PDP4E

Methods and tools for GDPR Compliance through

Privacy and Data Protection 4 Engineering

Risk management method for data protection and privacy

Project:	PDP4E
Project Number:	787034
Deliverable:	D3.4
Title:	Risk management method for data protection and privacy
Version:	V1.0
Date:	22/07/2019
Confidentiality:	Public
Author(s):	Yuliya Miadzvetskaya (KUL) Nicolas Diaz (UDE) Victor Muntés, Elena González, Jacek Dominiak (Beawre) David Sánchez (Trialog)

Funded by

Table of Contents

OBJECTIVE OF THE DOCUMENT 6
STRUCTURE OF THE DOCUMENT 6
RELATION WITH OTHER DELIVERABLES 6
1. RISK-BASED APPROACH TO PRIVACY AND DATA PROTECTION 7
1.1. GDPR AS A RISK-BASED REGULATION 7
1.2. DEFINITION OF RISK 8
1.2.1. Lack of explicit definition of the notion of risk in the GDPR 8
1.2.2. Distinct Interpretations of the notion of risk 10
1.3. COMPLIANCE VERSUS RISK MANAGEMENT DEBATE 15
3. THE LINDDUN PRIVACY THREATS MODELLING METHODOLOGY 21
3.1. THE LINDDUN METHODOLOGY STEPS 21
3.2. RATIONALE FOR ALIGNING THE GDPR AND LINDDUN 23
3.3. ALIGNING LINDDUN THREATS CATEGORIES WITH THE GDPR VOCABULARY 25
3.3.1. Linkability 26
3.3.2. Identifiability 28
3.3.3. Non-repudiation 30
3.3.4. Detectability 31
3.3.5. Information Disclosure 32
3.3.6. Unawareness 32
3.3.7. Non-compliance 33
3.3.8. Conclusion 34
ANNEX A EXTENDING LINDDUN METHODOLOGY 35
1.1. RATIONALE FOR EXTENDING LINDDUN 35
1.2. SPECIFICATION OF LINDDUN NON-COMPLIANCE THREAT 36
1.2.1. Unlawful ground 36
1.2.2. Undefined purpose 37
1.2.3. Undetected data subject categories 38
1.2.4. Undetected personal data categories 40
ANNEX B CONCLUSIONS WITH REGARD TO RISK IDENTIFICATION UNDER EXTENDED LINDDUN(+4U) 42
4. REFERENCES 45

Document History

Version	Status	Date
V0.1	Initial Table of Contents	18/04/2019
V0.2	UDE contribution integrated	13/05/2019
V0.3	Further KUL contribution	20/06/2019
V0.4	Beawre and Trialog contributions integrated	24/06/2019
V0.5	Suggestions by Beawre and CEA integrated by KUL	10/07/2019
V1.0	Suggestions by UPM integrated by KUL	22/07/2019

<table>
<tr><th colspan="3">Approval</th></tr>
<tr><th></th><th>Name</th><th>Date</th></tr>
<tr><td>Reviewer</td><td>Thibaud Antignac (CEA)</td><td>09/07/2019</td></tr>
<tr><td>Reviewer</td><td>Juan Carlos (UPM)</td><td>19/07/2019</td></tr>
<tr><td>Authorised</td><td>Antonio Kung (Trialog)</td><td>22/07/2019</td></tr>
<tr><th colspan="3">Circulation</th></tr>
<tr><th colspan="2">Recipient</th><th>Date of submission</th></tr>
<tr><td colspan="2">Project partners</td><td>22/07/2019</td></tr>
<tr><td colspan="2">European Commission</td><td>22/07/2019</td></tr>
</table>

List of Figures

Figure 1. The generic iterative process for carrying out a DPIA.. 12

Figure 2. Compliance approach using a PIA, CNIL.. 12

Figure 3. Protection goals... 13

Figure 4. DPIAs steps, ICO 14

Figure 5. Non-compliance tree from LINDDUN with root threats (circles), concrete threats (boxes), AND relation, OR relation.. 15

Figure 6. Example of comparison between the risk management methodology used in MUSA (inspired by CORAS and 31000) and ISO 29134. 18

Figure 7. Risk Management methodology for PDP4E's Risk Management tool. 18

Figure 8. Risk Matrix considering 3 generic incidents 20

Figure 9. The LINDDUN methodology steps... 21

Figure 10. The data flow diagram (DFD) of the Social network data 22

Figure 11. Mapping threat categories (Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, Non-compliance) to DFD element types. 22

Figure 12. Example of LINDDUN threat tree with root threats (circles), concrete threats (boxes), AND relation, OR relation 23
Figure 13. Taxonomy of Privacy Mitigation Strategies 23
Figure 14. Aligning the circular A-130 FIPPs to the Privacy Engineering and Security Objectives 24
Figure 15. Unlawful ground 37
Figure 16. Undefined purpose 37
Figure 17. Undetected data subject categories 38
Figure 18. Undetected personal data categories 41

List of Tables

Table 1. Description of risk based provisions in the GDPR 8
Table 2. Description of risk related provisions in the GDPR 10
Table 3. Distinct interpretations of the notion of risk 11
Table 4. Description of risk-based and rights-based approaches 15
Table 5. Description of Linkability under the GDPR lens 26
Table 6. Description of Identifiability under the GDPR lens 29
Table 7. Description of Non-repudiation under the GDPR lens 30
Table 8. Description of Detectability under the GDPR lens 31
Table 9. Description of Information Disclosure under the GDPR lens 32
Table 10. Description of Unawareness under the GDPR lens 32
Table 11. Description of Non-compliance under the GDPR lens 33
Table 12. Overview of child-specific provisions 40

Abbreviations and Definitions

Abbreviation	Definition
DPIA	Data protection impact assessment
GDPR	General Data Protection Regulation
ICT	Information and Communication Technologies
IOT	Internet of Things
DFD	Data Flow Diagram
PDP	Privacy and Data Protection
PDP4E	Privacy and Data Protection 4 Engineering
PDPbD	Privacy and Data Protection by Design
PET	Privacy-enhancing Technologies
TFEU	Treaty on the Functioning of the European Union

WP29	Data Protection Working Party
LINDDUN	Linkability, Identifiability, Non-repudiation, Detectability, Information Disclosure, Unawareness, Non-compliance
FIPP	Fair Information Practice Principles
IOI	Item of interest
ICO	Information Commissioner’s Office
ECJ	European Court of Justice
MS	Member States

Executive Summary

Objective of the document

This document details the contents of the risk management method steps based on LINDDUN and covers the adaptations made in order to ensure that LINDDUN takes into account the GDPR provisions [1]. LINDDUN is a privacy threat modelling methodology integrating 7 main privacy threat categories. In addition, an attempt will be made to asses how LINDDUN threat categories relate to GDPR provisions on data protection principles and data subject rights. The described method will be adapted to the feedback received from stakeholder validation after the first iteration.

Structure of the document

The first section of this document provides an overview of a risk-based nature of the GDPR. It also sheds some light on the lack of an explicit definition of risk in the GDPR. Moreover, compliance versus risk debate in the framework of DPIAs will be covered in the last part of this section.

In the second section, we describe the main steps followed by the risk management methodology.

The third section provides a description of LINDDUN methodology steps and explains the rationale for aligning LINDDUN with the GDPR vocabulary. In addition, an attempt will be made to translate LINDDUN threats categories into the GDPR lexicon.

Relation with other deliverables

This deliverable has been written in parallel to D3.1. Whereas D3.1 focused on the expected roles and their expertise, user needs and specification of the expected high-level functionalities, this document focuses on the methodological aspects of such risk management. Hence, the methodology has been depicted not only considering existing background on the topic, but to tackle with the objectives set in D3.1. During the preparation of the document, we have initiated discussions with the different technical work packages in relation to the touch points between a risk management process and the different disciplines considered in PDP4E. In particular, active conversations in relation with modelling of data flow diagrams, essential for the risk management method, have been conducted with WP4 (Requirements elicitation) and WP5 (Model-driven design). The reader may need to check WP4 and WP5 methodologies (D4.1 and D5.1) in order to fully understand the extent of the risk management method.

Guided by the development of the risk management tool, this document will be updated accordingly as part of D3.5. Moreover, discussions to better adapt the LINDDUN methodology to compliance of the GDPR (and, in particular, to risk managent as considered by the regulation) are still ongoing. Appendixes of this document describes the results of our discussions and the need for such adaptation.

1. Risk-based approach to privacy and data protection

This section provides an overview of a risk-based nature of the GDPR (1.1), of its risk-related provisions (1.2) and gives some insight into compliance versus risk debate (1.3).

1.1. GDPR as a risk-based regulation

The GDPR embraces a risk-based approach to data protection by encouraging controllers to perform the assessment of personal data processing operations in order to identify activities posing a high risk to data subjects and adopt tailored responses. The promoters of a risk-based approach argue that legal compliance should rather shift to the framing of responsible data use based on risk management [2]. Article 35 of the GDPR is the first risk management method enshrined in the European data protection law [3]. It provides for an obligation to carry out, prior to the processing, an assessment of the impact of the envisaged processing operations on the protection of personal data where it is likely to result in a high risk to the rights and freedoms of natural persons. The rights and freedoms of the data subjects primarily concern the right to privacy but may also involve other fundamental rights such as freedom of speech, freedom of thought, freedom of movement, prohibition of discrimination, right to liberty, conscience and religion [2].

Assessing data protection related risks equals to assuming that every personal data processing operation may entail risks for data subjects. Thus, Recital 75 GDPR refers to risks resulting from personal data processing which could lead to physical, material or non-material damage and provides for a non-exhaustive list of negative consequences such processing may have on individuals (e.g. evaluation of personal aspects for the purposes of work performance prognosis). Based on risk assessment conclusions, unacceptable privacy risks will be eliminated through the implementation of effective privacy controls as much as reasonable taking into account the state-of-the-art, cost and available mitigation controls. While completely eliminating all the privacy risks is almost impossible and is rather a wishful thinking, the privacy risk management aspires, first of all, to identify and eliminate as early as possible all the “unacceptable risks”. According to Recital 84, the supervisory authority should be consulted, “*where a data-protection impact assessment indicates that processing operations involve* ***a high risk which the controller cannot mitigate by appropriate measures*** *in terms of available technology and costs of implementation.”*

Risk-based approach (Recital 74)	The controller should be obliged to implement appropriate and effective measures and be able to demonstrate the **compliance of processing activities with this Regulation**, including the effectiveness of the measures. Those measures should take into account the nature, scope, context and purposes of the processing and **the risk to the rights and freedoms of natural persons**.

Moreover, the risk-based nature of the GDPR is translated through the requirement of a higher standard of protection with regard to some singled out cases, such as processing of special categories of data or child’s personal data. In addition, many provisions of the GDPR require the assessment of the likelihood and severity of the risk in order to determine what technical and organisational measures should be implemented and whether the personal data breach notification is required or not.

Risk level (high or not) based on	Risk-based compliance obligation
categories of data (sensitive) (Recital 51, 53)	Personal data which are, by their nature, particularly **sensitive** in relation to fundamental rights and freedoms **merit specific**

	protection as the context of their processing could create significant risks to the fundamental rights and freedoms.
categories of data subjects (children) (Recital 38)	**Children merit specific protection**, as they may be less aware of the risks.
likelihood and severity the risk for rights and freedoms of natural persons	**The higher the risk, the stricter the compliance obligation:** • the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures (Article 25) • the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security (Article 32) • the controller must notify the personal data breach to the supervisory authority (Article 33) • the controller shall communicate the personal data breach to the data subject without undue delay (Article 34, Recital 86) • DPIA (Article 35, Recital 84, 90, 91, 94) • obligation to notify the processing of personal data to the supervisory authorities (Recital 89) • obligation to keep records of processing activities (Article 30) • data protection officer (Articles 37-39)

Table 1. Description of risk based provisions in the GDPR

1.2. Definition of risk

This section will delve into the definition of risk and its different aspects, as set out in the GDPR (1.2.1) and analyse distinct approaches towards the notion of risk (1.2.2).

1.2.1. Lack of explicit definition of the notion of risk in the GDPR

While the GDPR relies on a **tailored "risk-based approach"** entailing the assessment of risk and the adjustment of mitigation strategies to its potential effect on data subjects' rights and freedoms, there is **no agreed definition of the concept of risk**. On the one hand, a lack of an explicit definition of the notion of risk under the GDPR causes lengthy debates on what should be captured by this term. On the other hand, it allows for **a greater flexibility** and a more tailored approach towards risk management. Thus, risk can be compared to a **black box**, where one can insert different elements depending on the nature, scope, context and purposes of the processing. Despite a lack of an explicit definition of risk under the GDPR, different elements of the notion of risk can be still grasped through the wording of its recitals and articles.

Risk related elements	GDPR definitions
Risk definition[1] (Recital 75)	The risk to the rights and freedoms of natural persons, of varying likelihood and severity, may result from personal data

[1] Recitals are interpretative tools in the EU legal order and can help to explain the purpose and intent of an act. However, they do not have any autonomous legal effect. The ECJ held that 'recital cannot be relied upon to

	processing which could lead to **physical, material or non-material damage.**
Non-exhaustive list of examples of **physical, material or non-material damage (Recital 75) to** data subjects	• Discrimination • Identity theft / fraud, financial loss • Reputation damage • Loss of confidentiality of personal data protected by professional secrecy • Unauthorised reversal of pseudonymisation • Any other significant economic or social disadvantage • Individuals deprived of rights and freedoms, or prevented from exercising control over their data • Processing sensitive data, including data on racial or ethnic origin, political opinions, religion or philosophical beliefs, trade-union membership; genetic data; health data; data concerning sex life; or data on criminal convictions and offences or related security measures • Profiling (personal aspects are evaluated [e.g. analyse or predict work performance, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements] to create or use personal profiles) • Processing children's and vulnerable persons' data • Processing large amounts of data affecting large numbers of individuals
Risks related to personal data processing (Recital 83)	• Accidental or unlawful destruction • Loss • Alteration • Unauthorised disclosure of, or access to, personal data
Aspects to take into account for risk assessment (**likelihood and severity**) (Recital 76)	• Nature • Scope • Context • Purposes of the processing
Criteria for risk level **(high or not)** assessment (Recital 76)	Risk should be evaluated on the basis of an **objective assessment**, by which it is established whether data processing operations involve a risk or a high risk.
Aspects to take into account for risk evaluation under DPIA (Recital 84)	• Origin • Nature • Particularity • Severity of a risk
Types of processing operations which are likely to result in **a high risk** to the rights and freedoms of natural persons (Recital 89, Recital 91, Article 35(3)) (to be complemented by DPAs)	• processing using new technologies • a new kind of data processing where no data protection impact assessment has been carried out before • personal data are processed for taking decisions regarding specific natural persons following any

interpret a provision in a manner clearly contrary to its wording'. (Judgment of the Court (Third Chamber) of 13 July 2006, Manfredi, ECLI:EU:C:2006:461).

	systematic and extensive evaluation of personal aspects relating to natural persons based on profiling those data or following the processing of special categories of personal data, biometric data, or data on criminal convictions and offences or related security measures • processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 10 • monitoring publicly accessible areas on a large scale, especially when using optic-electronic devices
Risk mitigation measures (Recital 28, Article 32)	• Pseudonymisation and encryption of personal data; • Ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services; • Ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident; • A process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.

Table 2. Description of risk related provisions in the GDPR

In the guidelines on DPIAs of Article 29 Working Party a risk is viewed as "*a scenario describing an event and its consequences, estimated in terms of severity and likelihood*" [4]. Thus, risk has 2 intrinsic elements: **an event and its consequences**. Gellert in his contribution on a notion of risk suggests an interesting exercise of identifying a risk under the GDPR with regard to those 2 elements. A new reading of Art. 35 (1) GDPR under this lens suggests that the "*high risk to the rights and freedoms*" would be the consequence, whereas the "*protection of personal data" comes under the notion of "event"* leading to these consequences [3]. Thus, If accountability obligations are not fulfilled by controllers/processors and all the necessarily organisational and technical measures are not implemented, it will necessarily lead to negative consequences to data subjects' fundamental rights. In other words, ***"the lower the compliance or the higher the "non-compliance event", the higher the risk to the data subjects' fundamental rights***" [3].

1.2.2. Distinct Interpretations of the notion of risk

Two different approaches towards the notion of risk can be singled out. First of all, those who do not consider non-compliance as risk to rights and freedoms of data subjects. The compliance with the legal framework is considered as an obligation, a mere obedience. It is assumed that compliance should always take place and risk mitigation measures should tackle other "uncertainties" on top of the compliance. Moreover, the supporters of this approach highlight that the process of identifying, assessing and mitigating risks of non-compliance with existing regulations is traditionally more focused on the risks for the organisation itself rather than on the risks and harms to individuals. Therefore, they advocate for a clear shift of focus from considering risk management as a process for building and demonstrating compliance to levelling it up to data subjects' oriented perspective.

The second approach recognizes that the compliance alone cannot mitigate all privacy risks, in particular in the era of digitalization and technological progress with laws still lagging behind. Nevertheless, the supporters of “risk of non-compliance” approach advocate that compliance should be integrated in risk analysis process due to the inherently scalable nature of compliance [3]. For instance, how much data minimisation and purpose limitation is enough for the processing of personal data and how much is enough for the processing of special categories of personal data? How can it be assessed that the compliance is achieved and maintained throughout all the data processing activities?

Distinct Interpretations of the notion of risk	
Risk of non-compliance	**Risk to data subjects’ rights**
“Compliance should be directly integrated in the risk analysis process, because compliance is inherently scalable”. Non-respect for data minimization principle may result in violation of data subjects’ fundamental rights. But how much data minimization do you need to be compliant?	Legal requirements could not be optional and there is no discretion to the data controller about the data subjects’ rights.
Criticised for being minimal requirements	Criticised for forgetting the scalable nature of compliance and its link to risks to data subjects’ rights.

Table 3. Distinct interpretations of the notion of risk

Despite a strong link between a risk of non-compliance and a risk to data subjects' fundamental rights [3], these two issues are thrown in two different baskets and are always examined separately. Almost all the existing methodologies advocate for their strict separation. In this regard, Article 29 Working Party in its guidelines on DPIA methodology also suggests a separation between compliance and risks, as demonstrated in Figure 1. As such compliance is not examined as a risk and the processing is assessed with regard to its proportionality and necessity. The notion of the risks to the rights and freedoms of data subjects arises only at a later stage, once the compliance is established. In addition, mitigating measures, as suggested by Article 29 Working Party[2], are also separated in two categories, as those envisaged to “*address the risks*” and those that aim to “*demonstrate compliance with the GDPR*” [4].

[2] The Article 29 Data Protection Working Group (“Working Party”) is a European advisory body comprising of representatives of the national data protection authorities. Although the opinions of the Working Party are not binding, significant authoritative value is attached to them, as all the Member States are represented in this body. Since the entry into force of the GDPR, it was replaced by the European Data Protection Board.

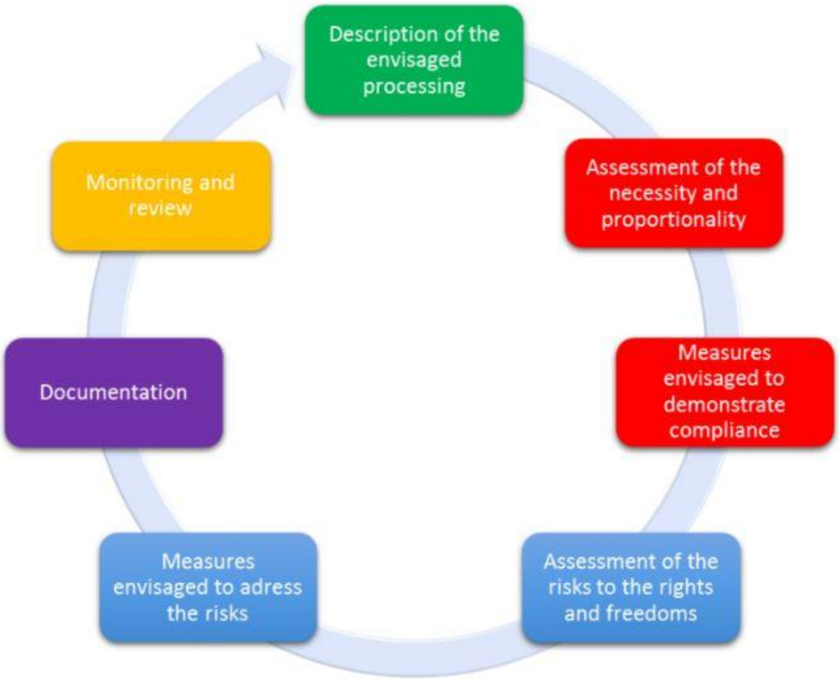

Description of the envisaged processing
Assessment of the necessity and proportionality
Measures envisaged to demonstrate compliance
Assessment of the risks to the rights and freedoms
Measures envisaged to adress the risks
Documentation
Monitoring and review

Figure 1. The generic iterative process for carrying out a DPIA

In addition, the methodology suggested by CNIL relies on the same assumption that compliance with "non-negotiable" **fundamental rights and principles**, established by law, should always take place (Figure 2). And the risk is viewed as "*a hypothetical scenario that describes a feared event and all the threats that would allow this to occur*" [5]. Thus, CNIL proposes to focus the risk analysis on privacy risks, "*related to the security of personal data and having an impact on data subjects' privacy" [5]*. One might question whether this approach doesn't mean a shift of privacy impact assessment to security impact assessment. While privacy shares enough features with security and both are inherently interconnected, privacy has its own meaning.

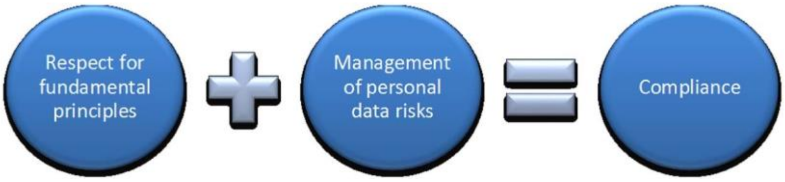

Respect for fundamental principles
Management of personal data risks
Compliance

Figure 2. Compliance approach using a PIA, CNIL

Bieker et al. methodology [6] relies on the same assumption that compliance is compulsory as a minimal requirement. Then this methodology refers to **protection goals** (1) availability, (2) integrity, (3) confidentiality, (4) unlinkability, (5) transparency, (6) intervenability, as shown in Figure 3. "*Each protection goal incorporates further, derived protection goals, each of which can be deduced from legal provisions in the GDPR*." [6] This approach raises some questions, because it assumes the compliance with the GDPR in the first place, but then proposes to complete each of the protection goals with the GDPR legal provisions. In his contribution, Gellert questions the "*utility to adopt events that are so closely related to compliance and whether the distinction between legal compliance and these events is not artificial*" [3].

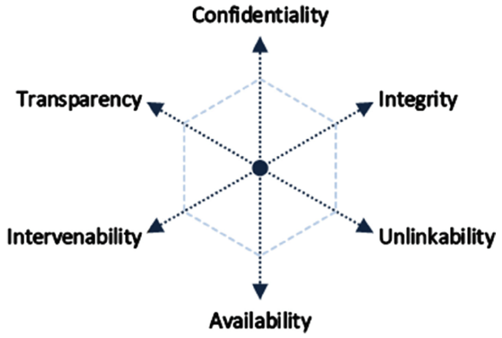


Figure 3. Protection goals[3]

In this way, many of the existing privacy risk management methodologies could be criticized for, on the one hand, using security risks as feared events and thus making it merely a data security methodology with privacy still lagging behind. On the other hand, they might be criticized for ignoring the inherently scalable nature of compliance and, thus, making an artificial separation between two inherently connected issues such as compliance with legal requirements and risks to rights of data subjects.

Information Commissioner's Office[4] takes a slightly different approach towards compliance and suggests to include associated compliance and corporate risks in step 5 of the methodology (Figure 4), notably "identify and assess risks". It seems that ICO admits that compliance and corporate risks may be intertwined with all other "risks" or even trigger all other risks. Therefore, depending on circumstances, there may be a need to integrate them in the risk analysis process.

[3] Felix Bieker, Michael Friedewald, Marit Hansen, Hannah Obersteller, and Martin Rost, "A Process for Data Protection Impact Assessment under the European General Data Protection Regulation", 4th Annual Privacy Forum, APF 2016, Frankfurt/Main, Germany, September 7–8, 2016 Proceedings.

[4] The UK's independent authority set up to uphold information rights in the public interest, promoting openness by public bodies and data privacy for individuals.

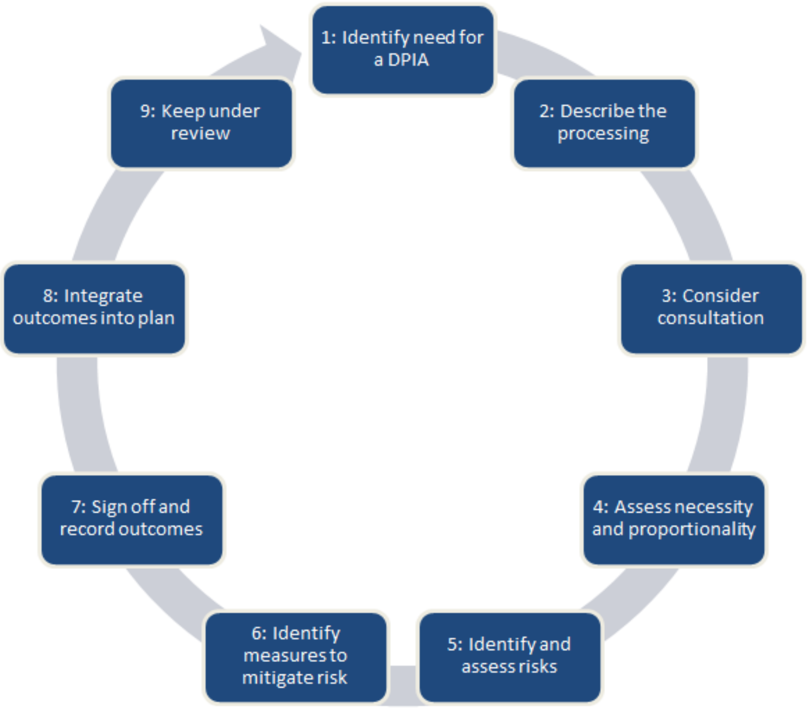

1: Identify need for a DPIA
2: Describe the processing
3: Consider consultation
4: Assess necessity and proportionality
5: Identify and assess risks
6: Identify measures to mitigate risk
7: Sign off and record outcomes
8: Integrate outcomes into plan
9: Keep under review

Figure 4. DPIAs steps, ICO

A different approach towards compliance is also suggested by LINDDUN methodology. LINDDUN includes non-compliance as one of its 7 threats types (Figure 5). Non-compliance under LINDDUN framework is closely related to legislation and policy with a particular focus on consent requirement. The compliance requirement applies to all the elements of DFDs and "*affects the system as a whole, because each system component (including data flow, data store and process) is responsible to ensure that actions are taken in compliance with privacy policies, legislative rules, and data subjects' consents" [7]*. LINDDUN approach is novel because it doesn't take compliance for "non-negotiable" legal principles and deals with it under the risk/threat perspective. Although "*LINDDUN is not a compliance technique, it does implement several principles imposed by data protection legislation (consent, awareness, data minimisation etc.) and explicitly draws attention to the need of regulatory compliance" [3]*.

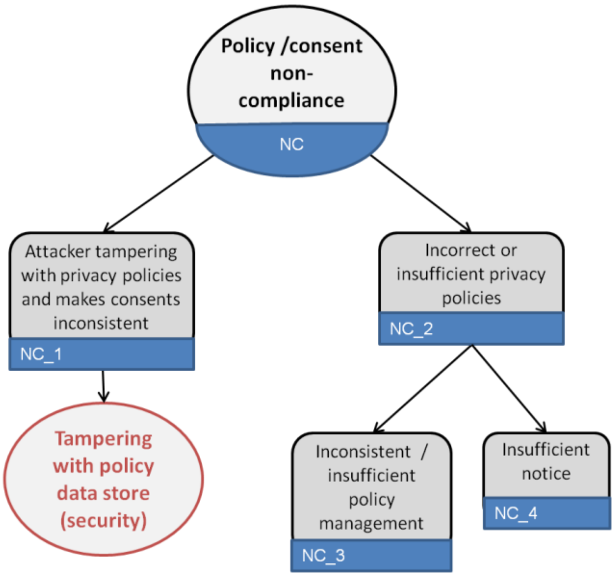

Policy /consent non-compliance
NC
Attacker tampering with privacy policies and makes consents inconsistent
NC_1
Incorrect or insufficient privacy policies
NC_2
Tampering with policy data store (security)
Inconsistent / insufficient policy management
NC_3
Insufficient notice
NC_4

Figure 5. Non-compliance tree from LINDDUN with root threats (circles), concrete threats (boxes), AND relation, OR relation

1.3. Compliance versus risk management debate

As examined above, the risk analysis, including the analysis of non-compliance and its consequences on the data subjects' fundamental rights, within one single risk calculation is not supported by current DPIA methodologies. The conventional practice towards privacy risk analysis consists in putting emphasis on other risks, going beyond the scope of compliance. And this approach towards risk has its historical explanation stemming from the debate between risk-based and rights-based approaches [2].

The risk-based nature of the GDPR was criticized for "*putting the focus of protection only when harms have arisen or are susceptible to*" [3]. Article 29 WP in its statement on the role of a risk-based approach noted that "*the risk-based approach is being increasingly and wrongly presented as an alternative to well-established data protection rights and principles rather than as a scalable and proportionate approach to compliance*" [2] and that data controllers should "*always be accountable for compliance with data protection obligations*"[2].This statement of Article 29 WP sets the basis for a clear separation between compliance and risks, which is now supported by a number of DPIA methodologies.

Risk-based approach	Rights-based approach
The level of protection afforded should be equivalent to the potential harms created by the processing of data.	The right to data protection should apply irrespective of the level of risk, and therefore provide for a uniform level of compliance or "minimum and non-negotiable level of protection for all individuals".

Table 4. Description of risk-based and rights-based approaches

However, this separation coming from the Statement of the Article 29 WP seems to ignore the scalable nature of compliance. How much data minimisation do you need to be compliant and how much data minimisation is enough to eliminate certain risks to rights and freedoms of individuals? Getting compliant may be compared to shooting at a moving target.

Therefore, the “*compliance should never be a box-ticking exercise, but should really be about ensuring that personal data is sufficiently protected*” [2]. For instance, it cannot be excluded that the controller/processor, while acting in good faith in ensuring their legal compliance, may still cause further risks to rights and freedoms of individuals stemming from involuntary non-respect for basic legal requirements.

2. Risk management methodology

In this section, we present the Risk Management methodology that we implement in WP3 tool (2.1). While all risk management methodologies presented in the literature and accepted by the international community through standards, scientific work or other best practices are similar, they also differ in different aspects. In this section, we briefly discuss some of the most well-known risk management methodologies and adapt them into a proposal that fits PDP4E's project requirements.

We explored best practices in industry and considered previous related FP7 and H2020 projects (in particular MODAClouds and MUSA) to come up with a proposal for PDP4E. In particular, we considered the following approaches:

- **Risk management methodologies used in MODAClouds and MUSA (and CORAS methodology implicitly)**: MODAClouds risk management methodology was inspired by the CORAS methodology [8]. The methodology implemented in these projects, proposed a simplified version of the CORAS methodology to improve the usability of the tools.
- **ISO 31000:2018**[5]: ISO 31000:2018 provides guidelines on managing risk faced by organizations. The application of these guidelines can be customized to any organization and its context. This standard provides a common approach to managing any type of risk and is not industry or sector specific. Therefore, it can be used throughout the life of the organization and can be applied to any activity, including decision-making at all levels. Because of the fact that it is the most generic standard to describe risk management activities and it is agnostic to a particular context, we take it as a general reference for PDP4E's Risk Management tool.
- **ISO/IEC 29134:2017**[6]: ISO/IEC 29134:2017 gives guidelines for: (i) a process on privacy impact assessments, and (ii) a structure and content of a PIA report. It is applicable to all types and sizes of organizations, including public companies, private companies, government entities and not-for-profit organizations. ISO/IEC 29134:2017 is relevant to those involved in designing or implementing projects, including the parties operating data processing systems and services that process personal data.

As an example of the comparisons performed among existing methodologies for risk management, in *Figure 6*, we show a visual summary of the main steps followed by the risk management methodology in MUSA and the steps suggested in ISO 31000:2018 and in ISO/IEC 29134:2017. While the vocabulary is not identical, the processes are very similar, and we were able to establish reasonable mappings among all the processes. For instance, in MUSA assets had to be defined and threats were identified with respect to those assets. We have also added a step to detect vulnerabilities following CORAS's recommendations, but we have considered this step as optional. In ISO 29134, the definition of assets and vulnerabilities is quite ambiguous, but they put the emphasis in the description of risk sources. Both methodologies or descriptions define *threats* (also called *unwanted incidents* in CORAS) and then *risks*. In general, a risk is to be considered an unwanted incident that has been assessed as a risk, *i.e.* its likelihood and impact/consequence have been evaluated. In ISO 20134, the analysis of impacts is treated separately, but in the rest of standards, this is usually part of the risk analysis step (this fact is captured through the orange arrow in the figure, indicating that impact analysis is done as part of the

[5] iso.org/standard/65694.html
[6] https://www.iso.org/standard/62289.html

risk assessment in most methodologies (like in CORAS). Some methodologies talk about treatments, while some other talk about controls. In general, these are all different terms to refer to mitigation actions.

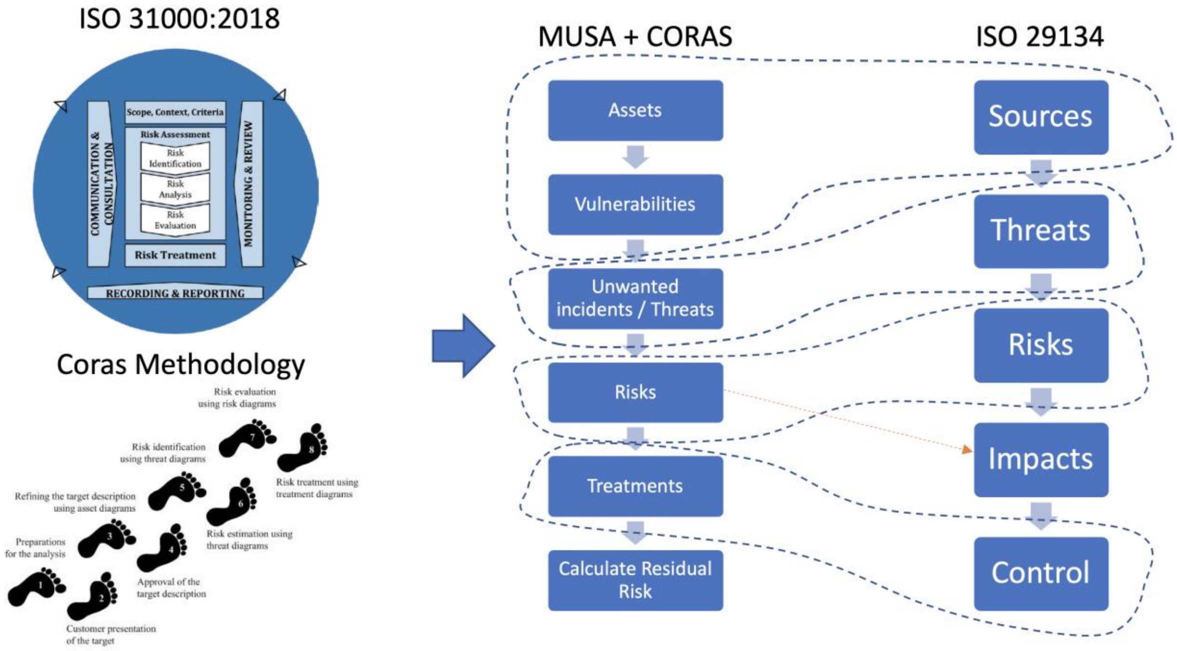

ISO 31000:2018
Scope, Context, Criteria
Risk Assessment
Risk Identification
Risk Analysis
Risk Evaluation
Risk Treatment
COMMUNICATION & CONSULTATION
MONITORING & REVIEW
RECORDING & REPORTING
Coras Methodology
Risk evaluation using risk diagrams
Risk identification using threat diagrams
Risk treatment using treatment diagrams
Refining the target description using asset diagrams
Risk estimation using threat diagrams
Preparations for the analysis
Approval of the target description
Customer presentation of the target
1
2
3
4
5
6
7
8
MUSA + CORAS
Assets
Vulnerabilities
Unwanted incidents / Threats
Risks
Treatments
Calculate Residual Risk
ISO 29134
Sources
Threats
Risks
Impacts
Control

Figure 6. Example of comparison between the risk management methodology used in MUSA (inspired by CORAS and 31000) and ISO 29134.

Based on this analysis, in *Figure 7*, we propose a methodology for risk management in PDP4E. In this figure we do not only depict the different steps of our methodology, but we also link these steps with the actors playing a central role in each step.

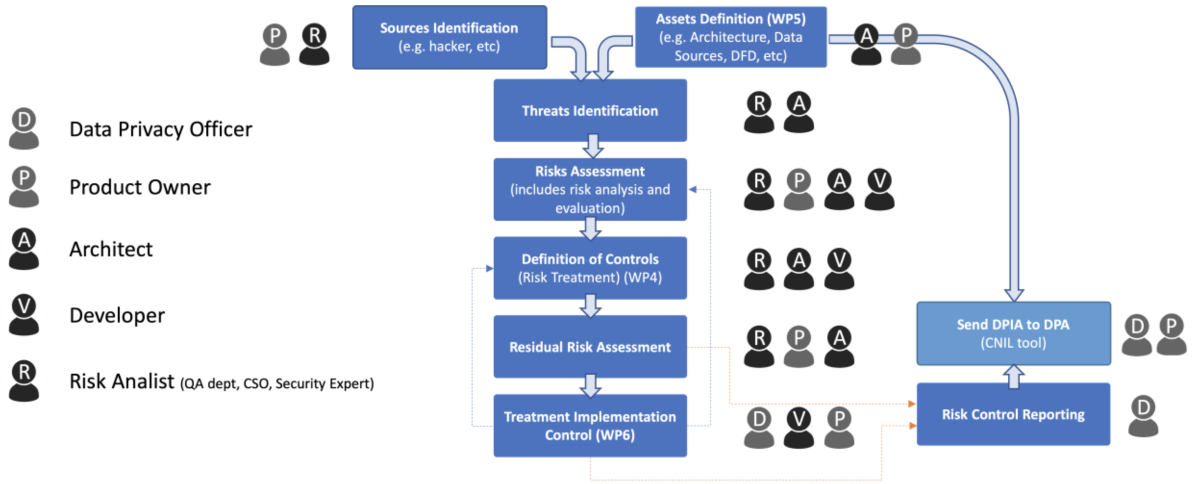

D Data Privacy Officer
P Product Owner
A Architect
V Developer
R Risk Analist (QA dept, CSO, Security Expert)
Sources Identification (e.g. hacker, etc)
Assets Definition (WP5) (e.g. Architecture, Data Sources, DFD, etc)
Threats Identification
Risks Assessment (includes risk analysis and evaluation)
Definition of Controls (Risk Treatment) (WP4)
Residual Risk Assessment
Treatment Implementation Control (WP6)
Send DPIA to DPA (CNIL tool)
Risk Control Reporting
P R
A P
R A
R P A V
R A V
R P A
D V P
D P
D

Figure 7. Risk Management methodology for PDP4E's Risk Management tool.

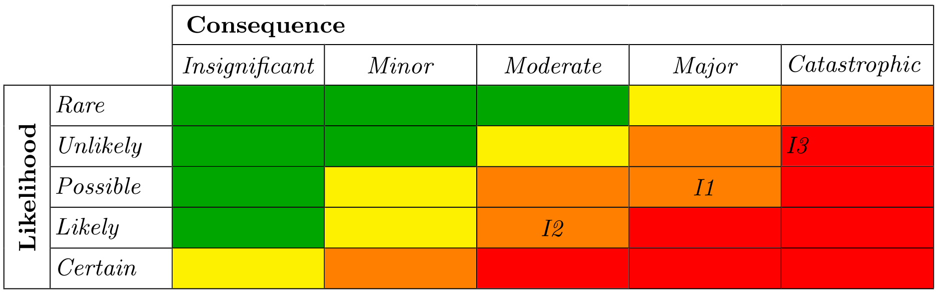
Our methodology, inspired by the previous analysis, can be summarized in 7 main steps. We add two more steps to emphasize the need for reporting and the contribution of the outcomes of our analysis to a broader-scope DPIA, through tools such as CNIL. The 7 main steps are:

- **Sources identification**: a specific risk may have one or more sources. These represent the root causes or the actors initiating the risk. Our methodology will allow expressing potential risk sources and to associated these sources to vulnerabilities, threats and risks later on in the process.
- **Assets definition**: most risk methodologies recognize the need to explicitly define assets. This is usually an essential part of the methodology as the risks are analysed with respect to the impact they may have on these assets. In PDP4E, our assets will mainly be DFDs defined in WP5 and the components of these DFDs. The information about the architecture of the system will also be taken into consideration, linked with the components represented in the DFD. We have not included a step to describe the vulnerabilities associated to a particular asset, as we consider this one an optional step in our methodology. However, PDP4E's tool should be able to provide the means for an organization to define the vulnerabilities related to a component of a DFD or a subset of components. Please note that new risks may arise when combining different types of components in a system and this is hardly managed with existing risk management technology.
- **Threats identification**: in this step, users are encouraged to identify threats that may affect the components in the described system. A previous detection of vulnerabilities in the previous step may be also helpful for threat identification since the previous definition of vulnerabilities may make some threats evident and it also allows for completeness checks at the end, checking for vulnerabilities that have not been mapped to the current list of defined threats.
- **Risk Assessment**: risk assessment is composed of two different steps: risk analysis, where risks are evaluated in terms of likelihood and consequence, and risk evaluation, where risks are accepted, or they are classified as risks that need to be mitigated. We discuss different approaches for conducting this particular activity at the end of this section (see "Approaches for risk assessment").
- **Definition of Controls**: mitigation actions are defined in the form of controls. A control can act as a mitigation action of different risks and a risk may require several treatments. Deciding what is the minimum number of treatments required to mitigate a risk may not be straightforward and our tool will support it.
- **Residual Risk Assessment**: once the mitigation controls are defined, the residual risks need to be reassessed. This involves again two steps: risk analysis, where likelihood and consequence are updated after the application of the control(s), and risk re-evaluation, where risks are analysed again, and they are classified as accepted or further mitigation actions required.
- **Treatment Implementation Control**: finally, we add a last step in the methodology that goes beyond many of the methodologies defined before. In particular, it involves the control of the implementation of the mitigation actions (or controls) proposed in the previous step. This step may be connected to the tools generated in WP6, to collect evidences from security and privacy monitoring in order to match them to controls and risks.

We foresee several roles involved in the usage of the PDP4E Risk Management tool as depicted in Figure 7, including architects, developers, risk management owners (e.g. DPO), product owners, risk analysts. More information about these roles can be found in D3.1.

Aproaches for risk assessment

One of the main challenges in risk management is precisely the estimation of the risk value corresponding to a particular unwanted incident. In security-oriented approaches like CORAS [8], risks are estimated using a *risk function* and the help of an expert in the field. Such risk function is often represented using a *risk matrix* like the one of Figure 8 which is divided in four sections each representing one of the risk levels: very low (green), low (yellow), high (orange), and very high (red). A risk level is obtained from the *frequency* of the unwanted incident (i.e. rare, unlikely, possible, likely, and certain) and its *consequence* (i.e. insignificant, minor, moderate, major and catastrophic). When analysing security threats, suck risk estimation is conducted over the systems' assets. That is, an expert elaborates the corresponding risk matrix for each asset and estimates the corresponding risks. Afterwards, treatments are proposed for those risks that are considered unacceptable for the particular software project.



		Consequence				
		Insignificant	*Minor*	*Moderate*	*Major*	*Catastrophic*
Likelihood	*Rare*	🟩	🟩	🟩	🟨	🟧
	Unlikely	🟩	🟩	🟨	🟧	🟥 *I3*
	Possible	🟩	🟨	🟧	🟧 *I1*	🟥
	Likely	🟩	🟨	🟧 *I2*	🟥	🟥
	Certain	🟨	🟧	🟥	🟥	🟥

Figure 8. Risk Matrix considering 3 generic incidents

Whereas an approach like the one described in CORAS [8] seems to suit a security threat analysis, a Data Protection Impact Assessment (DPIA) introduces new challenges. First, the GDPR introduces legal obligations that could be understood as treatments to pre-identified risks. For instance, the GDPR creates incentives to apply pseudonymisation[7] when processing personal data. One can easily assume that this is grounded on privacy risks that may occur if personal data of data subjects are not properly protected. For instance, a patient can get a higher fee from her insurance company if they find out that she suffers from specific diseases (i.e. unjustified discrimination). Under this premise, not following a legal obligation is a risk that is never acceptable for the company or institution in charge of the software project. Another difference is that whereas risks in security are estimated for the system's *assets*, risks in a DPIA are analysed over the *privacy rights* of data subjects. This not only means that when conducting a DPIA we are estimating risks on behalf of the data subjects, but also that such estimation must safeguard their privacy rights. This raises an ethical question of whether it is possible or not to accept some risks on behalf of the users and, consequently, not applying the corresponding controls.

[7] Art. 25 GDPR.

3. The LINDDUN privacy threats modelling methodology

This section provides a description of LINDDUN methodology steps (3.1) and explains the rationale for aligning LINDDUN with the GDPR vocabulary (3.2). In subsection 3.3 LINDDUN threats categories are translated into the GDPR lexicon.

3.1. The LINDDUN methodology steps

LINDDUN[8] is a privacy threat analysis methodology that integrates 7 main privacy threat categories [7]:

- **L**inkability (L) occurs when one can sufficiently distinguish whether 2 items of interest (IOI, such as requests from a user) are related
- **I**dentifiability (I) occurs when it is possible to pinpoint the identity of a subject (e.g., a user)
- **N**on-repudiation (Nr) occurs when it is possible to gather evidence so that a party cannot deny having performed an action
- **D**etectability (D) occurs when one can sufficiently distinguish whether an IOI exists, e.g., in a system
- **D**isclosure of information (Di) is the exposure of information to individuals who are not supposed to have access to it
- **U**nawareness (U) occurs when the user is unaware of the information he is supplying to the system and the consequences of his/her act of sharing
- **N**on-compliance (Nc) occurs when the system is not compliant with the (data protection) legislation, its advertised policies and the existing user consents

LINDDUN methodology steps are divided in problem space steps (step 1-3), which aim at describing privacy threats, and in solution space steps (step 4-6) necessary for the elicitation of mitigation measures and solutions corresponding to the threats identified.

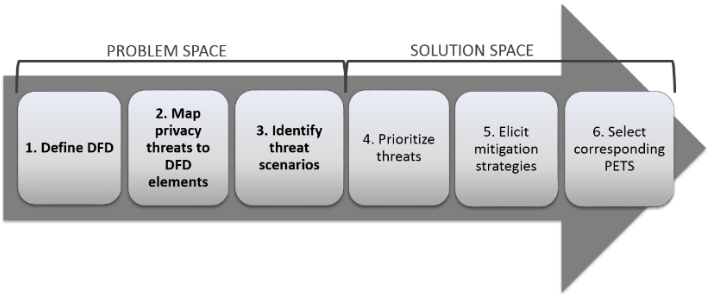

PROBLEM SPACE
SOLUTION SPACE
1. Define DFD
2. Map privacy threats to DFD elements
3. Identify threat scenarios
4. Prioritize threats
5. Elicit mitigation strategies
6. Select corresponding PETS

Figure 9. The LINDDUN methodology steps

Step 1 of the LINDDUN method implies the detailed system description with the recourse to major types of building blocks such as *external entities, data stores, data flows,* and *processes* (Figure 10).

[8] LINDDUN privacy threats modelling methodology, Available at: https://linddun.org/linddun.php# Last accessed on 17 April 2019.

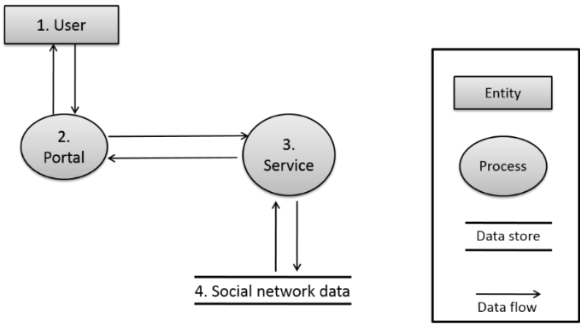
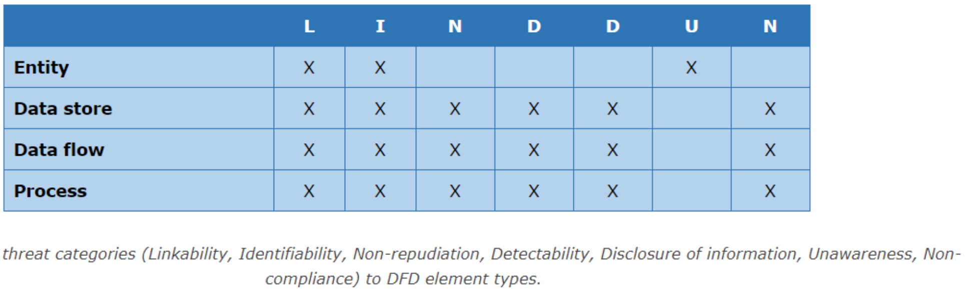


Figure 10. The data flow diagram (DFD) of the Social network data

Step 2 of the LINDDUN method entails creating a table corresponding to the case study where LINDDUN privacy threats are mapped to different blocks of the DFD (Figure 11).



	L	I	N	D	D	U	N
Entity	X	X				X	
Data store	X	X	X	X	X		X
Data flow	X	X	X	X	X		X
Process	X	X	X	X	X		X

Mapping LINDDUN threat categories (Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, Non-compliance) to DFD element types.

Figure 11. Mapping threat categories (Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, Non-compliance) to DFD element types.

Step 3 of the LINDDUN method comprises 3 substeps.

Substep 1 consists in examining each of the threat categories from the table above in order to determine whether they pose a threat to the system. It is done through the recourse to threat tree patterns (Figure 12. Example of LINDDUN threat tree with root threats (circles), concrete threats (boxes), AND relation, OR relation). In substep 2 all the branches, leaves and nodes of the tree are described and examined, where applicable. In substep 3 all other branches of the tree, which are not documented in step 2 should be explicitly documented as assumptions so that they could be easily tracked if there are any changes in privacy analysis results.

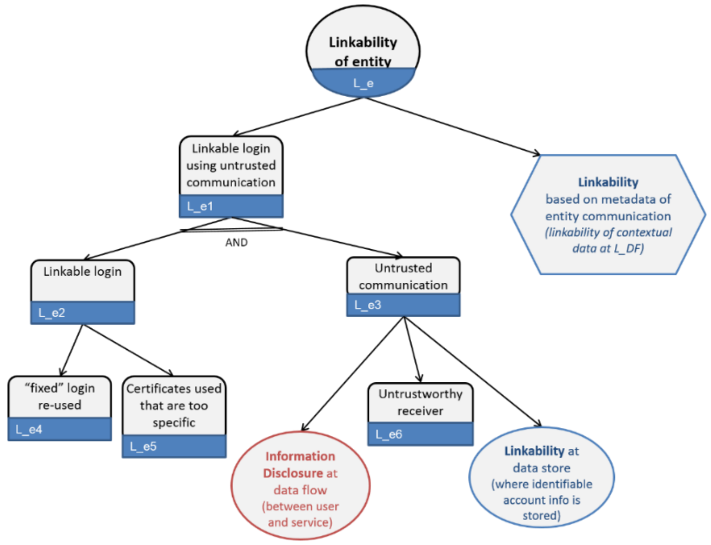

Linkability of entity
L_e
Linkable login using untrusted communication
L_e1
AND
Linkability based on metadata of entity communication (linkability of contextual data at L_DF)
Linkable login
L_e2
Untrusted communication
L_e3
"fixed" login re-used
L_e4
Certificates used that are too specific
L_e5
Untrustworthy receiver
L_e6
Information Disclosure at data flow (between user and service)
Linkability at data store (where identifiable account info is stored)

Figure 12. Example of LINDDUN threat tree with root threats (circles), concrete threats (boxes), AND relation, OR relation

In **step 4** all the privacy threats are assessed and evaluated via established risk assessment techniques.

In **step 5,** on the basis of the prioritization of risks established in step 4, the analyst determines appropriate *"mitigation strategies"* for each identified threat. *"Mitigation strategies"* refer to a wide spectrum of techniques addressing privacy threats. One possible taxonomy of strategies is proposed below. However, it does not aim at providing a complete overview of strategies and just shows a common set of mitigation strategies. Proactive approach to privacy is to ensure, for instance, that the user shares as little information as possible. And the reactive approach is in controlling and limiting the damage of the disclosure once it occurs.

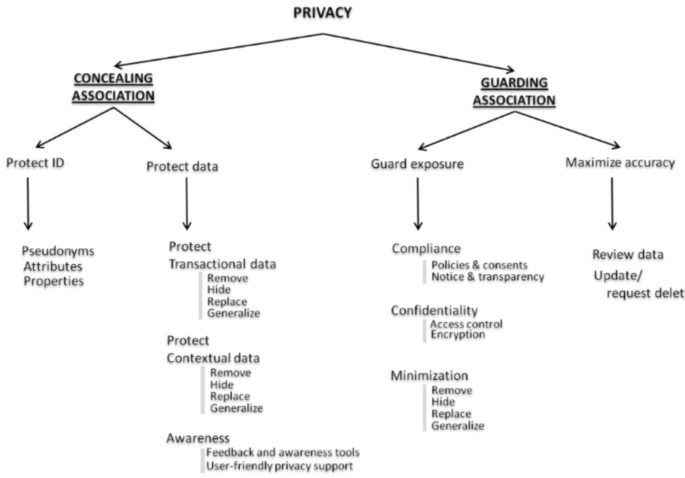


Figure 13. Taxonomy of Privacy Mitigation Strategies

In **step 6** mitigation strategies are translated into privacy requirements or solutions or are implemented directly as Privacy Enhancing Technologies that match with the mitigation strategies.

3.2. Rationale for aligning the GDPR and LINDDUN

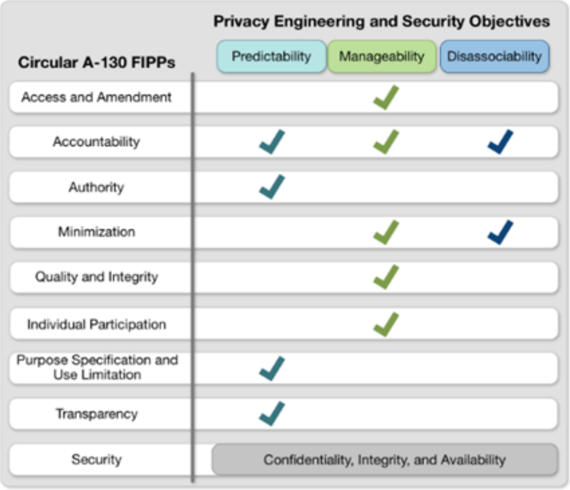
We will perform a preliminary exercise aimed at showing that despite the fact that the GDPR is a legal instrument and LINDDUN is an engineering method, they can be aligned in order to bridge the existing gap between legal and technical practices. This issue will be elaborated in a more detailed way in the next section.

The attempt to align LINDDUN and the GDPR replies to the demands of privacy engineering community of, first of all, translating complex legal texts into understandable by engineers principles/risks/threats. Secondly, this will contribute to the operationalisation of the GDPR, in particular in risk assessment process, and to ensuring that legal requirements do not live in total isolation from the practice.

A similar exercise was already performed in NISTIR 8062 [9], which sets out 3 privacy engineering objectives[9]:

- Predictability: Providing a reliable understanding about what is occurring with personal data processing within a system.
- Manageability: Administration of personal data with sufficient granularity so that the right level of control can be applied.
- Disassociability: Actively protect or "blind" an individual's identity or associated activities from unnecessary exposure during transactions.

NISTIR 8062 highlights that there is a correlation between these three objectives and 9 *Fair Information Practice Principles* (FIPPs)[10]. The FIPPs were proposed by the United States Federal Trade Commission as guidelines concerning fair information practice in an electronic marketplace. They can be considered as the foundation of all current data protection legislation. In addition, they correlate with the GDPR principles.



	Privacy Engineering and Security Objectives		
Circular A-130 FIPPs	Predictability	Manageability	Disassociability
Access and Amendment		✓	
Accountability	✓	✓	✓
Authority	✓		
Minimization		✓	✓
Quality and Integrity		✓	
Individual Participation		✓	
Purpose Specification and Use Limitation	✓		
Transparency	✓		
Security	Confidentiality, Integrity, and Availability		

Figure 14. Aligning the circular A-130 FIPPs to the Privacy Engineering and Security Objectives[10]

While analysing LINDDUN through the GDPR lens, one may draw the following conclusions.

1) Linkability (L), identifiability (I), detectability (D), and to some extent non-repudiation (Nr) are all pointing out to the existence of personal data, since the occurrence of one of these threats

[9] D2.2 PDP4E, Technical Gap Analysis and Synthesis of User Requirements, p. 13.

[10] NISTIR 8062 An Introduction to Privacy Engineering and Risk Management in Federal Systems. Available at: https://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8062.pdf

could lead to the identification of a natural person. According to the European legislation, the anonymous information does not require the compliance with the principles of data protection.[11] Anonymous data should be understood as information which does not relate to an identified or identifiable natural person.[12] However, "*in this era of big data, full anonymity is hard, if not impossible, and even more advanced anonymity techniques cannot guarantee full anonymity when data are linkable" [7].* The threat of linkability may necessitate a further analysis since it cannot be established without context whether the linkability of two items of interests would allow the identification of a natural person and, thus, qualify as the personal data.

2) Linkability might lead to identifiability (i.e. linking data to an identity). Once the data subject is identified or is identifiable, the information qualifies as personal data. And the application of the GDPR will be triggered if such data form part of a filing system[13] and we intend to process such data, while having been established in the EU and offering goods/monitoring EU residents.[14] Moreover, the GDPR will apply independently of the nature, content and format of the personal data.
3) Information disclosure threat could be linked with integrity and confidentiality principles under Article 5 of the GDPR. Personal data shall be processed in such a manner that ensures appropriate security of personal data, including protection against unauthorised or unlawful processing, accidental loss. Additional care should be taken in case of processing of sensitive data.
4) Unawareness threat could be linked to information requirements, meaning that the data subject must be given all the information about data processing activities. Unawareness could be also associated with non-respect for data minimisation requirement under the GDPR. You cannot collect more data than necessary and you should always opt for less intrusive means for achieving the same purposes. This will require creative thinking when dealing with big data projects.
5) Non-compliance threat could be associated with data protection by design requirement, accountability obligation under Article 24 GDPR, such as adopting appropriate technical and organisational measures ensuring the GDPR compliance or adopting internal privacy policies. For the most part we can speak about general GDPR non-compliance resulting in a pyramid of sanctions: from warnings to sanctions as a last resort.

3.3. Aligning LINDDUN threats categories with the GDPR vocabulary

This section provides the description of each LINDDUN threat type and its relation with the GDPR:

- **L**inkability (L)
- **I**dentifiability (I)
- **N**on-repudiation (Nr)
- **D**etectability (D)
- **D**isclosure of information (Di)
- **U**nawareness (U)

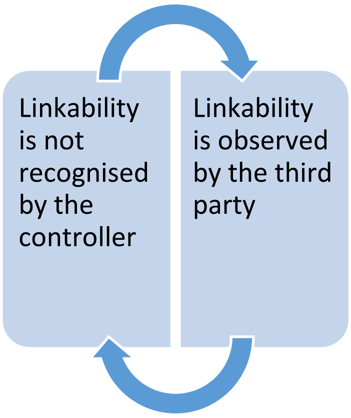
[11] Recital 26 GDPR.
[12] Ibid.
[13] Article 2 GDPR.
[14] Article 3 GDPR.

- **N**on-compliance (N)

3.3.1. Linkability

LINDDUN threat	Related GDPR principle	Related data subject right
Linkability = Being able to sufficiently distinguish whether 2 IOI (items of interest) are linked or not, even WITHOUT knowing the actual identity of the subject of the linkable IOI.	• Lawfulness, • Transparency, • Purpose limitation, • Data minimisation, • Storage limitation • Accuracy, • Integrity and Confidentiality, • Accountability	• Right to be informed, • Right of access, • Right to data portability • Right to rectification, • Right to be forgotten, • Right to restriction of processing, • Right to object, • Right not to be subject to a decision based solely on automated processing

Table 5. Description of Linkability under the GDPR lens


Linkability is not recognised by the controller
Linkability is observed by the third party

Linkability means *"being able to sufficiently distinguish whether 2 IOI (items of interest) are linked or not, even without knowing the actual identity of the subject of the linkable IOI"*[15]. Pfitzmann and Hansen give the following definition: *"unlinkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, etc.) 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."* [11] For instance, unlinkablity of a message sender/recipient to a message sent or received or relationship unlinkability between a sender and a recipient [11]. Unlinkability is one of prerequisites of anonymity. Nevertheless, failing unlinkability will not necessarily eliminate anonymity, but will decrease its strength [11].

From a legal perspective, linkability means that the failure to hide the link between different actions, identities or pieces of information could potentially result in the unexpected personal data processing (Table 5). For instance, the Article 29 WP provides for the following example: Titius has these fingerprints, this object has been touched by someone with these fingerprints and these fingerprints correspond to Titius, therefore this object has been touched by Titius [12]. Thus, linkability allowed to establish a link between one piece of information and the individual. The linking of different pieces of information can result in the misuse of the personal data by third parties. Such misuse can be caused by the failure to implement the necessary controls to ensure an appropriate level of protection of personal data (e.g., failed anonymization). If the controller is not aware itself of the personal data processing operation due to failed anonymization, it won't be able to comply with the GDPR data processing principles and, thus, will fail to ensure the respect for data subjects' rights. Thus, linkability

[15] LINDDUN privacy threats modelling methodology.

may result in the violation of a number of the personal data processing principles and of data subjects' rights listed in the GDPR.

First of all, the principle of lawfulness will be violated since there will be no lawful grounds for processing, as provided in article 6 of the GDPR. **Lawfulness** is deemed respected if the data subject has consented to the processing for specific purposes, if such processing is necessary for the performance of a contract or for compliance with a legal obligation, to protect the vital interests of the subject or of another natural person, or "*for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data*" and particularly when the data subject is a child.

Secondly, the principle of **transparency** will not be complied with, because **data subject will not be informed about the processing activities over their data.** The data subject might not be even aware at all that such personal data have been collected, used, consulted or otherwise processed and what is the extent of this processing.[16] Consequently, there will be no information provided relating to the processing of those personal data, in particular, on the identity of the controller and the purposes of the processing and further information to ensure fair and transparent processing.[17] Natural persons will not be made aware of risks, rules, safeguards and rights in relation to the processing of personal data and how to exercise their rights.[18]

Thirdly, **purpose limitation** principle will be also jeopardized since the controller, unable to establish the existence of the personal data, will not be able to ensure that the data collection is limited to "*specified, explicit and legitimate purposes*".[19] Moreover, in this case the controller will be collecting the personal data without knowing itself how and when these data will be used, since in its system the data is not identified as personal.

Moreover, the **data minimisation** and **storage limitation** principles will be also violated since the unawareness about the treatment of the personal data or its mere existence will not allow us to establish whether the same purpose can be achieved with a narrower collection of data and for a shorter retention period.

The inability to establish that the personal data exist in the system or that a third party can establish links between different pieces of information and, consequently, guess the existence of such data, will prevent us from ensuring that the data are accurate and kept up to date. As a result of this unawareness, controllers will not be able to ensure **accuracy** at all stages of collecting and processing of personal data and take every reasonable step to ensure that inaccurate data are erased or rectified without delay. Thus, contrary to the principle of accuracy, controllers will not make sure that outdated data are eliminated, or that data are correctly interpreted.

The compliance with the principle of **integrity** and **confidentiality** will be also jeopardized since the processing of the data, deemed as non-personal, will not be as secure as required for the personal data processing, "*including protection against unauthorised or unlawful processing and against accidental*

[16] Recital 39 GDPR.

[17] Ibid.

[18] Ibid.

[19] Article 5 (1) (b) GDPR

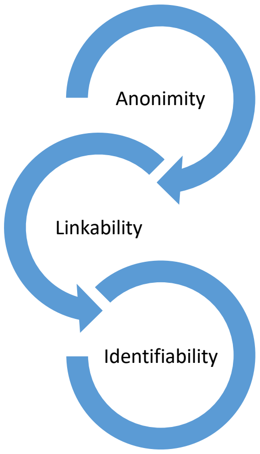
loss, destruction or damage, using appropriate technical or organisational measures"[20]. This will result in a lack of appropriate controls **to prevent unauthorised access** to the personal data as well as implement systemic quality controls in order to ensure that an appropriate level of security is reached. Moreover, the personal data will not be validated (e.g. using hashes), which might lead to some negative consequences, such as inability to guarantee its integrity and, consequently, the accuracy of that data.

According to the principle of **accountability**, the controller shall be responsible for, and be able to demonstrate compliance with, principles relating to processing of personal data and listed in Article 5 of the GDPR.[21] The non-respect for one of these principles will trigger the accountability obligation.

Since linkability in many cases is undetected because the personal data is not recognized as such and is not traceable in the system, the controller will not comply with information obligation, as substantiated in Articles 13-14. Thus, **data subjects will be deprived of the right to obtain information** about the processing activities over their data, the identity and the contact details of the controller, the purposes of the processing, the categories of the data and their recipients, and how the data are being controlled, monitored or used further.[22] The information obligation is the essential first step setting out the stage towards the exercise of other data subjects' rights. Neither **right of access**, nor **right to rectification** or **erasure** of personal data, nor **restriction** or **objecting** to their processing will be possible unless the data subject knows the personal data is processed by the controller.

3.3.2. Identifiability

"*Identifiability of a subject from an attacker's perspective means that the attacker can sufficiently identify the subject within a set of subjects*." [11] Identity can be explained and defined as the opposite of anonymity and the opposite of unlinkability [11]. In a positive wording, identifiability enables both to be identifiable as well as to link IOIs. The less is known about the linking to a subject, the stronger is the anonymity. The anonymity decreases with a growing linking [11].



The definition of identifiability provided in the technical literature seems not to be totally in line with the legal understanding of an identifiable natural person. While both the legal and technical literature recognise pseudonimisation as one of the techniques decreasing the likelihood of identifiability, the GDPR takes a stricter stance on pseudonimised data. For instance, Recital 26 GDPR sets out that "*pseudonimised personal data, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person*". And, thus, such data will be treated as personal under the GDPR, since pseudonym means that it is possible to backtrack to the individual and discover individual's identity. At the same time, the technical literature admits the flawlessness and high linkability potential of pseudonimised data, but still seems

[20] Article 5(1)(f) GDPR.
[21] Article 5(2) GDPR.
[22] Article 13 GDPR.

to treat pseudonimity as a concept in a slight opposition to identifiability [7]. "*Whereas anonymity and identifiability (or accountability) are the extremes with respect to linkability to subjects, pseudonymity is the entire field between and including these extremes*" [7].

LINDDUN threat	Related GDPR principle	Related data subject right
Identifiability = Being able to sufficiently identify the subject within a set of subjects (i.e. the anonymity set)	• Lawfulness, • Transparency, • Purpose limitation, • Data minimisation, • Accuracy, • Storage limitation, • Integrity, Confidentiality, • Accountability	• Right to be informed, • Right of access, • Right to data portability • Right to rectification, • Right to be forgotten, • Right to restriction of processing, • Right to object, • Right not to be subject to a decision based solely on automated processing

Table 6. Description of Identifiability under the GDPR lens

In addition the concept of identifiablity is not that straightforward. For instance, the GDPR provides a non-exhaustive list of identifiers in Article 4, such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. "*The natural person is "identifiable" when, although the person has not been identified yet, it is possible to do it*" [12]. But the likelihood of identifiability should be analysed on a case-by-case basis. For instance, a very common name will not necessarily allow to single out one particular person from the whole of a country's population [12], but can achieve the identification of a pupil in the classroom. In addition, the name, combined with some additional information can also allow the identification of someone as a result of this "unique combination" set. Even a very descriptive information about someone wearing a red hat can identify someone at the bus stop at a particular moment. Therefore, the identifiability depends on a case-by-case assessment and is context sensitive. For instance, a dynamic IP address was recognised as personal data by the ECJ (European Court of Justice) in Breyer case.[23] The ECJ held that "*even though the additional data necessary to identify the user of a website are held not by the **online media services provider**, but by that **user's internet service provider**, that dynamic IP addresses constitute personal data*".[24]

The identifiability is a dynamic process and, while it may not be possible to identify someone today with all the available means, it may happen at a later stage due to a technological progress. To determine whether an individual is identifiable, Recital 26 GDPR underlines, "*account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirect*". The likelihood of identification must be assessed in light of "*objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments*".

[23] Case C-582/14, Breyer, ECLI:EU:C:2016:779.

[24] Ibid.

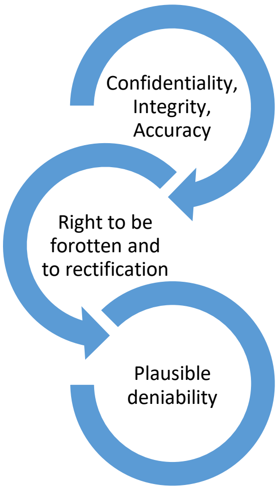
Since identifiability is closely related to linkability, it will affect the same GDPR principles and data subjects' rights (Table 6). Therefore, we decided not to provide a redundant explanation of the rationale behind each of them.

3.3.3. Non-repudiation

LINDDUN threat	Related GDPR principle	Related data subject right
Non-repudiation = Not being able to deny a claim. The attacker can thus prove a user knows, has done or has said something. He can gather evidence to counter the claims of the repudiating party.	• Integrity and Confidentiality, • Accountability, • Accuracy	Right to be forgotten and right to rectification

Table 7. Description of Non-repudiation under the GDPR lens

Non-repudiation is the opposite of plausible deniability. Plausible deniability from an attacker's perspective means that an attacker cannot prove a user knows, has done or has said something [7]. While the goal of the non-repudiation service is to provide irrefutable evidence concerning the occurrence or non-occurrence of an event, it must be admitted that some participants may desire that there is no irrefutable evidence concerning a disputed event or action [7]. Wuyts provides for some concrete examples where non-repudiation is a privacy threat. For instance, e-commerce applications, where the vendor can later use the signed receipt by the buyer as evidence that the user received the item. For other applications similarly users may desire plausible deniability in order to ensure that there will be no record to demonstrate the communication event [7].


Confidentiality, Integrity, Accuracy
Right to be forotten and to rectification
Plausible deniability

In an attempt to single out the most linkable GDPR principles (Table 7) with non-repudiation, we came to the conclusion that non-compliance with **integrity** and **confidentiality** requirements might lead to the loss of control over the personal data and increase the probability that it can be accessed by unauthorized parties. Logically, the controller will be held accountable for such incidents and for lack of appropriate confidentiality strategies. We consider that **right to be forgotten** and **right to rectification** are intrinsically linked with plausible deniability, since they allow for ex ante rectification of the personal data inaccuracies and the possibility to ask for erasure of those data, which are no longer necessary for the purposes for which it was collected or where such purpose ceases to exist, or where the data subject withdraws consent on which the processing is based.[25] Thus, **right to be forgotten** and **right to rectification** will prevent a priori the third parties from getting access to the information, which the data subject considers as inaccurate or compromising. Nevertheless, as provided in Article 17 GDPR some exceptions might apply to the exercise of the right to erasure, including the situations where there is a need to strike a right balance

[25] See Article 17 of the GDPR for more examples.

between public interests, freedom of expression and other competing rights and legitimate interests. In addition, Deng et al. notes with regard to plausible deniability that it ensures that “an instance of communication between computer systems leaves behind no unequivocal evidence of its having taken place” [13]. Thus, in relation to the right to be forgotten and right to rectification, one might ask whether the controller should store requests for personal data erasure or rectification. And wouldn’t such storage be detrimental to plausible deniability? Thus, the right balance should be again struck between accountability obligations and data subjects’ legitimate interests.

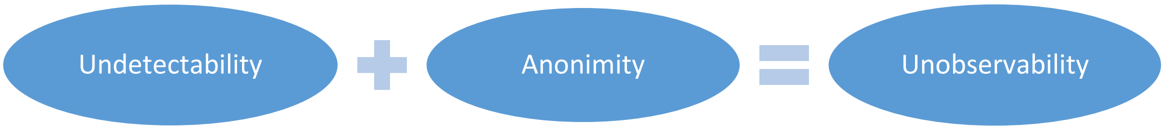
In addition, in order to guarantee plausible deniability the data controller may decide to make the data less accurate to “cover user’s tracks”. While the GDPR requires to keep the personal data up to date and ensure that inaccurate data are erased or rectified without delay[26], plausible deniability may require a different approach towards accuracy. On one hand, the accuracy of personal data should not be compromised, on the other hand, making personal data less discernible from the outside may be necessary for ensuring plausible deniability.

3.3.4. Detectability

LINDDUN threat	Related GDPR principle	Related data subject right
Detectability = Being able to sufficiently distinguish whether an item of interest (IOI) exists or not (e.g. by knowing that a celebrity has a health record in a rehab facility, you can deduce the celebrity has an addiction, even without having access to the actual health record)	• Lawfulness, • Transparency, • Purpose limitation, • Data minimisation, • Accuracy, • Storage limitation, • Integrity, Confidentiality, • Accountability	• Right to be informed, • Right of access, • Right to data portability • Right to rectification, • Right to be forgotten, • Right to restriction of processing, • Right to object, • Right not to be subject to a decision based solely on automated processing

Table 8. Description of Detectability under the GDPR lens

“*Undetectability of an item of interest (IOI) from an attacker’s perspective means that the attacker cannot sufficiently distinguish whether it exists or not. If we consider messages as IOIs, this means that messages are not sufficiently discernible from, e.g., random noise” [11].* The difference between unlinkability and undetectability is the following: in unlinkability, the IOI itself is not protected, but only its relationship to the subject or other IOIs is protected. For undetectability, the IOIs are protected as such [7]. Undetectability consists in, for instance, hiding the user’s activities or location [7]. Undetectability in the past was referred as unobservability. However, since unobservability comprises both anonymity and undetectability, LINDDUN method focuses on undetectability.



Detectability threat is strongly related to the context. It is impossible to establish without further

[26] Art. 5(1)(d) GDPR.

details whether detectability of one particular activity can lead to identifiability of an individual. But if we assume that detectability results in an identifiability of a natural person, the scope of the GDPR will be triggered in a similar way to linkability and identifiability (Table 8).

3.3.5. Information Disclosure

LINDDUN threat	Related GDPR principle	Related data subject right
Information Disclosure	- Lawfulness, - Transparency, - Purpose limitation, - Data minimisation, - Accuracy, - Storage limitation, - Integrity, Confidentiality, - Accountability	- Right to be informed, - Right of access, - Right to data portability - Right to rectification, - Right to be forgotten, - Right to restriction of processing, - Right to object, - Right not to be subject to a decision based solely on automated processing

Table 9. Description of Information Disclosure under the GDPR lens

Information Disclosure is the exposure of information to individuals who are not supposed to have access to it. Principles of integrity and confidentiality will be the most relevant to guarantee the security of the personal data processing. While Wuyts considers confidentiality as a security property, she empathises also its importance for preserving privacy properties, such as anonymity and unlinkability [7].

Similarly to linkability, information disclosure will also trigger all personal data processing related principles, since the data could be further collected, stored by third parties without specific purpose and without informing the data subject. Thus, data minimisation and storage limitation principles cannot be complied with either. In addition, the accuracy of the personal data can be also jeopardized (Table 9).

3.3.6. Unawareness

LINDDUN threat	Related GDPR principle	Related data subject right
Unawareness = Being unaware of the consequences of sharing information	- Fairness, - Transparency, - Data minimisation, - Accuracy, - Lawfulness, - Purpose limitation, - Accountability	- Right to be informed, - Right of access, - Right to data portability - Right to rectification, - Right to be forgotten, - Right to restriction of processing, - Right to object, - Right not to be subject to a decision based solely on automated processing

Table 10. Description of Unawareness under the GDPR lens

Unawareness occurs when a user is unaware of the information he/she is supplying to the system, and the consequences of his/her acts of sharing. In the era of digitalisation users tend to provide excessive information resulting in a loss of control of their personal information. Thus, awareness aims at ensuring that users are aware of their personal data and that only the minimum necessary information should be collected [7].

Unawareness points out to the violation of **fairness and transparency** requirements, since the data subject is not informed of all the risks related to the personal data processing and was not provided all the information required in relation to their personal data processing (Table 10). Transparency principle if further substantiated in Articles 13-14 GDPR referring to **information obligation** of controllers. Unawareness also leads to the fact that the data subject provides more personal information than required, and thus, the principle of **data minimisation** is violated [7]. According to **purpose limitation** principle, personal data should *collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes.* This correlates with The Platform for Privacy Preferences Project, as noted by Wuyts, which has been designed to allow websites to declare their intended use of the collected personal data [7]. In addition, since the data subject is not aware of some data processing activities, he/she is not able to ask for the information to be updated, which jeopardizes the **accuracy** of information [7]. Right to be informed together with the right of access constitute core prerequisites for the exercise of all other prerogatives granted to data subjects, in particular **right to data portability, right to rectification, right to be forgotten, right to restriction of processing, right to object, right not to be subject to a decision based solely on automated processing**. The detailed description of each of these rights can be found in PDP4E Deliverable 2.1.

Social Network Sites (SNSs) like Facebook or Twitter introduce additional challenges to the ones of fairness and transparency mentioned above. On their core, these platforms are spaces in which users make their private information publicly available to a large group of people. That is, users share different aspects of their lives with large and diverse audiences through posts, photos, videos and other type of media content. Although this is a common practice in the real world (people revel aspects of their private life to stablish and maintain social connections), in SNSs audiences are larger and harder to estimate by regular users. Consequently, private information sometimes reaches untrusted recipients causing unwanted incidents such as identity theft, reputation damage or financial fraud. Although privacy scholars have reported evidence in which users regret having shared personal information in SNSs concrete measures seem not to have been taken yet. Many argue that, like in the real world, risk information would help users making better and more informed privacy decisions. This could be done by following a similar approach to the one used by Health Warning Labels in cigarette packages or Nutrition Labels in food products. However, not much efforts have been made by SNSs to introduce mechanisms that inform the potential privacy risks of information sharing. Conversely, privacy researchers have already proposed awareness mechanisms for SNSs like Facebook that aim at supporting users in information disclosure activities within these platforms. Such mechanisms include wizards for defining access-control policies and the definition of risk patterns.

3.3.7. Non-compliance

LINDDUN threat	Related GDPR principle	Related data subject right
Non-compliance = Not being compliant with legislation, regulations, and corporate policies.	• Lawfulness limited to consent, • Transparency, • Accountability	All the existing legal frameworks are triggered

Table 11. Description of Non-compliance under the GDPR lens

Non-compliance is related to legislation, policy and consent and implies that the data subject should be informed by the controller about the system's privacy policy and allows the data subject to specify consents [7]. Wuyts gives some examples of such non-compliance, such as incorrect privacy policies provided to the user or when the policy rules are incorrectly managed by the system administrator [7].

Wuyts notes that policy specifies one or more rules with respect to data protection and these are general rules determined by the stakeholders of the system; a consent specifies one or more data protection rules and is determined by the user and only relate to the data regarding this specific user [7]. From the legal perspective, while the processing of personal data can be based on data subject's consent, lawfulness of the processing is not limited to consent compliance. The GDPR provides for 5 additional legal grounds where the processing of personal data is not based on consent: the performance of a contract, a legal obligation, the vital interests of individuals, the public interest and the legitimate interest of the controller. Thus, the personal data can be processed without data subject's consent if it relies on some other legal grounds.

When it comes to policy, Wuyts emphasizes the compliance with internal policies of the company. However, the compliance with internal policies of the company will not be enough if those policies are not correct, lack detail or are not user friendly with regard to privacy notices provided. Thus, non-compliance with policies should be related to broader issues covering also some external requirements and legal framework applying to controllers (Table 11).

Non-compliance threat, as described in LINDDUN, seems to be too generic and lacks in precision. Its current wording suggests that all the data protection related legal frameworks will be triggered. However, eliminating this threat is easier said than done, since the legal compliance is not an easy exercise.

Some further complexities of non-compliance threat will be provided in Annexe A. In Annex B we will proceed with the non-compliane risk identification through the negation of the GDPR provisions.

3.3.8. Conclusion

The connection between the GDPR and LINDDUN threat categories is very large since they rely on different vocabulary. This interdisciplinary exercise was an attempt to bridge the existing gap between the legal approach towards privacy risks and engineers approach towards privacy risks. The way to mitigate all the complexities of the tooling will need to be discussed at a later stage depending on the feedback received after the first iteration.

Annex A Extending LINDDUN methodology

This section identifies a number of gaps in Non-compliance threat, as described in LINDDUN methodology (1.1), and further provides a list of additional elements necessary for bridging this gap (1.2).

1.1. Rationale for extending LINDDUN

Non-compliance is mentioned as one of the threat categories under LINDDUN framework. Even though LINDDUN is not a compliance technique, it explicitly draws attention to the need of regulatory compliance. However, the wording of this threat is too generic and refers to the whole complexity of legal frameworks and policies. Thus, leaving the notion of non-compliance in its current vagueness and obscurity, will deprive non-compliance threat of its substance and make its analysis with regard to DFDs mapping extremely complex. Analysing the threat of non-compliance is not sufficient if it does not come along with technical and concrete measures to protect privacy and personal data in practice.

In addition, non-compliance under LINDDUN is limited to consent requirement. Even though the consent does constitute a legal basis for the personal data processing, it is not the only possible legal ground in this regard.[27] Therefore, it is not clear to the reader from the wording of non-compliance threat where the necessity to single out the consent issue comes from. Moreover, the consent requirement under LINDDUN framework does not meet the definition of consent, as provided in Article 4 GDPR, "*'consent' of the data subject means any* ***freely given, specific, informed and unambiguous*** *indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her*"[28]. LINDDUN fails to provide further details on the properties of the consent, notably that it should be given freely, in a specific manner, clearly and after the data subject was informed of the processing activities.

Moreover, LINDDUN does not cover fully purpose related requirements, which constitute a core prerequisite for deciding on other data processing related aspects, such as data quality requirements, relevance, proportionality, data minimisation, accuracy of the data collected and its retention period. While examining the interplay between **Solove's Taxonomy[29] and LINDDUN,** Wuyts notes that the use of the data for a different purpose, so-called "secondary use" under **Solove's Taxonomy,** is not considered in LINDDUN explicitly as it is closely related to data protection compliance. Wuyts further elaborates on this by stating the rule: "only use and share data if the data subject has consented to the specific purpose". It is not completely clear why Wuyts eliminates purpose from the scope of LINDDUN, and in particular with regard to Non-compliance threat, motivating this decision by its (purpose) too compliance oriented nature. While one agrees that purpose limitation principle will necessarily increase compliance with the legal framework and some GDPR principles notably, it seems difficult to understand the reasons why compliance is aimed at and avoided at the same time.

Thus, non-compliance under LINDDUN in its current status will be pointless if it is not further operationalized and extended with some GDPR requirements elaborated in the next section.

[27] See Article 6 GDPR for more information.

[28] Article 4 (11) of the GDPR

[29] Solove presents a taxonomy of privacy violations from a legal perspective.

1.2. Specification of LINDDUN non-compliance threat

This section provides an overview of the GDPR based threats deemed relevant for extending and specifying Non-compliance threat as referred in LINDDUN. As stated previously, non-compliance under LINDDUN is a catch-all threat, which covers everything and nothing at the same time. Therefore, non-compliance under LINDDUN shall be specified in a detailed manner and in connection with the GDPR, which entered into force almost one year ago. The aim of extension of this non-compliance issue is not to ensure the compliance with the whole GDPR text, but with some singled out issues deemed the most relevant in the framework of the software development life-cycle, such as lawful ground, purpose limitation, data subject categories and personal data categories. This version might be subject to further changes based on the feedback received after the first iteration.

LINDDU**N** (+4U)

Unlawful ground

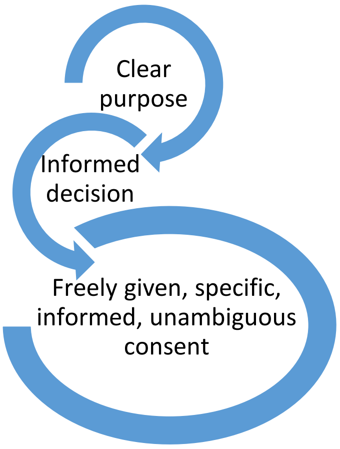
Undefined purpose

Undetected data subjects categories

Undetected personal data categories

1.2.1. Unlawful ground

Unlawful ground is the opposite of lawfulness and means that personal data are not processed by controller based on one of the **legal grounds** listed in the in Article 6 GDPR, such as (1) the consent, (2) the performance of a contract, (3) a legal obligation, (4) the vital interests of individuals, (5) the public interest and (6) the legitimate interest of the controller.



Consent means "*any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her*". If the data subject's consent is requested by electronic means, this request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided.

Consent and purpose are intrinsically related, since transparent and simple explanation of the purpose(s) of the processing of personal data allows a data subject to make an informed decision [14].

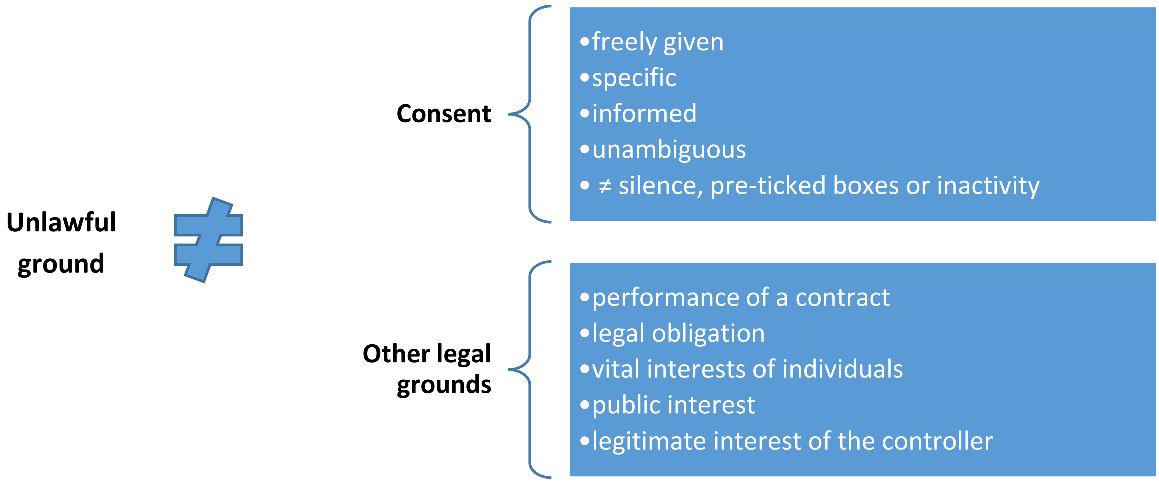

Unlawful ground
Consent
•freely given
•specific
•informed
•unambiguous
• ≠ silence, pre-ticked boxes or inactivity
Other legal grounds
•performance of a contract
•legal obligation
•vital interests of individuals
•public interest
•legitimate interest of the controller

Figure 15. Unlawful ground

1.2.2. Undefined purpose

Undefined purpose stands for the negation of purpose related requirements set out in Article 5(1)b GDPR: personal data shall be "*collected for **specified, explicit and legitimate purposes** and **not further processed in a manner that is incompatible** with those purposes".* Thus, the **undefined purpose** violates two main building blocks of purpose limitation principle: personal data must be collected for "specified, explicit and legitimate" purposes (**purpose specification**) and not be "further processed in a way incompatible" with those purposes (**compatible use**) [15]. First, specification of purpose is a core prerequisite for deciding on other data processing related aspects, such as data quality requirements, relevance, proportionality, accuracy of the data collected and its retention period [15]. Secondly, the principle of purpose limitation prevents the usage of the available personal data beyond the purposes for which they were initially collected. However, this does not rule out new, different uses of the data, if the parameters of compatibility are respected. Thus, principle of purpose limitation aspires to reconcile the need for "*legal certainty regarding the purposes of the processing on one hand, and the pragmatic need for some flexibility on the other*" [15].

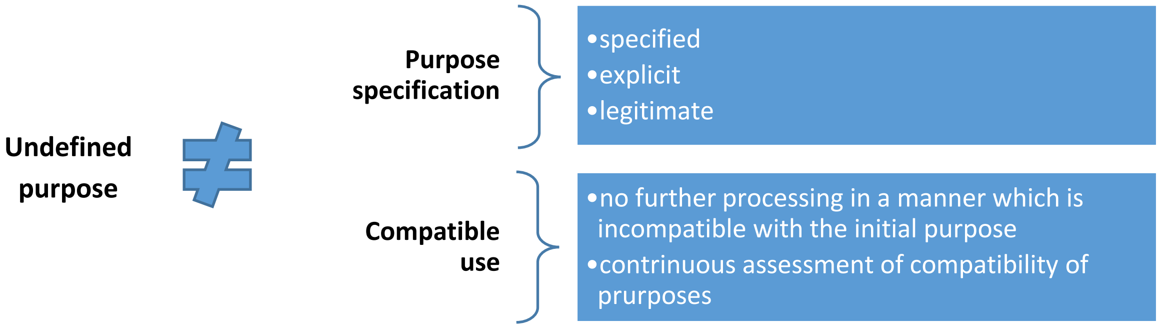

Undefined purpose
Purpose specification
•specified
•explicit
•legitimate
Compatible use
•no further processing in a manner which is incompatible with the initial purpose
•contrinuous assessment of compatibility of prurposes

Figure 16. Undefined purpose

First, any purpose must be **specified** prior to, and not later than, the time when the collection of personal data takes place [15]. Then the purpose of the collection must be detailed enough to understand what kind of processing is included within the specified purpose, and what data protection safeguards should be applied. At the same time, there is no need to overdo and provide anti user-friendly more detailed specifications. The approach of a "layered notice" to data subjects has been recommended in many situations by the WP29 [15]. If personal data is collected for more than one purpose, each separate purpose should be specified in enough detail to be able to assess the compliance with the law [15]. If processing operations relate to each other, the concept of an overall purpose, can simplify the task. However, the “overall purpose” practice should not be abused where processing operations are only remotely related to the initial purpose [15].

The purposes of collection must be **explicit** in order to ensure that there is no vagueness or ambiguity as to their meaning or intent. In other words, the specification of the purposes must be understood in the same way not only by the controller, but also by the data protection authorities and the data subjects concerned, irrespective of their different cultural/linguistic backgrounds [15]. This requirement contributes to transparency and predictability, reduces the risk that the data subjects' expectations will differ from those of the controller and allows data subject to take informed decisions.

Personal data must be collected for **legitimate** purposes. This requirement implies that the processing, of personal data in addition to the compliance with Article 6 GDPR requirements related to legal grounds, must be in accordance with the law, including data protection law along with other applicable laws such as employment law, contract law, consumer protection law.

Compatible use or prohibited incompatibility means that any further processing is authorised as long as it is not incompatible, provided the requirements of lawfulness are simultaneously fulfilled. Further processing refers to any processing operation occurring after the initial data collection stage. The fact that the further processing is for a different purpose does not necessarily mean that it is automatically incompatible and needs to be assessed on a case-by-case basis [15]. The legislators provided for some flexibility with regard to further use in order to allow for a better adjustment to the expectations of society or to situations when a need for an additional purpose was not detected by the controller nor data subject at the initial stage [15]. Thus, in some situations, a change of purpose may be permissible, provided that the **compatibility test is satisfied**.

Several purpose compatibility criteria are listed in in Recital 50 GDPR, notably: (1) the relationship between the purposes for which the data have been collected and the purposes of further processing, (2) the context in which the data have been collected and the reasonable expectations of the data subjects as to their further use, (3) the nature of the personal data and the impact of the further processing on data subjects and (4) the safeguards applied by the controller to ensure fair processing and to prevent any undue impact on the data subjects. It should be noted that *“further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes*”. However, the notions of scientific and statistical research are not clearly defined in the GDPR, which leaves considerable doubts as to the scope of that provision

1.2.3. Undetected data subject categories

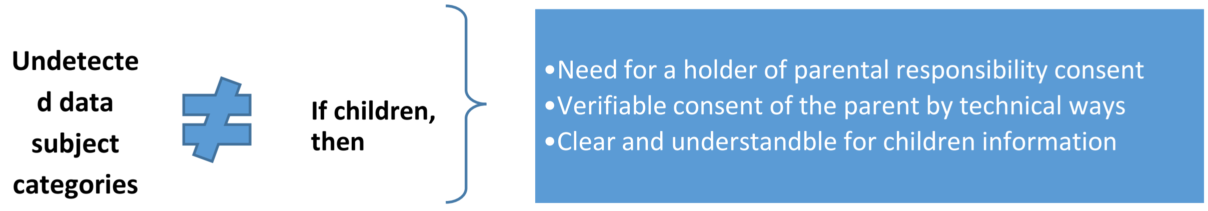

Undetecte d data subject categories
If children, then
•Need for a holder of parental responsibility consent
•Verifiable consent of the parent by technical ways
•Clear and understandble for children information

Figure 17. Undetected data subject categories

Undetected data subject categories is the opposite of the system enabled to detect when the data collected belongs to a child. Contrary to the Directive 95/46/EC, which did not contain any child-specific provisions, under the GDPR data controllers have to comply with a set of legal requirements for processing personal data of children [16]. This specific attention to children's personal data processing replies to the necessity to address the increased "datification" of children's lives. The Working Party emphasised on multiple occasions that the processing of children's personal data requires extra care and should comply with the principles of **data minimisation** and **purpose limitation** in a more stringent way [16].

Article 8 of the GDPR sets out the requirement for the **consent of the holder of the parental responsibility** in case of provision of information society services to children, if consent is the legal basis for the processing, as provided in Article 6(1a). Thus, controllers should make sure that they are able to **recognise children's personal data** and treat it in accordance with the GDPR provisions. In this regard, the controller shall ensure that its system has all the **necessary verification means and methods** to reasonably prove that the person providing consent is the parent of the child. However, the compliance with the consent requirement can be extremely difficult due to the age threshold divergences across the EU, since Article 8 does allow Member States to lower the age threshold of 16 years to a minimum of 13 years. This means in practice that different system requirements shall be implemented for different member states based on their national laws on the age limit. Moreover, it is not clear yet whether the data controller shall obtain fresh consent, when the child reaches the age of consent [16]. In this regard, the Article 29 Working Party provided that "*if the processing of a child's data began with the consent of their legal representative, the child concerned may, on attaining majority, revoke the consent. But if he wishes the processing to continue, it seems that the data subject need give explicit consent wherever this is required*." [17]

For PDP4E pilots, smart grids and connected/autonomous cars, special attention should be paid to circumstances in which personal data of children are processed in order to create personal or user profiles. Such practice is explicitly acknowledged as requiring extra protection [16]. It was not clarified in the GDPR what this extra protection entails in practice though. Moreover, a measure evaluating personal aspects relating to a data subject that is based solely on automated processing **should not concern children.** However, this is only prohibited as far as a decision produces legal effects for the child.[30] The golden rule shall be to adopt data minimisation as soon as the system detects the collection and use of the personal data for profiling, if such data belongs to a child. Otherwise, children's right to experiment and critically reflect upon their interactions risk to be undermined in the digital environment [16]. The children's right to explore and experiment with their identity can be further substantiated via **the right to be forgotten**. The GDPR empathizes its particular relevance for a child, who has given his or her consent and was not fully aware of the risks involved by the processing, and later wants to remove such personal data.

This table (Table 12) represents a detailed overview of all the children-specific provisions of the GDPR and is meant to help to adopt additional safeguards, when children's personal data is collected.

Children-specific elements In the GDPR	**Explanation of the GDPR provision**
Definition of the notion of a child	• No definition of who is a child

[30] Recital 71 GDPR.

	• Not clear until what age childhood lasts • The broad interpretation of children as under-18s was criticised as being unable to take into account the evolving capacities of children, and their level of maturity, in exercising their rights [16]
Specific protection (Recital 38 GDPR)	Children merit **specific protection** with regard to their personal data, as they may be less aware of the risks, in relation to the processing of personal data.
Cases of application of specific protection for children (Recital 38 GDPR)	• for the purposes of marketing • creating personality or user profiles • collection of personal data with regard to children when using services offered directly to a child
Child's consent in relation to information society services (Article 8)	• Where the child is below the age of 16 years, such processing shall be lawful only if that **consent is given by the holder of parental responsibility** over the child • The controller shall **verify** that consent is given or authorised by the holder of parental responsibility, taking into consideration available technology. • The consent of the holder of parental responsibility should not be necessary in the context of preventive or counselling services offered directly to a child.
What is information society service?	Any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services.[31]
Challenges of compliance	Article 8 does allow Member States to lower the age threshold of 16 years to a minimum of 13 years -> different age thresholds would apply throughout the EU
Information obligation (Recital 58) with regard to children	When provided to children, the information should be formulated in "*such **a clear and plain language** that the child can easily understand*"
Decision based on automated processing with regard to children (Recital 71)	A measure evaluating personal aspects relating to a data subject that is based solely on automated processing **should not concern children.** This is only prohibited as far as a decision produces legal effects for the child.
Right to be forgotten with regard to a child (Recital 65)	That right is **relevant in particular where the data subject has given his or her consent as a child** and is not fully aware of the risks involved by the processing, and later wants to remove such personal data, especially on the internet. The data subject should be able to exercise that right notwithstanding the fact that he or she is no longer a child.

Table 12. Overview of child-specific provisions

1.2.4. Undetected personal data categories

Undetected personal data categories threat refers to the system malfunction, which does not allow to detect whether the personal data collected is sensitive, related to criminal convictions and offences or just "normal" personal data. This issue is crucial for deciding upon the implementation of some

[31] Article 4 (25) GDPR refers to 'information society service' as "*a service as defined in point (b) of Article 1(1) of Directive (EU) 2015/1535 of the European Parliament and of the Council*".

additional safeguards in order to ensure a level of protection appropriate for each personal data category. Moreover, the personal data type impacts on whether the processing of the personal data can take place or not. For instance, the processing of sensitive data or data related to criminal convictions is prohibited in principle. Nonetheless, Article 9(2) and 10 establishes a number of exceptions to that prohibition, for instance when authorised by EU or MS laws. Thus, the exception to the general prohibition on the processing of the sensitive data is not only required to fall under one of the exceptions listed in Article 9(2) GDPR, but also to rely on one of the legal grounds specified in Article 6(1) GDPR. Sensitive data encompasses personal information "*revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation*", as provided in Article 9.

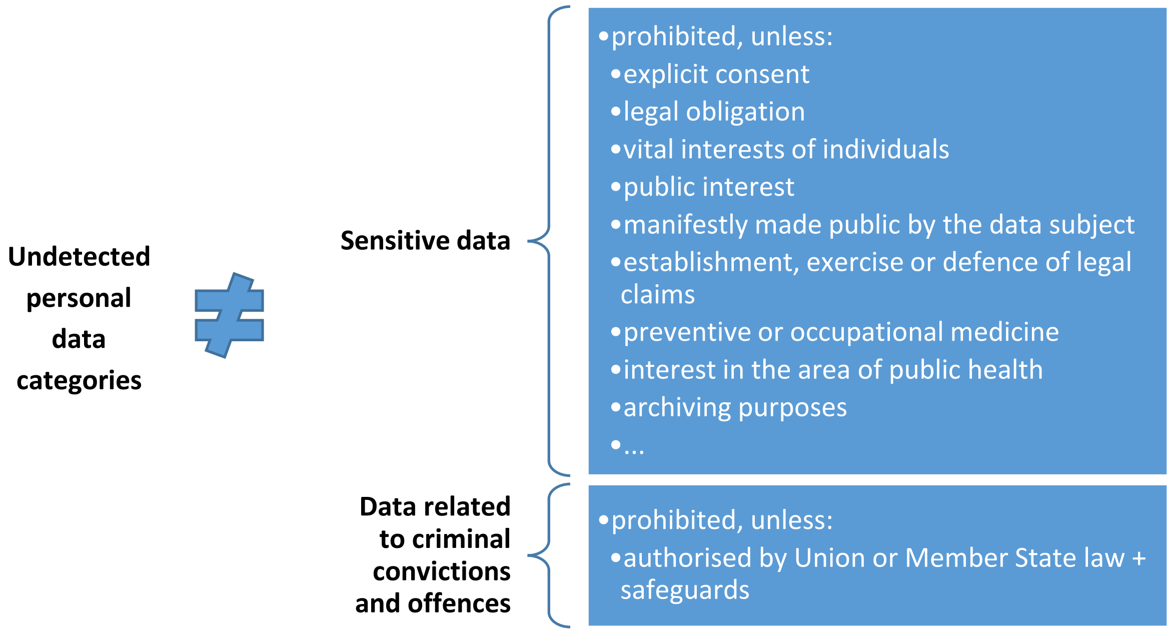

Undetected personal data categories
≠
Sensitive data
•prohibited, unless:
•explicit consent
•legal obligation
•vital interests of individuals
•public interest
•manifestly made public by the data subject
•establishment, exercise or defence of legal claims
•preventive or occupational medicine
•interest in the area of public health
•archiving purposes
•...
Data related to criminal convictions and offences
•prohibited, unless:
•authorised by Union or Member State law + safeguards

Figure 18. Undetected personal data categories

Annex B Conclusions with regard to risk identification under Extended LINDDUN(+4U)

In this section, an attempt was made to break down some GDPR provisions into potential risk scenarios. While some GDPR principles and related data subject rights can be mapped to DFDs, some more elaborate GDPR requirements can be difficultly accommodated even in the meta-model for the data protection architectural viewpoint, as suggested by Sion, Dewitte et al [18]. For instance, as such purpose can be caught per se, but its specified, explicit and legitimate nature cannot be accommodated. The same problem occurs with regard to lawful ground. The consent can be registered by the system, but it is more difficult to deal with other lawful grounds such as legitimate interests of controllers, vital interests of individuals, legal obligation, etc.

Consent-related risk scenarios	**PDP4E Risk Management Relevance**
Risk of not having a consent for a processing operation because of: a. Incorrect management of the record of consents and information provided at the time of the consent b. Incorrect identification of processing purposes	Incorrect management of the record of consents and information provided at the time of the consent
Risk of misusing consent as a backup option	Out of the scope

Risk of not having specific consent : failing to pair the consent with the purpose
a. Multiple processing ops -> one purpose (*many-to-one*);
b. Multiple processing ops -> multiple purposes (*many-to-many*);
c. One processing op -> multiple purposes (one-to-many).
	Out of the scope

Risk of not having informed, unambiguous consent: = Non-respect for information obligation	Depends on the risk source, if the data subject does not understand the information, then it will apply
Risk of having the consent of a wrong person (failed verification threshold)	Relevant
Risk of not having freely given consent : a. Because power imbalance between data subject and controller	Relevant
Other lawfulness-related risk scenarios	**PDP4E Risk Management Relevance**
Contract 1. Risk of collecting more than necessary 2. Risk of linking the collection of data to the contract where it is not necessary (Art.7(4))	Contract.1: During process (re-)engineering, not realizing that you are collecting more info than necessary.

	Engineering not following protocol. Contract.2: Borderline. Some mitigation actions can be implemented by engineering teams (e.g. policy stating that collection forms need to be reviewed by a peer).
Legal obligation 1. Risk of it ceasing to exist 2. Risk of it changing over time	Not relevant (Engineers need to revise the system and business environment periodically)
Legitimate interests Risk of an incorrect case-by-case assessment and balance against data subject rights	Not relevant (New condition to trigger risk analysis besides DPIA)
Data subject related risk scenarios	**PDP4E Risk Management Relevance**
Non-identifying a child?	Relevant (Weak authentication)
Failed verification threshold of a consent giver	Relevant (Weak authentication)
Misinterpretation/non-compliance of/with “specific protection” requirement as result of non-identification of a child	Relevant (Consequence of “non-identiyfying a child”)
Wrong assessment with regard to clarity of privacy policy to a child	Related to (consent) transparency
Taking automated decisions producing legal effects with regard to children as a result of wrong data subject categories assessments	Related to: Negative consequence to the data subject due to an unfair/unlawful automated decision. (Not only related to children)
Purpose related risk scenarios	**PDP4E Risk Management Relevance**
Incorrect assessment of the amount of data to be collected	Not relevant (Data minimization)
Incorrect assessment of purposes	Risk of wrong assessment during design (Purpose limitation)
Incorrect purposes compatibility assessment	Risk of wrong assessment during design (Purpose limitation)
Change of a purpose	Not relevant, needs to be addressed at a project management stage
Data categories related risk scenarios	**PDP4E Risk Management Relevance**
Failed anonymization	Relevant (also risk of wrong assessment of the PET – techniques does not work 100%)
Personal data is not recognized as such	Risk of wrong assessment during design Unknown external sources that identify DS.
Special categories of personal data are not recognized	Risk of hidden, or not so known, correlations between collected personal data and special categories. E.g. Postal code is related to ethnicity in some cities.

	Risk of wrong assessment during design (special categories are listed by the GDPR or supervisory authorities)
Unlawful processing of special categories of personal data	Not relevant, needs to be addressed at a project management stage
Incorrect balancing of interests in case of the sensitive data processing	Difficult to implement, falls under meta-risk category. It will be addressed through continuous risk management.
Explicit consent for the processing of the sensitive data is provided by a wrong person	Relevant

4. References

[1] LINDDUN privacy threats modelling methodology, Available at: https://linddun.org/linddun.php# Last accessed on 17 April 2019.

[2] Article 29 Working Party, Statement on the role of a risk-based approach in data protection legal framework, 30 May 2014. Available at: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp218_en.pdf Last access on 18 April 2019, p.2-3.

[3] Raphaël Gellert, “Understanding the Notion of Risk in the General Data Protection Regulation,” *Computer Law & Security Review,* 34, no. 2 (April 1, 2018), pp. 279–88.

[4] Article 29 Working Party,Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is “likely to result in a high risk” for the purposes of Regulation 2016/679, 17/EN, WP 248.

[5] CNIL Privacy Impact Assessment Methodology, February 2018 edition, p. 6. Available at: https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-en-methodology.pdf. Last accessed on 17 April 2019. Last accessed on 19 April 2019.

[6] Felix Bieker, Michael Friedewald, Marit Hansen, Hannah Obersteller, and Martin Rost, “A Process for Data Protection Impact Assessment under the European General Data Protection Regulation”, 4th Annual Privacy Forum, APF 2016, Frankfurt/Main, Germany, September 7–8, 2016 Proceedings.

[7] Kim Wuyts, Privacy Threats in Software Architectures, Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Engineering, January 2015, KU Leuven.

[8] Lund, M.S., B. Solhaug, and K. Stølen, Model-driven risk analysis: the CORAS approach. 2010: Springer Science & Business Media.

[9] Sean Brooks et al., “An Introduction to Privacy Engineering and Risk Management in Federal Systems”, Gaithersburg, MD: National Institute of Standards and Technology, January 2017, Available at: https://doi.org/10.6028/NIST.IR.8062. Last accessed on 17 April 2019.

[10] The code of Fair Information Practices. Technical report, U.S. Department of Health, Education and Welfare, Secretary’s Advisory Committee on Automated Personal Data Systems, Records, Computers, and the Rights of Citizens viii, 1973.

[11] Andreas Pfitzmann and Marit Hansen. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management (version 0.33 april 2010). Technical report, TU Dresden and ULD Kiel, 2010, p. 13.

[12] The Article 29 Working Party, Opinion 4/2007 on the concept of personal data, 01248/07/EN WP 136, p. 8.

[13] Mina Deng, Kim Wuyts, Riccardo Scandariato, Bart Preneel, Wouter Joosen, A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements, March 2011, Volume 16, Issue 1, pp 3–32, p. 8.

[14]B-J. Koops (2014) The trouble with European data protection law. International Data Privacy Law, Vol. 4, Iss. 4, 3; N. Fisk (2016)

[15] Article 29 Working Party, ‘Opinion 03/2013 on purpose limitation’ (WP 203).

[16] Eva Lievens and Valerie Verdoodt, “Looking for Needles in a Haystack: Key Issues Affecting Children’s Rights in the General Data Protection Regulation,” Computer Law & Security Review 34, no. 2 (April 1, 2018): 269–78, https://doi.org/10.1016/j.clsr.2017.09.007.

[17] Article 29 Data Protection Working Group (2009) Opinion 2/2009 on the protection of children's personal data (General Guidelines and the special case of schools), 5. Available at: https://www.garanteprivacy.it/documents/10160/10704/1619292.pdf/1ab4d295-c2b9-405f-a2df-1e0f7ec9cfd1?version=1.0 Last accessed on 22 April 2019.

[18] Sion, Dewitte et al, An Architectural View for Data Protection by Design, PRiSE project.