tr>
<td>XML</td>
<td>eXtensible Markup Language</td>
</tr>
</tbody>
</table>
Executive Summary

Currently we are living a rapid technological development and trends such as big data, ubiquitous computing on mobile devices/Internet of Things, cloud computing, schemes for interoperability of data (including identity and other personal attributes). The application of these development and trends in areas like smart grids, active ageing, intelligent transport systems, or smart cities are determining an unprecedented increase in the amount of data collected, shared and processed by public and private organisations. At the same time, increasing public reports about serious privacy breaches (notably the recent privacy violations by the NSA affecting EU citizens and institutions and tolerated by major Internet companies such as Google and Apple) have raised privacy concerns of citizens to new levels and have posed a real risk of generating a step backwards at societal level in terms of trust in ICT systems.

Despite relevant efforts made by CNIL, UK ICO or the EC (PIAF project) that are detailed in the Common grounds necessary to engineer truly trustworthy systems (see Section 5), Privacy by Design is still not a widely adopted practice in engineering ICT systems. There are different explanations for this current lack of adoption:

- There is no consistent and homogeneous terminology among stakeholders;
- It is difficult to translate privacy and security principles into real engineering practices;
- Current privacy practices are not easily available to policy makers, missing out on the chance to provide them with a source of inspiration;
- Lack of applicable guidelines and engineering methods for privacy by design;

Furthermore, it is impossible to ignore the dialectic close relationship between privacy and security; one cannot be achieved without the other.

PRIPARE endeavours to tackle the lack of adoption of PbD. Therefore this document aims to:

- Define a common terminology that can be understood by stakeholders (i.e. belonging to different disciplines like technologists and legal experts) without ambiguities. This will ease requirements and design discussions and the development of much needed effective synergies;
- List, discuss and agree on common understandings for key security and privacy principles and concepts that PRIPARE needs to consider. This understandings will be used in order to improve, from a procedural perspective, existing methodological ICT systems and project management lifecycle paradigms in order to enhance privacy;
- Review existing PbD literature to help establish a position of the consortium with respect to the key aspects to address in order to effectively operationalize the engineering of privacy and security by design and facilitate their uptake by different stakeholders. The outcome of this review will be summarized in a public position paper to be discussed in multi-disciplinary workshops (i.e. Annual Privacy Forum 2014 in May, 2014 [87]) with invited external experts and relevant stakeholders.
- Present the common grounds that have been identified as necessary by the consortium in order to engineer truly trustworthy systems;
• Identify most widely adopted project management and system development methodologies to envision challenges that PRIPARE’s methodology will have to overcome.

PRIPARE’s methodology will include current practices that optimize PIA processes and privacy risk management methodologies. It will also systematically integrate corresponding best practices and state-of-the-art recommendations with common system and project management methodologies. A full risk management process will be embedded in the methodology to promote the use of adequate security mechanisms in order to minimize the risk of exposure of private information. The EU Data Protection Directive enforces the concept of user-centricity by stating different rights that give the data subject control over his/her data (e.g., right to access, to object, the transparency principle). These user-centric aspects are central to privacy and will likewise be central to PRIPARE’s methodology.
1 Introduction

It is nowadays more essential than ever to create the conditions that allow to achieve high levels of public trust in ICT systems in order to facilitate faster adoption of new services and technologies that feature high and tangible levels of privacy and security embedded into their design and provided by default. Hence it will consequently increase the speed of innovation and creation of added value for a more competitive European ICT industry. Many courses of action have been taken in order to increase trust levels such as:

- Different EU Projects: e.g. as part of the FP7 programme, PARIS, PRECIOSA, PACT, POPCORN, PRESERVE, TERESA, PICOS, PIAF, or PIAw@tch;
- Legislation: the new Data Protection Regulation, a major reform of the EU legal framework on the protection of personal data and the commitment to achieve higher levels of data protection from Article 29 Data Protection Working Party, the European Data Protection Supervisor and Data Protection Authorities in the Member States;
- Standardisation efforts: HIPAA Privacy Rule, ISO 29100:2011, OASIS (PbD-SE & PMRM) TCs...

In order to propose an engineering approach to Privacy by Design (PbD) and to support PRIPARE’s Privacy and Security by Design (PSbD) methodology and reference model the first step is to set some common basis regarding:

- Terminology related to key privacy, risk and security concepts;
- Security and privacy principles;
- Common grounds to engineer truly trustworthy systems.

The use of terminologies offering disparate definitions for privacy, risk and security concepts and principles currently stands as a barrier for the adoption of PbD. Furthermore, stakeholders from different industry domains and professional areas, with different interests and backgrounds, use the same terms but with different meanings. In order to have a fluid, streamlined and effective communication, the terminology must be homogenised. Multiple sources (e.g. OASIS and ISO standards [6], [19], [17] or CNIL’s methodology for risk management process [11]) have been reviewed to collect privacy, risk and security relevant and controversial terms. These terms have been subject to a debate within the consortium and agreed understandings will be further validated with external experts to consolidate meanings that will be used within PRIPARE and hopefully adopted in other projects or relevant domains.

To determine if a system is privacy-respectful, its design and functions have to follow a set of previously agreed privacy principles. PRIPARE will provide its own set of principles, derived and extended from authoritative sources, that can in turn be modified or adapted for different contexts.

There are currently many technologies (Privacy Enhancing Technologies) or processes (i.e. PIAs) that help developing secure and privacy-respectful systems. PRIPARE’s methodology will try to combine the use of these tools into an effective and applicable methodology. This deliverable presents and analyses several of these state-of-the-art tools that may be further used to design the PSbD methodology in the Privacy instruments section (see section 5.5).
PIA
Since Blair Stewart’s first definition of a PIA in 1996 [37] many organisations and institutions have provided their own definition but there is a general consensus that a PIA:
- Is a process;
- Is systematic;
- Identifies and mitigates privacy risks;

PRIPARE is not only taking into account existing PIA frameworks but will also take a close look at PIA-focussed projects like PIAF and their recommendations [38].

Risk Assessment methodologies
Although PIAs generally include some kind of risk assessment steps, PRIPARE will study standalone risk management processes, especially privacy risks management methodologies. Best practices of these methodologies will be embedded in PRIPARE’s PSbD methodology.

OASIS
PRIPARE is not the only existing action trying to fill the gap between PbD theory and applicable engineering processes. OASIS has a technical committee trying to standardise the PbD documentation for Software Engineers, the PbD-SE TC and another one aiming to develop a Privacy Management Reference Model and Methodology (PMRM). While OASIS PMRM is mainly focused on the analysis phase of the ICT system engineering process, PRIPARE considers that its methodology must cover the engineering process from end to end, including all phases (e.g. design, implementation and verification). Anyhow, the work of both committees and their output will be closely examined looking for possible mutual benefits.

This document is organised as follows:
- Chapter 1 (this chapter) – presents an overview of the main purposes of this project and this specific document;
- Chapter 2 – Privacy by Design and Security by Design within PRIPARE: PRIPARE’s views on Privacy and Security by Design;
- Chapter 3 – Terminology: a comprehensive list of terms related to Privacy and Security by Design;
- Chapter 4 – Principles: security and privacy principles adopted within PRIPARE;
- Chapter 5 – Common grounds necessary to engineer truly trustworthy systems: defines the main agreements of the consortium on what characteristics and features the methodology should incorporate;
- Chapter 6 – The challenges of integrating into existing methodologies: lists the main project management and system engineering methodologies PRIPARE aims to complement and what can be the principal challenges of this complementation;
- Chapter 7 – Conclusions: summary of PRIPARE’s agreements and next steps;
2 Privacy by Design and Security by Design within PRIPARE

2.1 Privacy by Design and Security by Design Principles

PbD is usually defined as a number of principles that designers can apply from the start of the system development ensuring that privacy is addressed correctly including that the system achieves data protection compliance. For instance, in the well-known articulation of PbD proposed by Ann Cavoukian [4], the third principle states that: “Privacy Embedded into Design: Privacy is embedded into the design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that it becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality.” Similarly, the opinion of the EDPS that fosters the inclusion of mandatory PbD [40] in European regulation defines this principle as “privacy and data protection are embedded within the entire lifecycle of the technology, from the very early design stage, right through to their ultimate deployment, use and ultimate disposal”.

The proposal for a General Data Protection Regulation [13], as voted in LIBE Committee by the European Parliament, mentions the need for applying PbD, and mandates Data protection by design and by default: “Having regard to the state of the art, current technical knowledge, international best practices and the risks represented by the data processing, the controller and the processor, if any, shall, both at the time of the determination of the purposes and means for processing and at the time of the processing itself, implement appropriate and proportionate technical and organisational measures and procedures in such a way that the processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject, in particular with regard to the principles laid out in Article 5”.

These articulations describe high level activities that in general are difficult to translate to practice. While in all above definitions the PbD principle’s goal is clear, there is a lack of guidelines for engineers about how to carry out the task, hindering the adoption of PbD in everyday business practices. One of the main reasons for this disconnection between recommendations and their application lies in the limited past experience in designing systems with privacy in mind. Until the recent widespread of pervasive ICT systems strong privacy-preserving mechanisms and solution were mostly found in the academic realm, with some rare exceptions, e.g., the Tor network [41]. Only lately we started seeing some use of strong privacy-preserving cryptography in commercial products, such as Porticor [42] that used homomorphic encryption for key management protecting the key from Cloud providers, or CryptDB [43], a privacy-preserving database used by Google in their forthcoming Encrypted BigQuery client. These implementations are typically invisible or inaccessible to policy makers who discuss the principles of PbD, and hence cannot serve as inspiration for the definitions they have to enunciate.

The focus of PRIPARE is to develop a methodology to bring PbD into practice and to foster a risk management culture through educational material targeted to a diversity of stakeholders. However, it is important to note that most privacy-preserving solutions rely on basic security engineering mechanisms. In addition to data confidentiality, which is usually achieved by encryption, mechanisms for guaranteeing integrity or availability are necessary to ensure privacy. For instance, mechanisms to ensure the integrity of exchanged messages are necessary
to avoid active attacks that may be used to break the encryption algorithms and hence provoke a confidentiality breach that affects privacy. It has also been shown that availability can be crucial for privacy. Mittal et al. demonstrate in [44] that a well-designed Denial of Service (DoS) strategy allows an adversary to greatly degrade anonymity in low-latency anonymous peer to peer networks.

For this reason the project shall also consider Security by Design (SbD) principles, ensuring that the PRIPARE methodology encompasses security engineering activities such as risk and threat analysis when needed. In other words, the proposed methodology must promote the use of adequate security mechanisms to minimize the risk of exposure of private information. This dual perspective also ensures that the results of the project shall be of use to complement and reinforce SbD methodologies [25] and adopt the definition of existing SbD principles (e.g. the Application Security Principles from the Open Web Application Security Project [31]).

2.2 **PRIPARE’s methodology**

Within the PRIPARE project we revisit the existing articulations of PbD principles in order to develop a methodology that helps stakeholders to implement PbD in practice. The methodology shall bridge the gap between policy and engineering, establishing guidelines and procedures to transpose the principles into technological and organisational requirements that the engineers can use to embed PbD in the systems they develop. A crucial requirement for the methodology is that, despite its technological orientation to complement current regulation, it must be useful for all stakeholders involved in the process including regulators and policy makers, business managers and privacy officers, and not only the designers and engineers that ultimately have to implement the system.

A second requirement is to keep the PRIPARE methodology generic enough to accommodate privacy concerns and solutions in a broad range of ICT domains. Due to the fact that the intrinsic nature of PbD requires addressing privacy from the first steps of system analysis and design, the proposed methodology will be limited to integrating privacy in new ICT systems, consequently the integration of privacy protection measures when re-engineering systems is not in the scope of the project. Nevertheless, it must be noted that still some approaches may be relevant for re-engineering projects, and therefore nothing prevents stakeholders to re-use those aspects of the PRIPARE methodology that they deem useful in the context of such undertakings.

Since the project’s goal is to cover a broad range of ICT systems, privacy must also be regarded in a broad sense. The project will focus not only on personal data protection (informational privacy), but shall also address the other six types defined by Finn et al [39] (listed and discussed in the Privacy principles section 4.2). The methodology developed in PRIPARE shall provide designers with pointers to state-of-the-art privacy-preserving technologies addressing these privacy concerns, e.g., anonymous credentials or anonymous communication systems. This shall promote the use of these technologies and their integration in deployed systems, mitigating the risks for privacy that stem from the collection, distribution, aggregation and disposal of personal data.

The methodology developed in PRIPARE will establish the grounds for engineering PbD into real systems. Engineering systems with privacy in mind requires integrating privacy requirements
into the typical systems engineering activities. Hence, the PRIPARE methodology will elaborate on the following activities:

- Eliciting and analysing privacy requirements, at the same time as the system’s functional and non-functional requirements; and perform privacy risk and impact assessments.
- Developing designs that fulfil those requirements considering multilateral requirements.
- Implementing the design under realistic restrictions, and in consideration of state of the art technologies.
- Validating that the functional and privacy requirements are fulfilled.

The starting point of the methodology is the idea of minimizing the trust that users need to place on the data controllers or data processor which will be collecting, storing and processing their personal data. This principle implicitly ensures that the data minimization principle is fulfilled, since the best approach to minimize trust is to minimize the amount of data that needs to be entrusted\(^1\). The methodology will seek to minimize the amount personal data distributed to potentially untrustworthy parties, which in turn minimizes the risk of privacy breaches.

The methodology will pay special attention to what “technically comply with the data minimization principle” means. The guidelines provided will help designers and engineers to use and integrate state of the art privacy-preserving technologies in their systems. For example, technologies like anonymous credentials that allow to decouple the identity of an individual from the actions he or she performs; technologies that allow authentication based on attributes without the need to expose a persistent identity; anonymous communications, that ensure that anonymity cannot be revoked at the communication layer; or zero-knowledge proofs, a cryptographic mechanism that proves that a certain statement is true without revealing any further information (e.g., a zero-knowledge protocol is able to prove that the subject is over the age of 18 without revealing the actual date of birth, or any other information [46]). Similarly, systems can be developed where individuals are identified, but it is not possible to observe their activities. For example, Private Information Retrieval (PIR) mechanisms allow authorized users to query a database, but prevent the owner of the database from learning which database items have been queried. Within PRIPARE there are specific tasks that aim to identify privacy patterns that can be used as guidelines for engineers during the ICT system development process. These privacy patterns will aid engineers to choose the adequate PETs for systems. Privacy patterns will be classified indicating the right phase during the engineering process where to apply them, which qualities does the pattern provide and some legislation references to help legal practitioners.

The use of strong privacy-preserving technical protection mechanisms will ease the design of architectures and choice of mechanisms to minimize trust, and hence maximize privacy. It is foreseen that the socio-economic and legal context in which systems have to be deployed will influence the levels of protection beyond what is achievable by technological means alone. PRIPARE will also provide guidelines to provide privacy through organizational and technical forms of control and transparency.

\(^1\) “Protecting privacy by minimizing trust” is an on-going work from some of PRIPARE partners that will be published in the future.
3 Terminology

In privacy and data protection it is often the case that for some terms there are a variety of interpretations among different countries and/or professional areas. Hence the need to provide a comprehensive list of terms and definitions to be part of PRIPARE’s reference model. To provide the necessary unified terminology to describe privacy and PbD, PRIPARE’ partners will follow the steps listed below:

- Identify the most elusive and controversial terms using the following as sources:
  - ISO standards [6], [19];
  - OASIS PMRM [17];
  - CNIL’s methodology for privacy risk management [11];
  - Papers regarding PbD and security;
- For each term, concept or principle requiring a common understanding or agreed definition, create a discussion between the partners to reach consensus about the best source for the definition or to agree to a common definition;
- Present PRIPARE’s terminology at an upcoming multi-disciplinary methodological workshop to gain extra feedback from experts from different professional areas.

For the last step, PRIPARE consortium has agreed to present a position paper in the Annual Privacy Forum 2014 (May 2014) [87]; input from that workshop will be incorporated in the first iteration of the PRIPARE methodology.

As a result of these steps, the term definitions will be considered consolidated and be used as a reference for other outcomes of the project.

The result of the two first steps is presented in this section.

3.1 Terms and definitions

3.1.1 Security, risk management and other terms

Actor
According to PMRM reference model [17], an actor is “a data subject or a human or a non-human agent or (sub)system interacting with personal data within a Privacy Domain or System”.

PRIPARE’s methodology will make use of this definition and will endeavour to identify all the actors of the to-be-developed ICT systems.

Architectural tactics
Bachmann et al. [101] defines architectural tactics as “the means of satisfying a quality-attribute-response measure by manipulating some aspect of a quality attribute model through architectural design decisions”. The attribute-driven design (ADD) method [102] advises architects to use tactics when they need help coming up with a pattern, when an existing pattern isn’t quite right and they need to tailor it, or when they want to validate the choice of a pattern. Generally architectural patterns package tactics, and can be described in terms of them.

Architecture
Many relevant definitions exist for the architecture term depending on its context. For PRIPARE’s purposes the proper context is “System architecture”. IEEE defines architecture in its
glossary [8] as the “system fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and in the principles of its design and evolution”.

PRIPARE will define a system design methodology which elicits privacy requirements from the outset, which focuses on the specification of the relevant privacy requirements, the selection of appropriate architecture tactics (e.g. anonymous credentials), and the resulting proposal of appropriate privacy patterns.

Asset
An asset in a security context, according to Le Métayer [26], can be defined as “anything (concrete or abstract, part of the IT product or under its control) which has a value and needs to be protected by a security policy”.

PRIPARE will specifically use the term “privacy asset” to designate those assets related to the privacy domain.

Attack
According to Le Métayer’ definition in his Security Best Practices chapter [26], and extended to include privacy along with security, “an attack is the manifestation of a threat or a succession of threats (by attackers) resulting in a security or privacy breach”.

Boundary object
PMRM reference model [17] uses the boundary object concept, introduced by Susan Leigh Star and James R. Griesemer in a 1989 publication [96] and defines it as “a sociological construct that supports productive interaction and collaboration among multiple communities”. OASIS used this concept to define the PMRM process output. The output is accessible and relevant to all stakeholders, but each group takes from it and attributes to it what they specifically need. Boundary objects can facilitate collaboration across relevant communities. This is a very interesting point for PRIPARE. Methodology outputs should be considered boundary objects that will address different stakeholder needs (e.g. engineers, decision makers and customers).

Countermeasure
In FOSAD Security Best Practices [26], Le Métayer defines a countermeasure as “a (technical or organizational) mechanism or measure to reduce vulnerabilities”. A more privacy-related equivalent term would be the Privacy Enhancing Technologies (PETs). Le Métayer notes that “controls” can be used as a countermeasure synonym. Control is the preferred term as it is more aligned with ISO definitions (i.e. privacy controls).

Functional requirement
IEEE [8] defines functional requirement as “a requirement that specifies a function that a system or system component must be able to perform”. Generally acknowledged characteristics of good requirements, regardless if they are functional or non-functional, are: unitary, complete, consistent, atomic, traceable, current, unambiguous, verifiable and specify importance [100].

Generality
In the risk management domain, generality is the ability to cope with all aspects, including organizational, technical as well as management issues. The lack of generality, or the inability to provide a complete view of all security and privacy issues, may lead to overlook significant issues or to spend too much energy and time on minor items when other, more significant, aspects, are underestimated.

Incrementality
In the risk management domain, incrementality is crucial because, in practice, risk analyses can rarely be one shot undertakings. Most organisations prefer to start with a preliminary analysis, which should produce first conclusions as soon as possible at a moderate cost, before deciding to embark on more complex analysis.

**Measure**

CNIL [11] defines a measure, in the risk management domain, as “an action to be taken to treat risks. It may be to avoid, modify/reduce, share/transfer or retain them.

PRIPARE will help engineers to find the right measures to minimise privacy and security risks within an ICT system.

**Non-functional requirement**

Complementary to functional requirements, non-functional requirements describe “how a system is supposed to be”. Non-functional requirements are often called *quality attributes* and share with functional requirements the characteristics for good requirements (unitary, complete, consistent...).

**Rigour**

Rigour is obviously a virtue of any method, but it is a prerequisite for any method to be used to justify design choices, especially in the context of accountability, and when design decisions can have legal consequences. Rigour can be achieved through the application of systematic rules. Systematization itself brings additional benefits: it improves the efficiency of the process (and therefore reduces delays and costs, which are crucial factors in the case of innovative products for which time-to-market is often decisive) and enhances repeatability and maintenance.

**Risk**

Le Métayer [26] defines risk as “a measure of the potential impact of attacks (or losses incurred as a consequence of attacks)”. CNIL [11] describes a risk as a “Scenario describing a feared event and all threats that make it possible. It is estimated in terms of severity and likelihood.” CNIL’s definition has some differences that should be highlighted:

- While Le Métayer talks about measures, CNIL uses estimations given the uncertainty of the events;
- CNIL estimations are based on both, the severity and likelihood of the event while Le Métayer’s definition only uses the potential impact, ignoring the likelihood.

For PRIPARE, CNIL’s definition is preferable. The concept of *privacy risk* is introduced as that risk that threatens privacy.

**Privacy risk management**

According to CNIL [11], a privacy risk management process is an “Iterative process that allows to objectively manage the privacy risks on the data subjects concerned by a processing of personal data. It essentially consists in appreciating them (identification, estimation in terms of severity and likelihood, and evaluation for comparison), treating them (determining and implementing proportionate measures), accepting residual risks, communicating (stakeholder consultation, results presentation...), and monitoring changes over time (context, risk, measures...).”

PRIPARE, within its methodology will define the necessary steps to conduct a privacy risk management process.

**Quality attributes**
See Non-functional requirement’s definition.

Risk source
CNIL [11] describes a risk source as a “Person or non-human source that can cause a risk, accidentally or deliberately”.

Security breach
Daniel Le Métayer defines a security breach and it can be easily adapted to the privacy domain as “a violation of the security or privacy policy with detrimental impact on the assets (asset destruction, asset disclosure, denial of service, etc.).”[26].

Security by design
The European Security Research and Innovation Forum (ESRIF) in its 2009 final report [98] defines the Security by design concept as “to embed security in the technology and system development from the early stages of conceptualisation and design”. It also recommends institutionalizing security considerations in the organisations. This term is closely entangled with PbD. Security and privacy are related in such a way that they cannot exist one without the other and PbD aims to do for privacy the same as SbD has done for security.

The relation between PbD and SbD is better and more extensively described in Ann Cavoukian’s paper [25] and her view is shared among PRIPARE consortium:
“Privacy must be proactively incorporated into networked data systems and technologies, by default. The same is true of security. Both concepts must become integral to organizational priorities, project objectives, design processes, and planning operations.”

Security requirement
OWASP defines security requirements in its quick reference guide [99] as “a set of design and functional requirements that help ensure the software is built and deployed in a secure manner”.

PRIPARE endeavours to elicit not only security functional requirements but also non-functional and privacy requirements (functional and non-functional).

Severity
CNIL [11] describes severity as “Estimation of the magnitude of potential impacts on the data subjects’ privacy. It essentially depends on the level of identification of the personal data and prejudicial effect of the potential impacts”.

System
According to OASIS’ PMRM reference model [17], a system in the privacy domain is “a collection of components organized to accomplish a specific function or set of functions having a relationship to operational privacy management.”

PRIPARE’s methodology will make use of this definition and will endeavour to identify all the interacting systems or subsystems during the engineering process.

Threat
CNIL [11] describes a threat as a “Typical action used by risk sources that may cause a feared event”. According to Le Métayer, it is also valid to define a threat in terms of vulnerability: “A threat is a potential for an attacker to exercise a vulnerability” [26]. There is sometimes the misconception that threats and risks are synonymous; however CNIL, Le Métayer and other members of the security community clearly differentiates between threats, potential events, and risk, the probability of that event to happen adjusted with the impact that the event would have.
Trust boundaries

According to OWASP in its quick reference guide [99], “Typically a trust boundary constitutes the components of the system under your direct control. All connections and data from systems outside of your direct control, including all clients and systems managed by other parties, should be considered untrusted and be validated at the boundary, before allowing further system interaction”.

This concept is closely linked to the touch points as defined in PMRM reference model [17] and reflected in the privacy terms section 3.1.2.

Vulnerability

In its paper [26], Le Métayer defines vulnerability; PRIPARE adopts his definition and expands it to include the privacy domain. This way vulnerability can be described as “a flaw or weakness in the security and privacy procedures, design or implementation of the product that could be exercised and result in a security or privacy breach”.

3.1.2 Privacy terms

Accountability

Accountability is a much overloaded word with different meanings and interpretations. Its complexity is well documented in existing literature. (e.g. [1], [2] and [3]). At present, the new European Data Protection Regulation Draft [13] includes the concept accountability as one of its new principles and its draft text is aligned with the existing EDPS glossary definition: “Principle intended to ensure that controllers are more generally in control and in the position to ensure and demonstrate compliance with data protection principles in practice. Accountability requires that controllers put in place internal mechanisms and control systems that ensure compliance and provide evidence — such as audit reports — to demonstrate compliance to external stakeholders, including supervisory authorities.” [3]. It is important to highlight that accountability is not only about to put controls in place but also to provide proof of these implemented mechanisms.

It makes sense to use this definition for PRIPARE as it is an ad-hoc definition for privacy that will be common to the EU and will allow discussing accountability within PRIPARE regardless the complexity of the word itself.

Accountability by design

As defined in the accountability principle in the European Data Protection Regulation Draft [13] and in the EDPS glossary [3], accountability not only means the obligation or responsibility to comply with data protection but to demonstrate this compliance. Accountability, as a central principle for data protection, should be included along with privacy in the initial phases of system design and not as an add-on because they are a core part of the system and due to its strong impact on the architecture design.

One foundational principle of PbD is: “Privacy must be embedded into design” [4]. PRIPARE will adopt this principle and will ensure it is applied during the design phases of system development.

Analysis

The Oxford Dictionary [14] defines analysis as “detailed examination of the elements or structure of something”. In the context of systems and software engineering, the analysis phase main aim is to gather the functional and non-functional requirements that can be also split into user and system requirements.
Relating it to PRIPARE’s methodology, there will be an analysis phase what will go deep into the system description trying to detect potential privacy risks.

**Anonymity**
The ISO 29100 [6] defines anonymity as a “characteristic of information that does not permit a personally identifiable information principal to be identified directly or indirectly”. Normally when data is considered anonymous controllers and processors are exempt from applying the principles of protection (e.g. EU Data Protection Directive [5]). A study should be conducted to determine what the de-anonymisation risks are, especially if pseudo identifiers are used. Depending on the context and the availability of data, an attacker can succeed in re-identifying anonymised data.

**Anonymisation**
The ISO 29100 [6] defines anonymisation as the “process by which personally identifiable information (PII) is irreversibly altered in such a way that a PII principal can no longer be identified directly or indirectly, either by the PII controller alone or in collaboration with any other party”. Hence anonymisation is the process of adding the anonymity’s characteristic to personal data.

**By design**
The Oxford Dictionary [14] describes the by design concept as “as a result of a plan; intentionally”. Privacy and security by design means that privacy and security in a system must be intentionally planned and not left as an add-on for the final phases. If privacy or security is not present in the design phases, systems may have to be re-built from scratch in order to become secure and privacy-respectful.

**Consent or Informed consent**
There are various definitions from several sources for consent that match (e.g. [3], [5] and [6]). According to the Data Protection Directive [5] consent and informed consent are synonymous. The EDPS states that “In data protection terminology, consent refers to any freely given, specific and informed indication of the wishes of a data subject, by which he/she agrees to personal data relating to him/her being processed (see Article 2 sub (h) of Data Protection Directive 95/46/EC and Article 2 sub (h) of Regulation (EC) No 45/2001.”. Article 29 Working Party issued a working document [95] providing guidance on how to obtain consent for the use of cookies and similar technologies in compliance with EU legal requirements. This document clearly defines what informed consent should mean in the data protection ambit.

The new EU Data Protection Regulation [13] will tighten the meaning of consent and all consent will be required to be explicit (before, only consent to processing sensitive data needed to be explicit), requiring the data controller to bear the burden of proof for the data subject’s consent.

PRIPARE’s methodology will have to ensure that informed consent is achieved for any personal data processed by the system.

**Context (contextual privacy)**
There are books (e.g. Privacy in context: Technology, Policy and the Integrity of social life [9]) that refer to the relation between privacy and context. For PRIPARE the relevant fact is that privacy is dependent on context. There is no absolute privacy (independent from context). For some contexts such as a social network, privacy requirements are less stringent than in a patient history system in a hospital. In any case, privacy requirements must be elicited and considered during the analysis and design phases of system development.
Control (over personal data)
This term is closely related to the EU Data Protection Directive principle [5] Right of Access which aims to guarantee that the data subject is able to access, rectify, block or erase his data. PRIPARE partners have agreed to design PRIPARE’s methodology user-centric, which will enforce engineered systems to provide users with the necessary controls to manage their data.

Data minimisation or information minimisation
EDPS states that “The principle of data minimisation means that a data controller should limit the collection of personal information to what is directly relevant and necessary to accomplish a specified purpose. They should also retain the data only for as long as is necessary to fulfil that purpose. In other words, data controllers should collect only the personal data they really need, and should keep it only for as long as they need it.” [3]. This is one of the principles to be followed to develop privacy-respectful systems. PRIPARE considers this principle as one of the starting points of the methodology and that it should be followed during the project definition phase to ensure that the minimum personal data is processed. Minimising data collection what will lead to a minimisation of the need of trust in the data controllers and processors, increasing the levels of trust in ICT systems.

Design
According to the Oxford Dictionary [14], design is “a plan or drawing produced to show the look and function or workings of a building, garment, or other object before it is made”. For PRIPARE the relevant context for this term should be System design. The NTIA, of the USA defines Systems design as “A process of defining the hardware and software architecture, components, modules, interfaces, and data for a system to satisfy specified requirements” [12].

Domain Owner
PMRM reference model [17] defines a Domain Owner as “An entity having responsibility for ensuring that privacy controls and privacy constraints are implemented and managed in business processes and technical systems in accordance with policy and requirements”

Identifiable
There is no explicit definition for this term in the related literature but its meaning can be extracted from the Personal Identifiable Information term. Identifiable is the characteristic of data or information that enables it to be used to identify the subject to whom such information relates.

Identity
According to the ISO29100 [6] definitions identity is a “set of attributes which make it possible to identify the personally identifiable information principal”. In PRIPARE’s context, the principal is the data subject.

Legitimacy of processing
The EU Data Protection Directive [5] defines the legitimate purpose principle (art. 6.b) and states that “Personal data can only be processed for specified explicit and legitimate purposes and may not be processed further in a way incompatible with those purposes”. Within PRIPARE this principle will be closely related to the informed consent principle. The users must always be informed of and consent to the purpose of the collection of data.

Lifecycle Data Protection Management
This is a new term introduced in the latest draft of the EU Data Protection Regulation draft [13] adopted by the LIBE Committee (Recital 71a). It is the new title of section 3 - Chapter 4 and
refers to management of personal data in the entire lifecycle of data from collection to deletion. It includes the need to carry out assessments of the impact of the envisaged processing operations on the rights and freedoms of the data subjects.

PRIPARE will support the design and development of systems that follow the privacy and security principles during the whole lifecycle of the data.

**Management**

The Oxford Dictionary [14] defines management as “the process of dealing with or controlling things or people”. Management concept is usually related to other terms like data or data protection (i.e. data management or data protection management).

**Methodology**

The Merriam-Webster [15] dictionary defines a methodology as “A body of methods, rules, and postulates employed by a discipline: a particular procedure or set of procedures”. PRIPARE will provide a methodology that will describe the set of procedures, methods, tools, rules and postulates used to develop privacy-friendly and secure-by-design systems.

**Opt-in and opt-out**

According to the ISO29100 [6] opt-in is a “process or type of policy whereby the personally identifiable information (PII) principal is required to take an action to express explicit, prior consent for their PII to be processed for a particular purpose”. The opposite term, opt-out, is used when consent is implicit and the subject must take an explicit action to withhold or withdraw the consent. According to the PbD foundational second principle [4], “Privacy as the default setting”, opt-in should be the right choice.

**Personal Identifiable Information (PII) or Personal data**

PII and Personal data are synonyms according to the different sources so the choice of which one to use can be considered more a matter of style. While ISO opts for the PII style and names all the related terms accordingly; the EU prefers to talk about personal data (EU Data Protection Directive). The European Data Protection Directive defines personal data as “any information relating to an identified or identifiable natural person (data subject); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;” [5].

Privacy-respecting systems designed with PRIPARE’s methodology must ensure that personal data is kept private and secure according to provisions in the data protection regulatory framework and best available practices. PRIPARE’s consortium has agreed to use the EU wording style and will use terms like personal data, data controller or data subject instead of its ISO equivalents.

**PII controller or data controller**

The EU Data Protection Directive [5] defines a data controller as “the natural or artificial person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; (art. 2 d)”. It can be considered a synonym to the definition of the PII controller term in the ISO29100 [6]. The EDPS [3] completes the definition by stating that “In particular, the controller has the duties of ensuring the quality of data and, in the case of the EU institutions and bodies, of notifying the processing operation to the data protection officer (DPO). In addition, the data controller is also responsible for the security measures protecting the data. The controller is also the person or entity that receives a request from a data subject to exercise his or her rights”.
PRIPARE will adopt the definition of EDPS of a data controller’s as it is more complete.

**PII principal or data subject**
The ISO29100 [6] defines PII principal as a “natural person to whom the personally identifiable information (PII) relates” but states that the synonym “data subject” can be used instead of the PII principal term (the EU Data Protection Directive [5] uses data subject). PRIPARE will use *data subject* to identify the natural person to whom the personal data relates.

**PII processor or data processor or processor**
The PII processor (ISO29100 [6] term) and data processor (EU Data Protection Directive [5] term) are synonyms for the entity who processes personal data according to the data controller’s instructions. PRIPARE will adapt ISO’s definition to EU terminology defining a data processor as a “privacy stakeholder that processes personally data on behalf of and in accordance with the instructions of a data controller”.

**Privacy**
Westin [7] defines privacy as “an individual’s right “to control, edit, manage, and delete information about them[elves] and decide when, how, and to what extent information is communicated to others”. There are several other definitions and privacy is certainly not a universal concept that can be applied across all technologies and all situations. PRIPARE does not pretend to choose one as the “right one”. When thinking about PbD, it is beneficial to keep in mind a taxonomy of privacy that can help to organise and classify privacy issues. Finn, Wright and Friedewald propose a taxonomy of seven types of privacy [39] that can be used in PRIPARE’s PSbD methodology and that is presented in the Privacy Principles section (4.2.1).

**Privacy by architecture and Privacy by policy**
There are two concrete approaches to privacy engineering that can be combined to build privacy-friendly systems. Privacy by policy is an approach that is mainly focused on FTC’s Fair Information Practice Principles (FIPPs) notice and choice. This approach is sometimes considered to fall short as it does not take into account strong attacks like hacking, which would compromise users’ privacy. It is based on the assumption that data controllers and processors can be trusted. A more complete approach, *privacy by architecture*, uses other types of strategies (i.e. data minimisation) to build more privacy supporting systems. Privacy by architecture’s goal is to build systems with non-identifiable users or those, which even when compromised, require a considerable effort to identify the users. PbD combines these two approaches and PRIPARE will make use of them in order to propose architectures that embed privacy controls to produce privacy-friendly and trustworthy systems.

**Privacy Domain**
PMRM reference model [17] defines a Privacy domain as “A physical or logical area within the use case that is subject to the control of a Domain Owner(s)”. PRIPARE will enforce the identification of different privacy domains in the system’s use cases so that personal data, privacy controls and other elements can be harmoniously organized within them.

**Privacy Enhancing Technologies (PET)**
According to the IS029100 [6] definition that is in accordance with that from the EDPS [3] “PET is a privacy control, consisting of information and communication technology (ICT) measures, products, or services that protect privacy by eliminating or reducing personally identifiable information (PII) or by preventing unnecessary and/or undesired processing of PII, all without losing the functionality of the ICT system”
PRIPARE will have to provide the necessary steps to choose, if appropriate, existing PETs and to provide in the design hooks for a foreseeable future where an adequate PET can be linked into the system.

**Privacy patterns and antipatterns**

In 1977 the term *pattern* was coined relating it to the architectural context in the book “A Pattern Language: Towns, Buildings, Construction” [10] as a description of a problem and the solution. This definition has been adapted and expanded for other contexts as software development. Generally a pattern can be considered as “a reusable solution to a commonly occurring problem”. This definition can be contextualised for privacy issues. Part of PRIPARE work is to define a template to describe privacy patterns and to make an effort to collect and classify existing privacy patterns and best practices. Privacy patterns will be complemented with legal and ethical references to give them a non-technical perspective that can be used for non-engineering stakeholders.

As the opposite concept, antipatterns are “common responses to a recurring problems that are usually ineffective and risk being highly counterproductive”. PRIPARE may attempt to identify these privacy antipatterns to guide stakeholders in the right engineering choices.

**Privacy policy**

According to the ISO20100 definition a privacy policy is an “overall intention and direction, rules and commitment, as formally expressed by the personally identifiable information (PII) controller related to the processing of PII in a particular setting” [6].

**Privacy principles**

According to the ISO20100 definition a privacy principle is a “set of shared values governing the privacy protection of personally identifiable information (PII) when processed in information and communication technology systems” [6].

**Privacy preferences**

According to the ISO20100 definition privacy preferences are “specific choices made by a personally identifiable information (PII) principal about how their PII should be processed for a particular purpose” [6].

**Privacy risk assessment**

According to the ISO20100 definition a privacy risk assessment is an “overall process of risk identification, risk analysis and risk evaluation with regard to the processing of personally identifiable information (PII)” [6]. It is actually an important part of a Privacy Impact Assessment.

PIAs are a de-facto standard for designing privacy-respecting systems and given that the draft Regulation on Data Protection requires Data Protection Impact Assessments to be conducted if certain processing operations are likely to present specific risks, PRIPARE will use them as a core part of its methodology.

**Privacy stakeholder**

The ISO20100 defines privacy stakeholders as any “natural or legal person, public authority, agency or any other body that can affect, be affected by, or perceive themselves to be affected by a decision or activity related to personally identifiable information (PII) processing”. At some early stage PRIPARE’s methodology will encourage methodology adopters to communicate with the privacy stakeholders of a system, to elicit privacy requirements and to foresee any privacy issue that may arise at a later date.
Private data
See Personal Identifiable Information (PII) or Personal data.

Profile (profiling)
At present, the new European Data Protection Regulation Draft [13] defines the act of profiling as “any form of automated processing of personal data intended to evaluate certain personal aspects relating to a natural person or to analyse or predict in particular that natural person’s performance at work, economic situation, location, health, personal preferences, reliability or behaviour”.
PRIPARE will accept this definition as it is the one used for the Data Protection Principles that are adopted within PRIPARE.

Proportionality
The EU Data Protection Directive [5] mentions (i.e. article 11) that some rules or principles need not be applied when a disproportionate effort is required. The principle of proportionality in information society has been discussed in the law case Promusicae vs. Telefonica (C-275/06). The principle requires a fair balance to be struck between the various fundamental rights protected by the EU legal order. Additionally, in the general EU doctrine, there are four criteria to verify the proportionality of a measure (proportionality test): there must be a legitimate aim for a measure, (ii) the measure must be suitable to achieve the aim, (iii) the measure must be necessary to achieve the aim, (iv) the measure must be reasonable, considering the competing interests. The European Court of Justice (ECJ) often merges the last two criteria into one.

Pseudonymity
The ISO15408 states that “pseudonymity ensures that a user may use a resource or service without disclosing its identity, but can still be accountable for that use. The user can be accountable by directly being related to a reference (alias) held by the TSF, or by providing an alias that will be used for processing purposes, such as an account number” [19]. Pfitzmann and Hansen define this in a much broader way and state that accountability and disclosure are not necessarily required. Its definition is much simpler: “Pseudonymity is the use of pseudonyms as identifiers” and “A pseudonym is an identifier of a subject other than one of the subject’s real names” [16]. When Pfitzmann and Hansen talk about Pseudonymity they imply that only the holder of the pseudonym is able to link it to their identity.

Quasi-identifier
The Springer’s encyclopedia of Cryptography and Security [20] defines a quasi-identifier as “a set of attributes that, in combination, can be linked with external information to reidentify (or reduce uncertainty about) all or some of the respondents to whom information refers”. There is an example of quasi-identifiers mentioned in the Motwani and Xu’s paper: “Neither gender, birth dates nor postal codes uniquely identify an individual, but the combination of all three is sufficient to identify 87% of individuals in the United States” [21].

Reference model
According to OASIS “A reference model is an abstract framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist. A reference model is not directly tied to any standards, technologies or other concrete implementation details, but it does seek to provide a common semantics that can be used unambiguously across and between different implementations” [17].
While a methodology is more oriented to provide a set of procedures, the reference model tries to unify the concepts to which the methodology will refer. PRIPARE will provide a comprehensive reference model for the privacy environment that will be used within PRIPARE’s methodology.

**Sensitive PII or Sensitive personal data**

According to the ISO20100 [6] definition privacy preferences are a “category of personally identifiable information (PII), either whose nature is sensitive, such as those that relate to the PII principal’s most intimate sphere, or that might have a significant impact on the PII principal”. Current EU Data Protection directive specifies what those categories are (religious beliefs, political opinions, health, sexual orientation, race, and membership of past organisations) and what special measures have to be taken while processing this data. In the new EU Data Protection Regulation new categories like gender identity, biometric data, administrative sanctions, suspected offences or philosophical beliefs will be added.

**Third party**

The ISO20100 defines a third party in the privacy environment as a “privacy stakeholder other than the personally identifiable information (PII) principal, the PII controller and the PII processor, and the natural persons who are authorized to process the data under the direct authority of the PII controller or the PII processor”[6]. PRIPARE will encourage engaging with these stakeholders (i.e. data protection authority), whenever necessary, in the early stages of the system conception so that their concerns are considered during the system design and building.

**Touch point**

OASIS PMRM methodology specifies that data flows carrying PI and privacy constraints among the different privacy domains must be identified. According to OASIS, “the intersection between the data flows and those domains, or systems within those domains” [17], are considered touch points.

**Unlinkability**

The ISO 15408 [19] defines that “unlinkability ensures that a user may make multiple uses of resources or services without others being able to link these uses together. [...] Unlinkability requires that users and/or subjects are unable to determine whether the same user caused certain specific operations in the system.” Pfitzmann and Hansen removes the focus on the user and generalizes talking about Items of Interest (IOIs): “linkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker’s perspective means that within the system (comprising these and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not” [16]. PRIPARE is more aligned with Pfitzmann and Hansen’s definition as it is more general and covers not only unlinkability of users but also of messages, actions...

**Unobservability or Undetectability**

The ISO 15408 [19] defines that “unobservability ensures that a user may use a resource or service without others, especially third parties, being able to observe that the resource or service is being used.” Pfitzmann and Hansen describes unobservability in means of IOIs as “undetectability of an item of interest (IOI) from an attacker’s perspective means that the attacker cannot sufficiently distinguish whether it exists or not” and exemplifies it like “If we consider messages as IOIs, this means that messages are not sufficiently discernible from, e.g., random noise” [16].
User-centricity / User-centred design

This term, also known as human-centred design, is defined by the W3C as “a user interface design process that focuses on usability goals, user characteristics, environment, tasks, and workflow in the design of an interface” [24]. Limiting the user-centricity concept to the user interface of a system is an error as it should be embraced in the whole system, since its inception. The ISO 9241-210:2101 [23] uses a broader definition: “approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques” which is preferred. The same ISO standard notes that “usable systems can provide a number of benefits, including improved productivity, enhanced user well-being, avoidance of stress, increased accessibility and reduced risk of harm”. In PRIPARE’s scope, user centred design will also increase privacy and trust.
4 Principles

To consider a system secure, trustworthy or privacy-respectful, it has to follow a list of common principles on security and privacy. These principles have to be borne in mind throughout the system design and development phases. To ensure this, any methodology used to build trustworthy systems should have those principles deeply rooted in it. While the list of principles can vary from one ICT system to another depending on the industry or domain of the system to which it belongs, a comprehensive common list will be defined within PRIPARE to ensure the common basis of privacy and security for any system is covered in the reference model.

For security principles there is a consensus on the list of eight principles provided by Saltzer and Schroeder in 1975 [27]. However OWASP (an open community dedicated to enabling organizations to conceive, develop, acquire, operate, and maintain applications that can be trusted) presents its own list [31], an extension of Saltzer and Schroeder’s. These principles will help to design and implement systems without security flaws.

Regarding privacy principles there is a more extensive list of available sources. Many organizations and government bodies have provided various lists of principles. Such lists can be explicit like OECD (Organisation for Economic Co-operation and Development) and FTCs (Federal Trade Commission) or be more implicit in directives and regulations like in the EU Data Protection Directive or the to-be EU Data Protection Regulation, currently in draft status. The list of sources used to build PRIPARE’s privacy principles compendium is:

- FTC FIPPs (Fair Information Practice Principles)
- OECD [28] and [29]
- The proposal for a General Data Protection Regulation [13], as voted in LIBE Committee by the European Parliament;

For the selection of principles, main principle sources have been evaluated trying to find principles that are common among them. Principles have been subject to discussion between PRIPARE’s partners until a consensual list was built.

4.1 Security principles and best practices

Security principles must be embedded in PRIPARE methodology and reference model so developed systems have the necessary security controls to be trustworthy. These security principles will be applied during the design of the methodology and within the methodology during all the system development phases.

Defence in depth (Complete mediation)

Layered security mechanisms improve security in systems. If one mechanism fails there are still others that may still provide security. Multiple security mechanisms are able to address different attack vectors increasing the security.

This principle does not relate to any specific control, it is oriented for guidance purposes during the selection of security controls for the system.

PRIPARE methodology will have to ensure that different layers of the system or different subsystems have their own security controls to ensure this principle is followed.
Positive security model (Whitelisting)
There are two antagonist security models: whitelisting and blacklisting. Blacklisting is a negative security model meaning that it only defines what is rejected, everything else is accepted. Whitelisting is the opposite concept, it defines what is accepted and everything else is rejected. The benefit of using the whitelist model is that new attacks and situations, not anticipated during the design and development, will be prevented.
PRIPARE will have to ensure that the authorisation model of the ICT systems follows this principle to avoid security issues.

Fail securely (Fail-safe defaults)
To achieve secure and resilient systems all errors must be handled securely. In case of error the system should behave in the most conservative way, avoiding authorisations that would not happen in the normal course of the system.
PRIPARE’s methodology will ensure that developed systems follow this principle by checking the existence of the most suitable controls enabling this behaviour among the different privacy and security services.

Run with least privilege (Least privilege)
This principle is also known as the Principle of Least Authority. “The principle of least privilege recommends that accounts have the least amount of privilege required to perform their business processes. This encompasses user rights, resource permissions such as CPU limits, memory, network, and file system permissions.”[31] This principle is closely related to the whitelisting principle. To ensure the least amount of privilege is assigned to system users, a good approach is to whitelist those privileges.
Notice that the principle does not apply to users but to accounts, this can include other entities i.e. systems and middleware.
To provide a secure authorisation control in the ICT system, PRIPARE’s methodology must ensure that authorisation components and services follow this principle.

Open design (avoid security by obscurity)
Security by obscurity is the reliance of the security of a component or a system in the secrecy of its implementation. Designing secure controls should rely on open and well-known principles like i.e. cryptography. The open design principle will ensure that instead of assuming that possible vulnerabilities will not be exploited because they are secret, systems will use established and proven standards, protocols, algorithms and methods.
PRIPARE will encourage the use of existing and well-known security controls to provide security for the system based on transparency of their specifications.

Simplicity (Economy of mechanism)
Any system, subsystem or component of a system must be as simple as possible. As the complexity increases it is more difficult to foresee the possible security issues or risks.
PRIPARE’s methodology will provide and encourage the use of design and privacy patterns in order to ensure that systems are simple and easy to understand and that well-known and tested solutions are re-used in the simplest manner to achieve desired goals.

Detect intrusions (compromise recording)
Designing a perfectly secure system is impossible. Therefore to minimise the impact of security breaches, intrusions must be detected so that further, more refined, attacks can be prevented.
Detecting intrusions mostly relies on manual or automated inspections of system logs. Logging all security relevant information is of great importance as it is the only data that can be analysed after an intrusion. By solely logging the information no intrusion will be detected. Forensic tools must be developed or provided and procedures should be established to ensure a regular monitoring of the logs.

PRIPARE’s methodology will endeavour to ensure that logging components and services are layered across the system architecture while encouraging the establishment of the necessary organisational procedures to ensure that the logs are monitored to detect intrusions.

Don’t trust external infrastructure or services
Communicating with external infrastructure or services usually means dealing with systems that cannot be audited, have different security policies and cannot be examined to determine whether they have been compromised or not. These circumstances make it impossible to establish a trust relationship and imply that a special focus has to be placed in communications with such systems. Specific safeguards should be implemented on any data exchange with external infrastructure or services.

PRIPARE will analyse all data flows and will identify those whose sources or destinations are untrustworthy. Specific controls will be required for those data flows.

Establish secure defaults (psychological acceptability)
This principle is closely related to the privacy by default principle that will be presented later. Users should have, by default, the most secure configuration settings activated. However they should be able to, explicitly, lower the security to enhance their user experience.

PRIPARE will provide user-centred privacy and design patterns to ensure that security and user-experience can be combined in a positive sum developing secure and user-friendly systems.

4.2 Privacy principles
Privacy is a complex topic with various definitions and meanings. Protecting such a complex and, at times, abstract notion brings with it a new set of problems. The right to privacy has traditionally been argued as a right to be let alone. However, privacy should be considered in relation to legislative definitions, particular rights accruing in relation to the space or place an individual inhabits, a need to protect the integrity of the body, a specific value placed on the content and accessibility of information, and/or the constitution of a set of boundaries [89]. As Viseu, Clement and Aspinall argue, privacy is a “loose concept encompassing a variety of meanings” [90]. The term is sometimes used to refer to boundaries around, and opposition between, “the public” and “the private” [91]. This can take the form of privacy in certain spaces, such as the home, in relation to medical records, or in relation to shopping habits (such as loyalty cards). Moving away from the concepts of public and private, privacy is also understood in the context of rights and legislation [92]. A variety of possible solutions for protecting privacy have been offered. These range from encouraging organisations to adopt Fair Information Practice Principles (FIPPs) to online privacy protections and other types of Privacy Enhancing Technologies, seals and certificates. However, there are also those who argue that “privacy is dead”, or at least if it is not yet dead that it is impossible to protect in the face of rapid technological developments [94].

Privacy is certainly not a universal concept that can be applied across all technologies and all situations. Finn et al. argue that current attempts to capture the complexities of privacy issues in reactive frameworks are inadequate. They state that: “Rights to privacy, such as those
enshrined in the European Charter of Fundamental Human Rights, require a forward-looking privacy framework that positively outlines the parameters of privacy in order to prevent intrusions, infringements and problems” [39]. Finn et al. build upon Clarke’s influential conceptualisation of privacy and his taxonomy of privacy based on four categories. They suggest that Clarke’s taxonomy is no longer adequate for addressing the range of privacy issues that have arisen with regard to a new and emerging set of systems and technologies. They therefore suggest an approach that encompasses seven types of privacy: privacy of the person, privacy of behaviour and action, privacy of communication, privacy of data and image, privacy of thoughts and feelings, privacy of location and space and privacy of association. Under this taxonomy, privacy of the person is defined as the right to keep body functions and body characteristics private. Privacy of behaviour and action refers to sensitive issues such as political activities and religious practices. Privacy of communication relates to interception of communications such as recording and access to e-mail messages. Privacy of data and image involves the right of the individual to exercise control over personal data, rather than such data being available to organisations and others by default. Privacy of thoughts and feelings refers to the individual’s right to not have to share her thoughts and feelings, or to have these revealed. Privacy of location and space encompasses the right of the individual to freely move about in public or semi-public space, without being monitored or tracked. Privacy of association refers to the right of the individual to associate with others without being monitored.

This approach is beneficial in terms of navigating the various definitions of privacy in the literature to date. Rather than focusing only on personal data and personal communications, as has been the case to date in data protection legislation, the taxonomy proposed ensures that different types of privacy are protected. With regard to the PRIPARE project, it would be beneficial to keep this taxonomy in mind when thinking about PbD. Rather than getting caught up in the myriad and diverse definitions of privacy, basing the PRIPARE methodology on this taxonomy of seven types of privacy will move the debate forward as opposed to reinventing the wheel.

4.2.1 Data protection principles

In the EU, all Member States are signatories to the Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR). The right to privacy is incorporated under Article 8 of the ECHR. At the European level, the Data Protection Directive 95/46/EC [5] regulates the collection and processing of personal data. Currently, most national data protection is based on the following eight EU principles of data protection:

Data must be:
1. Fairly and lawfully processed
2. Processed for limited purposes
3. Adequate, relevant and not excessive
4. Accurate
5. Not kept for longer than necessary
6. Processed in accordance with the individual’s rights
7. Secure
8. Not transferred to countries without protection.
However, this Directive currently only provides guidance, which is then implemented in national data protection regulation across the EU. Data protection law is not currently harmonised across the EU – the Directive is an attempt at harmonisation but in reality legislation at the national level overrides this EU-wide framework.

4.2.1.1 Data Protection Regulation

The current proposal to reform the EU Data Protection Directive includes a proposition to transform this guiding framework into a Regulation, which is directly applicable in the Member States. Furthermore, the current reform highlights the need for privacy protection to be included throughout the entire lifecycle of a product, under Article 23 and the proposal for a PbD requirement. Although the majority of this reform concentrates on the Internet and personal data, there are also important potential changes for the use of video surveillance particularly in terms of digital traces. The reform proposes to reinforce individuals’ rights, to strengthen the EU internal market, to ensure a high level of data protection in all areas, to ensure proper enforcement of the rules and to set global data protection standards. In terms of the key changes, the EU is proposing: a “right to erasure” (Article 17, at some stage it was called the “right to be forgotten”), which concentrates on the individual’s right to decide when they no longer want their data to be processed, which will lead to the deletion of data; expiry dates with regard to the deletion of data; that consent will have to be given explicitly rather than be assumed; that individuals will have the right to refer in all cases to their home national data protection authority, even when their personal data is processed outside their home country; companies and organisations will have to notify serious data breaches (Article 32) without undue delay (and, where feasible, within 72 hours2); a single set of rules on data protection, valid across the EU; increased responsibility and accountability for those processing personal data; a need for privacy protection to be included throughout the entire lifecycle of a product; and finally that national data protection authorities will be strengthened so that they can better enforce the EU rules at home. The proposal also includes the obligation to appoint a Data Protection Officer (DPO) under Article 32a in the case of the processing of sensitive personal data, processing of personal data related to more than 5,000 data subjects in a 12-month period, and/or where processing operations involve the regular and systematic monitoring of data subjects. The Draft Regulation also includes a mandatory Data Protection Impact Assessment (DPIA) for data controllers and processors whenever the process raises specific risks to the rights and freedoms of data subjects. This Draft Regulation also includes some major changes that will have to be considered within PRIPARE’s methodology and reference model:

- Scope modification: the Regulation will be applied in territories outside the EU to those data controllers who aim services at EU citizens, and which monitor EU citizens in any way;
- Profiling: the Regulation introduces a definition for profiling, already reflected in the privacy terms section 3.1.2, and it is expressly permitted for the purposes of fraud monitoring and prevention and to ensure the security and reliability of services provided by controllers. On the other hand, profiling that has a discriminatory effect in sensible categories (e.g. race, ethnic origin and political opinions) is prohibited outright;

2 In the original 25 January 2012 version of the Data Protection Regulation, the EC had specified 24 hours. However, in the version of the Regulation emerging from the European Parliament’s LIBE Committee in October 2013, this was changed to 72 hours.
• Anonymity: the Regulation accepts that, technically, a total anonymity is almost impossible and that it depends in part on the costs and the amount of time required identifying the data subjects to whom the data relate.

The principles of data protection included in the PRIPARE project for discussion come from the European Data Protection Directive 95/46/EC and are extended with those that are new to the Draft Regulation. These principles include: safeguarding personal data, proportionality and data minimization, compliance with the data subject’s right to access and amend their personal data, accountability and the right to deletion. These principles are important in terms of the data lifecycle, from the collection of personal data (and an individual consenting to this collection of their personal data), to processing (and the right to the individual to object to this processing and the principle of proportionality), to deletion of personal data (and the right of the individual to have their data retained only for a set time period and to have their data erased after this time). To date, the project consortium has agreed on the principles listed; however there may still be a need for the PRIPARE project to include a reference to the use of state-of-the-art technologies and the need for engineers to build in new technological solutions to minimize privacy risks. The data protection principles, including issues such as – “what is meant by consent?” – will be further discussed with stakeholders as the project progresses. The new EU Draft Regulation modifies the notion of consent to define it as explicit and informed, rather than implicit. The PRIPARE project will take these new developments into account.

4.2.1.2 List of principles

In terms of defining principles the consortium has agreed to the following definitions:

Principle of data quality
This principle is also known as Safeguarding quality of personal data. Quality of data and transparency are key targets that need to be ensured. Data should be accurate and, where necessary, kept up to date. In order to follow this principle, PRIPARE must ensure that the engineered system provides the necessary privacy controls to enable data subjects to control their personal data. The methodology will emphasize the need for a user-centric approach while enforcing the adoption of this principle.

Principle of data minimisation / proportionality
This principle aims to limit the processing data and ensuring data avoidance and minimisation, processing only adequate and relevant personal data; its use must not be excessive in relation to the purposes for which they are collected and/or further processed.

In relation to the two principles outlined above, the PRIPARE project must keep in mind that the limitation of processing is dependent on context. The terms included in this principle, such as proportionality, adequate and relevant, cannot be defined in concrete terms but are dependent on context. The proportionality term has already been discussed in the terms and definitions section 3.1.2.

The PRIPARE methodology must ensure that data minimisation is taken into account during the analysis phase and at the outset of the project.

Principle of purpose specification and limitation / Finality Principle
Legitimacy of processing personal data must be ensured either by basing data processing on consent, contract, legal obligation, etc. Personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. The purpose of the processing should be defined at the latest at the moment of the collection
of the data. The PRIPARE project must therefore ensure that the purpose of data collection is defined at the outset. Privacy policies are an extended practice that ICT systems apply in order to follow this principle.

**Principle of purpose specification and limitation (specific for sensitive data)**

This principle is also known as *Legitimacy of processing sensitive personal data*. It is a particular case of the *Principle of purpose specification and limitation* where the personal data processed is considered sensitive, according to the applying legislation.

Legitimacy of processing sensitive personal data must be ensured either by basing data processing on explicit consent, a special legal basis, etc. Personal data must be collected for specified, explicit and legitimate purposes and not further processed in any way incompatible with those purposes. The purpose of the processing should be defined at the latest at the moment of the collection of the data. PRIPARE will encourage using stricter controls whenever sensitive data is collected or processed.

**Transparency Principle / Openness Principle**

This principle ensures compliance with the data subject’s right to be informed. It must be ensured that the data subject is informed about the collection of his data in a timely manner (e.g. providing information about: identity of the controller, purpose of processing, recipients of the data, etc.). Data must be processed fairly and lawfully (respecting the applicable national legislation as well as the rights and freedoms of individuals). This principle also protects individuals against unlawful or arbitrary discrimination. Normally this is addressed by providing a description of the data processing activities required for service delivery, ensuring internal and external transparency. Transparency of automated decisions vis-à-vis individuals must be ensured especially. An envisioned privacy instrument to enforce systems to follow this principle is *data transaction logs*, accessible by data subject and that reflect the personal data transactions (i.e. whenever personal data is sent to a processor).

**Right of Access**

Compliance with the data subject’s right of access, rectification, erasure or blocking of data: It must be ensured that the data subject’s wish to access, rectify, erase and block her data is fulfilled in a timely manner and as appropriate when the processing of such data does not comply with the provisions of the European Data Protection Directive in particular because of the incomplete or inaccurate nature of the data. Access right includes facilitating the provision of information about processed data and purpose. This principle is closely related to the *data quality* principle. To ensure the quality of the data it is imperative that the data subject is allowed to access it.

**Right to Object**

Compliance with the data subject’s right to object: Facilitating the objection to the processing of personal data, direct marketing activities and disclosure of data to third parties. Facilitating the objection to being subject to decisions that are solely based on automated processing of data. Just like the *right of access*, this principle is firmly related to the *data quality* principle, and the three of them can be followed making use of user-centric approaches, which will be encouraged by PRIPARE.

**Safeguarding confidentiality and security of processing**

Preventing unauthorised access, logging of data processing, network and transport security and preventing accidental loss of data are key targets that need to be ensured. This principle clearly reflects the relationship of privacy and security. A system cannot be considered private if it is
not also secure. PRIPARE also adopts security best practices and principles that will be embedded into the engineered ICT systems.

**Compliance with notification requirements**
Obligation to notify the supervisory authority: Notification about data processing, prior compliance checking and documentation are the key targets that need to be ensured.

**Conservation or Retention principle**
Compliance with data retention requirement: Retention of data should be for the minimum period of time consistent with the purpose of the retention or other legal requirements. Personal data must be kept in a form that permits identification of data subjects for no longer than what is necessary for the purposes for which the data were collected, or for which they are further processed.

**Accountability**
Accountability can be described as a demonstrable acknowledgement and assumption of responsibility for having in place appropriate policies and procedures, and promotion of good practices that include correction and remediation for failures and misconduct. It encompasses expectations that organizations ---will report, explain and be answerable--- for the consequences of decisions about the protection of data. The principle of accountability is of particular importance for the PRIPARE project. Under the draft Regulation organisations will be expected to demonstrate compliance and to be able to show evidence of this. The PRIPARE project will therefore provide best practice examples of how or what services should be included in the design and implementation phases.

**Right to erasure**
Require the data controller to take all reasonable steps to have individuals’ data erased, including by third parties *without delay*, for the personal data that the controller has made public *without legal justification*. The PRIPARE methodology will ensure ICT systems include an option for the deletion of data upon request by an individual. It will also ensure that the necessary technical and organisational measures are available to forward the deletion request to the affected stakeholders in order for them to fulfil the request.

**Privacy by design / Data protection by design**
Requires data protection to be embedded within the entire lifecycle of the technology, from very early design stage, right through to its ultimate deployment, use and final disposal (Recital 61 Draft regulation). Adopting this principle as part of a PbD methodology can seem redundant, but is also can be considered as the central principle for PRIPARE. The methodology is designed to be a systematic and engineering approach to PbD.

The PRIPARE methodology must ensure that privacy is embedded throughout the lifecycle of the ICT systems and of personal data.

**Privacy by default / Data protection by default**
Requires users'/data subjects’ control on the distribution of their personal data. This translates to explicit consent each time personal data processing is intended. The PRIPARE project must ensure that the issue of explicit consent is embedded in the methodology from the outset and that the privacy preferences are automatically set to its most privacy-preserving configuration.
5 Common grounds necessary to engineer truly trustworthy systems

PRIPARE’s PSbD methodology aims to be holistic. This means that:

- It can be applied to systems or subsystems that compose it, even if being designed separately. For instance an application can rely on solutions and services provided by an underlying technology platform (e.g. a smart phone operating system or a cloud service). The methodology must be applied separately to the underlying technology subsystems and to the application subsystem. A well-known problem is the Internet Protocol (IP) which was not designed with the proper requirements for privacy preservation. Applying PRIPARE methodology would have ensured that anonymisation capabilities would be an integral part of the design of the IP subsystem. Another key subsystem is the set of policies for privacy and security (user stated policies, security policies) and their enforcement. Two identical systems in terms of hardware and software could have totally different privacy protection depending on policies. Therefore PRIPARE methodology must also apply to the design of policies.

- It takes into account the fact that engineering process follows domain specific standards. For instance the IEC 61508 standard covers the complete life cycle in many different domains: automotive systems, railways systems, process industries and nuclear power plants. PRIPARE methodology must be adaptable to the specific aspects of each variant.

- It must also care for the various scales of systems, from the small application developed in a few person-months (e.g. a smart phone application) to huge applications necessitating hundreds of person-months (e.g. a complex e-commerce website).

PRIPARE will propose a simple typology of projects based on the positioning of the focus:

- Projects focusing on application features: being application feature as something useful to the customer, for instance as smart meter application
- Projects focusing on technology or platform features being platform features something useful to the application designer, for instance an operating system, a protocol system

PRIPARE has examined different state-of-the-art practices and assessments in order to determine the necessity, desirability or viability of incorporating them into the methodology.

5.1 Ethical Impact Assessments (EIAs)

New information and communication technologies not only raise privacy but also wider ethical concerns. Wright and Mordini propose conducting an Ethical Impact Assessment (EIA), to complement a privacy impact assessment (PIA). They suggest that these two forms of impact assessment could be carried out concurrently [88]. Similarly to a PIA, an EIA can ensure that any potential ethical implications of a system are assessed and examined by stakeholders during the development phase. Wright and Mordini point out that considering ethics in this context is not new and point out that Moor, in his 1985 essay, raised concerns about computer ethics, the nature of technology and societal and ethical values. Wright and Mordini propose a framework for an EIA based on various sources. In relation to values, they draw on the values stated in the EU Reform Treaty, such as human dignity, freedom, democracy, human right protection,
tolerance and justice, amongst others [88]. They point out that these values are also stated in the Charter of Fundamental Human Rights of the European Union. In terms of the role of ethics with regard to information and communication technologies, Wright and Mordini state that “the crucial ethical, privacy-related issue seems to be the one of control, the risk individuals bear as they automatically lose control of their personal information. Individuals’ awareness and genuine consent is lacking in many circumstances”. They suggest that “an ethical inquiry into these processes and ethical assessment of technologies can help to clarify such issues in a rational way” [88]. PRIPARE has carefully considered including ethical aspects in the methodology and has decided that:

- Engineering ethical principles if a more complex task, if possible, than embedding privacy ones;
- An EIA can be seamlessly integrated with PIA processes.
- Those ethical aspects that may affect privacy are already embedded in the privacy principles stated in the EU DPD and Draft Regulation;

Anyhow, the methodology will give hints of where it can be extended to include EIAs and PRIPARE will endeavour to connect collected privacy patterns (part of PRIPARE’s tasks) with relevant ethical aspects.

5.2 Best practices on PIAs

Although PbD is often cited by regulators as best practice in relation to privacy and data protection, there is an obvious gap between the abstract principles of PbD and more concrete processes of operationalising those principles. There is a lack of guidance provided to industry and/or engineers in terms of how to think about PbD principles in practice. A step forward in operationalising PbD is to conduct a privacy impact assessment. A Privacy Impact Assessment (PIA) can help to identify privacy risks. Identifying these risks can highlight areas where PbD principles can be applied and embedded. PbD should therefore use Privacy Impact Assessments (PIAs) as a basic element.

A recent PIA framework developed for RFID has been cited as being a “landmark PbD document” [80]. The framework is the first of its kind in being sector-specific and developed by industry. It provides guidelines on how to process data specifically related to RFID applications, and how to assess privacy and data protection issues through PIAs. Spiekermann states that the framework “suggests concrete privacy goals and describes a method to reach them” [81]. In order to be effective, PIAs need to move beyond legal compliance checks in order to “offer a prospective identification of privacy risks before systems and programmes are put in place” and that they “have to consider privacy risks in a wider framework which takes into account the broader set of community values and expectations about privacy” [82].

PIAs should not be considered as simply legal compliance checks, which ask: “If we did X, would we be in compliance with the law and the fair information principles upon which the law is based?” [82] Nor should they be considered to be privacy audits, used to assess existing technologies, although, as Wright argues, a PIA can enable an organisation to demonstrate compliance with legislation in the case of a privacy audit or complaint. According to David Wright, undertaking a PIA can “provide evidence that the organisation acted appropriately in attempting to prevent the occurrence. This can help to reduce or even eliminate any liability,
negative publicity and loss of reputation.”[83] A 2007 Linden Consulting report for the ICO states that they are most useful for new programmes, services or technologies. However, they are not simply used to warn against potential risks but also to mitigate these risks, and to change the development process accordingly. PIAs, therefore, move beyond the legal compliance to assess and address the “moral and ethical issues posed by whatever is being proposed” [84]. The Ontario Data Protection guidance states that the “cyclical nature of the information lifecycle must be supported by appropriate policies, practices, procedures, tools and contracts”. With reference to this lifecycle of information, the guidance states that “risk must be properly identified, minimised to the extent possible, and appropriately managed where it can't be eliminated” and “a proper contemplation of the information lifecycle includes these concepts”. A privacy impact assessment is one of the ways that the information lifecycle can be managed and privacy risks minimised [85].

PIA principles should be reviewed in the context of the system or technology being developed. Conducting a PIA in order to fulfil the requirements of PbD includes thinking about issues of: accountability, data minimisation (including minimising linkability, observability and identifiability), transparency and compliance in the context of risks identified by the PIA process.

Wright suggests that there is currently a “growing interest in Europe in privacy impact assessment” [83]. The UK introduced the first PIA methodology in 2007, although PIAs have been used in Australia, Canada, New Zealand and the United States since the mid-1990s. Conducting a PIA is now mandatory for government agencies in the UK, Canada and the US. It has been found that “unless they are mandatory, many organisations may not undertake them even though their projects, technologies or services have serious privacy impacts”. The EU specifies in its Draft Regulation a mandatory DPIA prior to the data processing whenever “processing operations are likely to present specific risks to the rights and freedoms of data subjects by virtue of their nature, their scope or their purposes”[13]. In terms of best practice, Wright concludes that a PIA process should include:

- An assessment of privacy risks an organisation might face in relation to a new project (although he cautions that a PIA on its own will not highlight all privacy risks and/or issues associated with a new project)
- A process of engaging stakeholders
- Examples of specific risks;
- Recommendations and an action plan
- Third party reviews;
- Benchmarks that organisations could use to test how well they are following the process;
- Publication of the PIA report.
- PIA updates if there are changes in the project.

This is followed up with the recommendation that a third party review and/or audit of an organisation’s PIA be conducted as “it is all too easy for project proponents to say initially that they accept and will implement suggested changes, only to find reasons later to back-slide, and either partially or wholly abandon their initial commitment” [83]. In terms of best practice, Wright also suggests that, in addition to a third party review, accountability mechanisms, such as mandatory reporting requirements, should be implemented. Finally, Wright argues that tying
PIAs to budget submissions for new projects and programmes can ensure a greater number of PIAs are actually undertaken, as well as enhancing accountability.

PRIPARE will embrace and incorporate this view in PIAs to its procedure and reference model approaches while it will also attempt to improve them with other recommendations from projects like the EC-funded PIAF project.

The steps proposed by the EC-funded PIAF project [38] and ISO/IEC WD 29134 [86] for conducting a PIA are as follows:

1. Determine whether a PIA is necessary
2. Identify the PIA team and set the terms of reference, resources and time frame
3. Prepare a PIA plan
4. Determine the budget for the PIA
5. Describe the proposed project to be assessed
6. Identify stakeholders
7. Describe the information flows and other privacy impacts
8. Consult with stakeholders
9. Check the project complies with legislation
10. Identify risks and possible solutions
11. Formulate recommendations
12. Prepare and publish the report
13. Implement the recommendations
14. Third-party review and/or audit of PIA
15. Update the PIA if there are changes in the project
16. Embed privacy awareness throughout the organisation and ensure accountability.

5.3 Risk management process

The first phase of a risk management process is a risk analysis aimed at assessing the situation in order to prepare the subsequent decision making step. The decision may typically consist in the implementation of additional technical or organizational countermeasures to address the risks identified during the security analysis phase. Risk management process is sometimes part of PIAs. The decision making task itself should involve all the stakeholders identified in the PIA to ensure multiple points of view are considered. In keeping with the principles of PbD, the risk management process should start before the design of a product to ensure that privacy issues are taken into account from the outset (rather than considered as an afterthought, to mitigate bad design choices). The risk analysis does not have to be confined to the design phase though. On the opposite, all best practices guides stress that risk analysis and risk management should be continuous processes, which should be applied, as part of a continuous improvement procedure, at every stage of the lifecycle of a product.

Definition of the main technical terms has already been provided in the Terms and definition section (3.1.1) but to illustrate these definitions; let us consider the information management system of a large hospital:
The most valuable assets are the patients’ health records. The privacy policy should, inter alia, protect them against disclosure to any unauthorized staff or third party.

A privacy breach would be the disclosure of some health records to an unauthorized staff or third party.

A vulnerability could be a lack of buffer overflow control in the implementation of the access control component of the information system.

A threat could be the potential for exploitation of the vulnerability by a hacker to get root level privileges on the machine containing the health records.

An attack could be the manifestation of the threat, followed by a copy of the health records of all patients and their publication on the Internet.

A countermeasure could be a set of additional controls to avoid buffer overflows.

The risk could be qualified as high both for patients and for the hospital because the impact on its image would be catastrophic; obviously the trust of the patients in the hospital would be undermined.

A variety of risk analysis methods have been defined for security, for example commercial methods such as FRAAP [47], STRIDE [48] and [49], ASTRA [26] and Cigital [50], international standards such as ISO ([19], [51], [52], [6], [53]) and national standards such as SEI (Software Engineering Institute - Carnegie Mellon University), NIST (National Institute of Standards and Technology) [54] or EBIOS [55]. Each of these methods has its own specificities, scope and emphasis. Some of the methods are general in their purpose (e.g. covering company wide security analysis as well as product analysis) when others are more focused (e.g. dedicated to product assessment); some of them provide general guidance when others are more constraining; some of them stress organizational issues when others are more technically oriented.

### 5.3.1 Common notions and principles

Beyond their differences of scope, focus and vocabulary, all methods are based on common notions and principles. We first sketch these principles before discussing the specific features of privacy and their impact on the risk analysis process.

First, threats are generally associated with a number of attributes, such as:

- The attacker or the threat agent. Attackers can be insiders or outsiders; they can be individuals or organizations;
- The assets concerned by the threat;
- The motivation of the threat agent;
- The resources of the attacker (e.g. expertise, material and funding);
- The actions perpetrated by the attacker.

Several techniques can be used to identify vulnerabilities but they can be classified into two main categories: **checklist based methods** and **testing methods**. These two approaches are complementary but their applicability depends on the lifecycle phase of the product: testing methods, which rely on the existence of the code, can obviously not be applied on a system which has not yet been developed. Generally speaking, building checklists is a very effective way to capitalize on past experience and reduce the dependency of an organization with...
respect to a small group of experts. Checklists can also exploit public vulnerability databases, lists published by computer incident response teams or any other industry sources.

The role of countermeasures is to minimize the likelihood of an attacker exercising a vulnerability. Most methods are based on catalogues of countermeasures which are supposed to be available to the security designer. These countermeasures are the results of decades of research in computer security and their presentation is much beyond the scope of this document. In addition, the adaptation of security analysis methods to privacy requires the consideration of new countermeasures generally called PETs (Privacy Enhancing Technologies).

The attack analysis can be seen as the convergence point where the threats, vulnerabilities and countermeasures identified in the first three steps are combined to form the overall picture of the security of the system. One of the most natural ways to represent attacks in a systematic way is to resort to attack trees. Attack trees can be seen as a security variant of the fault trees that have been used for decades in reliability analysis.

Assuming that attacks have been identified, the next step consists in assessing the risks associated with these attacks to be in a position to take appropriate decisions. Whereas the previous steps were essentially technical, risks have to do with the consequences of the attacks and their evaluation necessarily involves business related considerations such as, typically, the direct and indirect damages to the company’s assets, to its customers’ assets or to its whole business. The impact of an attack can result from the disclosure of confidential or personal data, from the destruction or inappropriate modification of data, from denial of service, etc. Two parameters are generally used to evaluate risks: the probability of successful attacks and their impact. The two parameters can be combined in different ways to establish a final estimation of the risk.

Risk assessment can be qualitative in the sense that it provides a relative evaluation of risks on a fixed and limited scale (e.g. three possible values for two parameters in the above example, such as low, average and high probability of successful attacks). The only possible use of the result of a qualitative analysis is to compare risks and define priorities. In contrast, quantitative risk assessments aim at measuring risks in terms of a numerical variable (usually money but sometimes also in potential loss of life). A typical way to get a quantitative measure of risk is to define it as the product of the estimated value of the threatened assets by the expected probability or frequency (e.g. annual rate) of successful attacks. The result is then an estimation of potential (direct) damages that can be balanced with the costs of potential countermeasures. The output of a quantitative risk assessment can thus be used directly by the decision maker to evaluate the return on investment for a given countermeasure.

The objective of all the previous security analysis steps was to prepare the final one, namely the decision making step. The goal of the decision maker is to strike the best economic balance between the level of risk and the costs of potential technical, business and legal countermeasures. Basically, there are only four main options available to the decision maker depicted in the next figure.
However, since these methods focus on security, they need to be adapted to deal with privacy. As regards risk analysis, privacy differs from security in several ways: first, technically speaking, privacy gives rise to specific risks, such as de-anonymization (or re-identification) through statistical or logical inference, and can be supported by specific tools (PETs). Also, the decision making process should not be confined to the company deploying the IT system: it should involve all stakeholders, including representatives of the subjects whose privacy could be jeopardized by the system. So the transposition of security risk analysis to privacy analysis is not straightforward and deserves careful consideration. An example of effort in this direction is the risk analysis guidelines proposed by the CNIL [11].

5.3.2 CNIL

CNIL has proposed a risk a methodology for risk management that specifically addresses privacy risks [11] which is based on the EBIOS security risk analysis method. The CNIL methodology is based on 5 phases:
The context includes the perimeter of the system, hardware and software to be protected, applicable regulation and privacy policies, and sources of risks. The events to be avoided are related to the assets to be protected (primary assets such as personal data and supporting assets such as storage hardware or software). Threats can be purely technical (spying, eavesdropping, retrieval of deleted information, etc.) or also involve organizational or social aspects (phishing, social engineering, bribery, etc.). The decision making step consists in choosing the appropriate way to address each type of risk. Risks with a high severity and likelihood must be avoided or reduced by implementing security measures that reduce both their severity and their likelihood. At the other end of the spectrum, the decision can be taken to accept risks with low severity and likelihood. In any case, it is of prime importance to provide justifications of the decisions based on the risk analysis. The CNIL also provides a catalogue of good practices to address a list of typical privacy threats [56].

To conclude this subsection, we would like to emphasize some key qualities that any privacy risk analysis should possess to improve the PbD approach, namely rigour, generality and incrementality, already defined in the security and risk management terminology section 3.1.1.

5.4 User empowerment

Once a piece of information has been introduced into a system and linked to a user identity, it must be properly managed by allowing users to participate in aspects that are relevant to their privacy interests. Identity management refers to the set of processes that administers the collection, authentication, use, and deletion of an identity, and the data linked to it (identity attributes), within an organization and across its boundaries. It has evolved from silo-like approaches, where all the identity information is kept and used within a single organization, to federated, or network-centric, approaches where the underlying infrastructure enables a participating entity (the identity provider) to share their users’ personal information with other service providers (the relying parties).
In this context, the term user-centricity broadly refers to the empowerment of the users of identity management systems. It implies providing users with a better control over the lifecycle of their personal information (Lifecycle Data Protection Management) while considering Human-Computer Interaction (HCI) aspects [75].

5.4.1 Identity management

Identity Management Systems (IMS) are ICT systems or technologies that were developed in order to increase security and productivity while decreasing cost. Today, these electronic IMS allow users not only to manage their identity but also their privacy. For example, URL-based systems such as OpenID allow users to choose the entity storing their personal information, OAuth enables users to decide what pieces of information to share, Kantara User Managed Access (UMA) lets an individual control the authorization of data sharing and service access made between online services on the individual’s behalf, and card-based systems further allow users to store the pieces of information to be shared.

The state of the art of user-centric identity management realizations provides users with different degree of control of transactions involving their data and the data itself, usually following a user in the middle approach: Either the identity data or the authorization to gain access to it flows through the user’s client. This feature has led to present user-centric systems in opposition to previous, network-centric identity management solutions.

For an identity management system to be truly user-centric it must adhere to the data minimization principle, disclosing as little personal identifiable information as possible e.g. for some transactions it might be enough to provide proof that a user is under 18 instead of revealing their actual age. Identity management solutions based on anonymous credentials [78] and zero-knowledge proofs [79] enable users to provide authenticated anonymous attributes without identification i.e. attribute-based credentials. Microsoft’s U-Prove, IBM’s Idemix, and the architecture proposed by the European project ABC4Trust are examples of technologies enabling this approach.

Hence, for an identity management solution to be user-centric, it must abide by some constraints and requirements. These requirements translate into specific user interaction features which implement them. Some key points to consider when defining these features are, according to Fischer-Hübner et al [76]:

- How to map the legal requirements (e.g. informed consent) to user interaction patterns (e.g. comprehension, consciousness, control, consent) and user interface designs (e.g. presenting privacy notice and gathering informed consent).
- User-friendly representation of complex concepts such as privacy policies and preferences.

Several approaches have been proposed to address them, including:

- Privacy-protecting defaults for privacy policies and preferences, which make things easier for the end user.
- Simplified customization of defaults, which allows for enhanced and finer-grain control.
• Solutions for getting truly informed consent such as Just-In-Time-Click-Through Agreement, multi-layered format, two-clicks, privacy policy as XML so that it is processed and displayed by clients, menu-based selection, etc.
• Development of privacy metaphors easier to understand for individuals such as privacy icons or role-based paradigms.

Data protection principles recognized by privacy regulation reinforce the idea of user-centricity. For example, the European Data Protection Directive provides users with several Rights regarding their personal information which dictate that data subjects must be able to control (e.g. access, rectify, erase, block and object) different stages of the lifecycle of their data. The Transparency Principle further dictates that the data subject has the right to be informed about the collection of their data in a timely manner, which is usually translated to providing notice and obtaining (informed) consent. HCI aspects must be considered to provide adequate notice that users are able to understand so as to provide an informed consent.

5.4.2 User-centricity
As previously described, user-centric aspects are central to privacy, and so they must be to the PRIPARE methodology.

User-centricity is indeed a broader concept originating in the field of HCI. In this sense, the concept of user-centred design (UCD) is described as a design process that pays attention to the user needs throughout the process, testing the designers’ assumptions with actual users. In this sense, and in the context of identity management, a solution is deemed to be user centric because not only it empowers the user in managing their identity, but first and foremost it responds to the user’s true needs and wants for that management.

The ISO 9241-210:2010 standard describes 6 key principles that will ensure a design is user-centred [23]:
1. The design is based upon an explicit understanding of users, tasks and environments.
2. Users are involved throughout design and development.
3. The design is driven and refined by user-centred evaluation.
4. The process is iterative.
5. The design addresses the whole user experience.
6. The design team includes multidisciplinary skills and perspectives.

These principles are aligned with some other steps of a privacy-by-design methodology. For example, current practices regarding Privacy Impact Assessment require users to be involved as system stakeholders for further consultation.

Furthermore, for a design to be user-centric, users should be involved in several stages of the process. Thus the PRIPARE methodology must involve users throughout the design process. In this regard, different UCD strategies exist that can be applied to different scenarios. Some imply a deep user involvement (e.g. cooperative, participatory or contextual design methodologies), while others reuse existent knowledge condensed in the form of process and product guidelines –notwithstanding, they always take actual users in real world scenarios into account.
Likewise, these principles can be applied to different process models, but they are always substantiated in a particular way of dealing with specific activities. For instance, analysis and specification start dealing with the context of use: who will use the product, what for and under which conditions (platform, environment, preferences, etc.). On the other hand, validation includes user evaluation through usability testing.

It has been made clear that one of most important elements to achieve user empowerment are identity management systems, and that these must be user-centric designed for them to ensure they help to reach acceptable privacy levels. On the other hand, it is important to remark that user-centricty does not only focus or can be applied in ICT systems. Privacy policies and preferences, and even regulations and directives must revolve around the user to provide him with the necessary power to manage their own privacy.

5.5 Privacy instruments

During the last decades a wide variety of technologies and tools have been proposed to improve privacy protection ([57], [58]). Their functionalities can be classified into four main categories:

- Information hiding (e.g. anonymisation, encryption, etc.);
- Information management (subject privacy policies, user interfaces, etc.);
- Transparency (dashboards, controller privacy policies) and
- Accountability (traceability, log management, etc.).

Most efforts of the PETs (Privacy Enhancing Technologies) community have actually focused on the first category and a wide array of techniques have been proposed to reduce the disclosure of personal data (following the data minimization principle) [60], [97], [57]. These techniques include secure multi-party computation, commitments, private information retrieval (PIR), or trusted platform modules (TPM), etc. These PETs have been used to provide strong privacy guarantees in a variety of contexts such as smart metering, electronic traffic pricing, ubiquitous computing or location based services [61], [62]. One of the problems of engineering PbD is the lack of a process for selecting PETs, other than the sometimes relative scarcity of them.

However, the scope of PbD extends beyond the use of privacy enhancing tools. It is also necessary to have appropriate methodologies for PbD, defining precisely the issues to be considered, the design steps, the actors involved and assessment criteria. Several authors (Gürses, Troncoso and Diaz [59], Hoepman [63], Kerschbaum [64], Mulligan and King [65]) have already pointed out the complexity of engineering PbD as well as the richness of the design space of data minimization [59], calling for the development of more general and systematic methodologies for PbD. A step in this direction is described in ISTPA’s Privacy Framework [66] which uses a set of practices necessary to support privacy principles (the fair information practices) and defines a group of “privacy services” (such as audit, control, enforcement, negotiation, validation, access, etc.) to be parameterized and integrated into a privacy framework. In the same spirit, Rakesh Agrawal et al. [67] provide guidelines for the design of databases complying with general privacy principles (hippocratic databases). Another proposal, Hoepman’s, relies on the use of design patterns to define eight privacy strategies called respectively: “Minimise, Hide, Separate, Aggregate, Inform, Control, Enforce and Demonstrate” [63]. PRIPARE will define a privacy design pattern template and will collect and classify a comprehensive list of them to help engineers to select the most appropriate PETs or to apply privacy best practices. As far as privacy mechanisms are concerned, Kerschbaum [64] points out
the complexity of their implementation and the large number of options that designers have to face. To address this issue and favour the adoption of these tools, Kerschbaum proposes a number of guidelines for the design of compilers for secure computation and zero-knowledge proofs. In a different context (designing information systems for the cloud), Machanavajjhala et al. [68] also proposes implementation techniques to make it easier for developers to take into account privacy and security requirements.

A complementary and essential aspect is the evaluation of the level of privacy protection provided by a solution. Several privacy metrics such as k-anonymity [69], l-diversity [68] or differential privacy [70], McSherry and Talwar [71] have also been proposed as ways to fulfil this aim. Differential privacy provides strong privacy guarantees independently of the background knowledge of the adversary: the main idea behind differential privacy is that the presence (or the absence) of an item in a database (such as the record of a particular individual) should not change in a significant way the probability of obtaining a certain answer for a given query. Several methods (e.g. [71], [72] and [73]) have been proposed to design algorithms achieving these privacy metrics or to verify that a system achieves a given level of privacy [74].

5.6 Architectural selection

There has been no successful approach to engineering PbD that has tackled the problem of architecting privacy. Most promising approach, PMRM, in its methodology comes only to the analysis stage, enabling mapping privacy controls to privacy services, while it misses to help in one of the most delicate phases, the design. PRIPARE endeavours to fill this gap by proposing a series of processes that will try to use the privacy requirements (or qualities) in order to identify the architectural patterns and tactics that will guide the selection of the suitable architecture of the system which is crucial for security and privacy. The set of identified tactics and patterns will be the basis for the development of truly trustworthy systems.
6 The challenges of integrating into existing methodologies

Any organisation or institutional body has its own methodologies in place for ICT systems engineering. These methodologies can have been explicitly and consciously deployed or they may have emerged spontaneously. In any case, the working process of almost any organisation can be reduced to a list of steps that they take during the system or project development. Organisations choose their methodology based on their industry, organisation structure, project and/or customer typology.

Neither methodologies nor organisations are static. They are constantly evolving to self-adapt themselves until they achieve an optimal point where they work together in a successful way.

In some cases it is difficult to impose a new methodology on organisations that have been successful during years by promising more privacy-respectful or secure systems or which, on the contrary, are not convinced of the importance of taking a privacy-preserving approach. In other cases it is simply impossible to modify the current methodology for standard reasons; some domains have some constraints in particular systems with safety concerns and they must be compliant with specific process. Hence, the best way for adoption is to create a PSbD methodology that can complement those already in existence in an unobtrusive way.

6.1 Software and system engineering methodologies

There are as many system engineering methodologies as organisations practicing system engineering. Each organisation has its own methodology with its own peculiarities but almost all can be classified into four big families that comprehend virtually any system development lifecycle.

6.1.1 Waterfall

The concept of the waterfall model was first formally described in Royce’s 1970 article [32] but did not use the term waterfall. This article was to show that a commonly used software practice was flawed.

Note that even Royce’ referred its paper to computer program development, this process can be generalised to any system development.

Royce defined a common set of steps to all computer development process:

![Figure 3: Waterfall methodology phases](image-url)
The waterfall model maintains that one should not move to the next phase until the preceding one is completed and perfected. The waterfall model has some benefits and some drawbacks:

- **Benefits:**
  - Focuses on documentation and source code;
  - Simple approach;
- **Drawbacks:**
  - Customers cannot change requirements after the first phase;
  - Some design problems can only be found during the implementation stage;
  - Project outcome can only be seen at verification stage while other methodologies can provide some outcome at early stages;

PRIPARE’s methodology will have to provide a linear alternative with sequential steps so it can match this methodology. Steps should be grouped into equivalent stages in the PSbD methodology for a seamlessly methodology integration.

### 6.1.2 Iterative or incremental

Given the limitations of the waterfall model some modifications were made to make a more usable and realistic model.

![Figure 4: Iterative methodology phases](image)

The iterative model basically has the same steps as the waterfall model but they are repeated during multiple iterations. The idea behind this method is to build smaller portions of the system and using feedback from the process to start a new iteration. The main difference between iterative and incremental models is that while iterative continuously refines all the outputs e.g. system design, system implementation, the incremental focuses each iteration in specific components or systems that are individually delivered.

In the iterative process, for example, coding problems discovered during the implementation phase can be used to improve the design in the next iteration. It also can be used to get user feedback and use it to create new requirements.

The iterative model has some benefits that the waterfall model lacks:
- Intermediate delivery steps, enabling the measure of the alignment with the project goals;
- Supports refactoring, enabling better designs;
- Uses developer and user’s feedback;
As for the waterfall model, PRIPARE’s methodology steps should be grouped into stages that should match the iterative process ones. Special consideration can be made for the initial planning stage that can be used to allocate project privacy related resources or to engage with project stakeholders to elicit specific privacy and security requirements.

6.1.3 Prototype

The original purpose of this model is to evaluate non-final proposals for the design of the system by trying them out rather than interpreting its descriptions. It is also useful to present prototypes to end users so they can test them and input their feedback before the production is developed.

The prototype model has some benefits that the waterfall and iterative models lack:

- Reduced time and costs; by providing end users with prototypes, the quality of elicited requirements will improve and minimise changes during implementation where the costs grow exponentially;
- Improved user involvement; engaging the customer or the end users during the prototyping phases allows them to provide more complete specifications that help to build a satisfactory product;
Although some of the stages of the methodology match the ones defined in the waterfall or iterative models, PRIPARE will have to tackle the specific prototyping stages and define the necessary or recommended steps to follow during that phase in order to develop secure and privacy-friendly systems.

6.1.4 Agile

In 2001 a group of software developers gathered to discuss lightweight development methods in reaction against other software development methods that were more heavy-weighted, harder to follow and flawed with several handicaps (e.g. inability to handle changes, higher cost and burden of documentation). As a result of the discussion they published the Manifesto for Agile Software Development that includes twelve principles [36]. This agile manifesto for software development can and has been adapted to a more generic approach that can be used as a project management methodology. Many methodologies follow the agile manifesto (e.g. scrum, Kanban and extreme programming).

Even if it is the case that Scrum is only one of multiple agile methodologies, it is the most extended. Given the similitudes between agile methodologies i.e. Scrum with Kanban, by defining how PRIPARE’s PbD methodology complements Scrum, it is expected to set the basis to cover most of other agile methodologies.

Scrum steps are can be simplified as:

- Customer builds the product backlog by describing user stories that includes all the features he wants in the system; each user story is accompanied with a value that measures the importance of the story for the customer;
- The scrum manager gets the developers team to make an estimation of the difficulty of the user stories (e.g. time or story points);
- The scrum manager, the customer and the development team agree a sprint: *a closed list of user stories to be implemented in a specified amount of time based on the story value and the estimation of difficulty of the user story*;
- During the sprint, the developers team work in the user stories included in it;
- At the given time the sprint is released and the customer can expect to have the functionalities described in the user stories previously agreed;
- The process starts again to implement the next batch of user stories.
It will be a challenge for PRIPARE’s PbD methodology to complement other agile methodologies. Agile methodologies lack an explicit design phase and expect the design of the system to emerge during the various sprints. The absence of a global design of the system will difficult the integration of privacy patterns, especially for developers who are new to privacy concerns. These types of issues can be tackled by suggesting the implantation of a Sprint-Zero [34] that helps to create a minimum design or skeleton of the system. This first sprint could ensure that privacy and security controls will be easily hooked during forthcoming sprints. Another alternative to ensure that security and privacy controls are embedded in the system is to define security and privacy requirements as user stories themselves or as constraints to functional user stories.

6.2 Project management methodologies

Organisations agree that project management is critical to business performance and organisational success. According to the PWC global project management report of 2012 [35] many organisation uses a combination of different project management methodologies. The most used methodology is PMBOK, or modifications of it. Agile methodologies adoption is growing and more than 34% of the respondents stated that they used Agile PM (project management) methodologies. The report also mentions PRINCE2 as one of the most used PM methodologies.

6.2.1 PMBOK

The PMI (Project Management Institute) developed the PMBOK® guide in an attempt to document a standard terminology and a compendium of guidelines and good practices for project management.

PMBOK® defines a set of processes in terms of inputs, tools and techniques and outputs. The process list is organized into five groups and ten different knowledge areas that should be followed during the project lifecycle.

<table>
<thead>
<tr>
<th>Knowledge areas</th>
<th>Process groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initiating</td>
</tr>
<tr>
<td>Integration</td>
<td></td>
</tr>
<tr>
<td>Scope</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td></td>
</tr>
<tr>
<td>Human resource</td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td></td>
</tr>
<tr>
<td>Risk</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td></td>
</tr>
<tr>
<td>Stakeholders</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 7: PMBOK process and knowledge matrix](image)

To complete the PMBOK® methodology to manage privacy and security-concerned projects a new knowledge area may be added, Privacy & Security. PRIPARE’s PbD methodology will try to enlighten what kind of processes should be defined inside the PMBOK® methodology to ensure privacy and security principles are followed.
6.2.2 PRINCE2

PRINCE2 is a project management methodology developed by the Office of Government Commerce (OGC). It is widely used in the UK but has also been adopted more globally (mostly in Australia or Europe [35]). Despite the fact that PRINCE2 was initially developed for ICT project development; its last version is compatible with any project typology.

PRINCE2 is a process driven methodology that is based on:

- Seven principles: continued business justification, learn from experience, defined roles and responsibilities, manage by stages, manage by exception, focus on products and tailored to suit the project environment;
- Seven themes: business case, organization, quality, plans, risk, change and progress;
- Seven processes: Starting up a project, initiating a project, directing a project, controlling a stage, managing stage boundaries, managing product delivery and closing a project;

Each process provides the specific activities that have to be performed and the inputs and outputs that should be used or provided. Principles and themes should be followed during all the project management processes.

![Figure 8: PRINCE2 seven processes](image)

PRI-PARE’s methodology will suggest specific activities, inputs and outputs to include in the PRINCE2 methodology to ensure the privacy and security principles are followed during the system development.

PRI-PARE’s methodology will consist of a series of steps. These steps will have to be mapped to those of existing project management or system development methodologies. Whilst this mapping may be trivial in some methodologies, others lack steps that are very significant for PRI-PARE (e.g. scrum methodology lacks the design step). For some steps of existing methodologies, PRI-PARE will define the necessary processes or tasks, inputs and outputs that are not covered by existing PIAs or risk management processes in order to ensure that privacy is embedded at the core of new ICT systems[^3].

[^3]: Re-engineering of legacy ICT systems falls outside PRI-PARE’s methodology see more details in Section 0.
7 Conclusions

Throughout the definition of PRIPARE’s terminology and principles, several terminology and principles sources have been thoroughly analysed in order to find similarities, differences, contradictions and ambiguities. PRIPARE does not endeavour to redefine terms or create new principles but to find the minimum starting point that will enable stakeholders to better discuss privacy and engineer truly trustworthy systems.

During the development of this document, surveys have been carried out and relevant discussions held between PRIPARE consortium’s experts in order to come to a final agreement regarding a consolidated terminology and principles list. The term privacy will remain undefined, given the inherent complexity of the term and the nuances and richness of perceptions that we feel need to be respected; instead, the use of a privacy taxonomy has been recommended in order to enable stakeholders to discuss privacy issues and requirements. This terminology will be used from now on in all PRIPARE’s outputs as it is based on a multidisciplinary consensus of the community (legal, engineering and business).

With regard to the selection of the security and privacy principles that PRIPARE will embed in ICT systems engineered with its methodology, the consortium has agreed to adopt:

- OWASP security principles; event though OWASP is about Web Application Security, there is no reason that prevents applying these security principles into any other type of ICT system;
- EU Data Protection Regulation principles;

Special attention will be paid to the data minimisation principle that can be followed by minimising the trust users need to have on the ICT systems, data processors and data controllers. It is also important to reflect on the agreement within the consortium to support the methodology with mechanisms to ensure that state-of-the-art technologies are evaluated and applied during the system engineering process. This will assure that the privacy and security principles are followed.

A number of agreements have been reached within the consortium as to what common grounds need to be considered in order to increase the chances that PRIPARE’s methodology effectively succeeds in engineering truly trustworthy systems:

- Using PIAs best practices, adapting existing PIAs processes and taking into account agencies and organisations’ recommendations;
- Including a complete risk management process to ensure privacy risks are identified and minimised;
- Ensuring user-centric systems are engineered, where data subjects can manage their identity;
- Helping stakeholders identify the most appropriate privacy instruments for a system;

In order to encourage the adoption of PRIPARE’s methodology, this should not aim to replace existing system engineering and project management methodologies but should endeavour to complement them, ascertaining common points and how these can be exploited, minimising
the adoption effort and maximising results. This task will be difficult to accomplish taking into account the variety and disparity of methodologies.

Thoughts and conclusions of the consortium are expected to be presented and discussed in the Annual Privacy Forum 2014 [87]. This document leaves the door open for the introduction of changes that reflect the feedback of experts’ opinions.

According to PRIPARE’s roadmap, the next steps for the methodology are:

- Presentation of a position paper on PbD to be discussed at a multi-disciplinary methodological workshop. The partners will invite experts to participate in the discussion of the terminology and principles relevant to PbD;
- Definition of PRIPARE’s methodology in the “Privacy and Security by Design Methodology” deliverable;
- Collection of feedback from stakeholders during the training and presentation throughout dissemination events, seminars and workshops of the initially defined methodology and reflected in the “Training Results Report” deliverable and in the project’s website and forum.
- Consolidation of Privacy and Security by Design methodology in the “Updated Privacy and security-by design Methodology” deliverable.
8 References


[41] The Tor Project. https://www.torproject.org/


