Technical University of Munich
TUM School of Computation, Information and Technology

A data protection reference model for GDPR necessities in transparent utilization of privacy-enhancing technologies

Dmitry Prokhorenkov

Complete reprint of the dissertation approved by the TUM School of Computation, Information and Technology of the Technical University of Munich for the award of the Doktor der Naturwissenschaften (Dr. rer. nat.).

Chair: Prof. Dr. Jens Großklags

Examiners:

1. Prof. Dr. Uwe Baumgarten
2. Prof. Dr. Claudia Eckert

The dissertation was submitted to the Technical University of Munich on 18.12.2024 and accepted by the TUM School of Computation, Information and Technology on 20.05.2025.

To my Family and in memory of my father Prokhorenkov Victor

Abstract

Imposed privacy threats, demands, and changes in assessing the anonymization level of personal data underscore the need for many organizations to fulfil various compulsory regulations, including GDPR. Given imposed threats and considering the value of personal data and data subject rights, organizations have been continuously facing complex challenges in addressing privacy, data protection, and compliance needs since the introduction of GDPR in 2016 year. Nevertheless, introducing for this matter privacy-enhancing technologies such as Differential Privacy partly facilitates the referred implications. Yet, it leaves the organizations to demonstrate fulfilment, interpretation and translation of such goal-based requirements between technical and legal domains for diverse stakeholders without specific means or the need to utilize outdated frameworks and approaches.

We present a comprehensive analysis of the studied interdisciplinary domain, tackling implications from general approaches to mitigate referred weaknesses, considering diverse stakeholders' needs with the utilization of mixed methods for developing a dedicated reference model for these needs. The main contribution of current dissertations is a comprehensive reference model built on solid theoretical foundations, incorporating domain-specific requirements derived from the utilization of differential privacy, combined with a solid synthesis of knowledge in the privacy and data protection domain. The explicit process of the development of the reference model, besides the multiphased strategy to validate the quality of the developed artefact considering diverse data protection recommendations, was crafted and tuned with multiple phases of improvement with input from diverse experts. The interdisciplinary perspective tackled in the current study provides and enables an alternative way for fulfilling GDPR requirements while utilizing privacy-enhancing technologies and also delivers rich insights into how organizations can benefit from the application of the developed model, particularly applied by diverse stakeholders with varied levels of experience.

Our results demonstrate and contribute not only to the main goal of the current dissertation with a proposed reference model but also highlight that it is possible to interpret goal-based regulation in a practical way for enhancing the privacy of data subjects and also demonstrate the limitations of methods chosen. Furthermore, the results suggest the need to extend efforts to investigate interdisciplinary gaps between the technical and legal domains.

Zusammenfassung

Auferlegte Datenschutzbedrohungen, Anforderungen und Veränderungen bei der Bewertung des Anonymisierungsgrades personenbezogener Daten unterstreichen die Notwendigkeit für viele Organisationen, verschiedene verpflichtende Vorschriften, einschließlich der DSGVO, zu erfüllen. Angesichts der auferlegten Bedrohungen und in Anbetracht des Wertes personenbezogener Daten sowie der Rechte von betroffenen Personen stehen Organisationen seit der Einführung der DSGVO im Jahr 2016 kontinuierlich vor komplexen Herausforderungen, den Anforderungen an Datenschutz, Datensicherheit und Compliance gerecht zu werden. Dennoch erleichtert die Einführung datenschutzfördernder Technologien wie Differential Privacy in gewissem Maße die genannten Herausforderungen. Allerdings bleibt es den Organisationen überlassen, die Erfüllung, Interpretation und Umsetzung solcher zielbasierten Anforderungen zwischen technischen und rechtlichen Bereichen für verschiedene Stakeholder nachzuweisen – oft ohne spezifische Mittel oder unter der Notwendigkeit, auf veraltete Frameworks und Ansätze zurückzugreifen.

Wir präsentieren eine umfassende Analyse des untersuchten interdisziplinären Bereichs, die Implikationen von allgemeinen Ansätzen zur Minderung der genannten Schwächen angeht und dabei die Bedürfnisse verschiedener Stakeholder berücksichtigt. Dabei werden Mixed Methods genutzt, um ein speziell für diese Anforderungen entwickeltes Referenzmodell zu erstellen. Der Hauptbeitrag der vorliegenden Dissertation ist ein umfassendes Referenzmodell, das auf soliden theoretischen Grundlagen basiert und domänenspezifische Anforderungen einbezieht, die sich aus der Nutzung von Differential Privacy ableiten. Dieses Modell wird mit einer fundierten Synthese von Wissen im Bereich Datenschutz und Datensicherheit kombiniert. Der explizite Entwicklungsprozess des Referenzmodells, einschließlich der mehrphasigen Strategie zur Validierung der Qualität des entwickelten Artefakts unter Berücksichtigung verschiedener Datenschutzempfehlungen, wurde in mehreren Phasen mit Feedback von Experten kontinuierlich verbessert und angepasst. Die interdisziplinäre Perspektive, die in der vorliegenden Studie behandelt wird, bietet und ermöglicht eine alternative Herangehensweise an die Erfüllung der DSGVO-Anforderungen unter Einsatz datenschutzfördernder Technologien. Sie liefert darüber hinaus wertvolle Erkenntnisse darüber, wie Organisationen von der Anwendung des entwickelten Modells profitieren können, insbesondere wenn es von verschiedenen Stakeholdern mit unterschiedlichen Erfahrungsstufen angewendet wird.

Unsere Ergebnisse zeigen und tragen nicht nur zum Hauptziel der aktuellen Dissertation mit dem vorgeschlagenen Referenzmodell bei, sondern unterstreichen auch, dass es möglich ist, zielbasierte Regulierung auf praktische Weise zu interpretieren, um den Datenschutz von betroffenen Personen zu verbessern. Gleichzeitig zeigen sie die Einschränkungen der gewählten Methoden auf. Darüber hinaus legen die Ergebnisse nahe, dass verstärkte Anstrengungen erforderlich sind, um interdisziplinäre Lücken zwischen technischen und rechtlichen Bereichen zu untersuchen.

Acknowledgment

First, I would like to thank my supervisor, Professor Uwe Baumgarten, for providing me with the opportunity to join his department and finally starting my doctoral track during such not an easy time in Covid. Furthermore, I want to thank him for providing flexibility and independence in conducting the research, continuously supporting me in this endeavour and spin-offs, and sharing his knowledge.

I also gratefully thank Professor Claudia Eckert for the opportunity to join and contribute to such fascinating research projects and for giving me more time to revise my dissertation, without which the dissertation would not have been, I suppose, finished.

I also want to thank Professor Hervé Glotin, who highlighted and demonstrated what actual research is supposed to be and made further recommendations.

Further, I want to express huge gratitude to the colleges and all other project members. It was a joy to work with you, exploring unsolved and complex issues and finding solutions together. Additionally, I'm deeply thankful to my friends for their unwavering support. Their encouragement and shared moments have been a source of strength and joy, making this journey more bearable and memorable.

Last but not least, I want to express my gratitude and thank my family, Corinna, Elis, and Nikola, for the brightest moments during this journey.

Contents

List of Figures

List of Tables

Acronyms

BES Best Evidence Synthesis

BIDT Bavarian Research Institute for Digital Transformation

CICA Canadian Institute of Chartered Accountants

CIT Computation Information and Technology

CNIL Commission nationale de l'informatique et des libertés

DPA Data Protection Authority

DPIA Data Protection Impact Assessment

DPbD Data Protection by Design

DPG Data Protection Goals

DPO Data Protection Officer

DPMS Data Protection Management Systems

DSR Design Science Research

ENISA European Union Agency for Cybersecurity

EPDP European Data Protection Board

GDPR General Data Protection Regulation

HA Hybrid Approach

IP Internet Protocol

IS Information Systems

ISMS Information Security Management System

ISO International Organization for Standardization

IT Information Technology

ITIL Information Technology Infrastructure Library

MM Maturity Model

PDCA Plan-Do-Check-Act

PDLC Privacy Development Life Cycle

PET Privacy Enhancing Technologies

PIA Privacy Impact Assessment

PII Personally Identifiable Information

PIMS Personal Information Management Systems

PMM Privacy Maturity Model

PMRN Privacy Management Reference Model and Methodology

PbD Privacy by Design

PSMO Privacy Security Managers Officer

RM Reference Model

RMDP Reference Model for Differential Privacy

SDLC Software Development Life Cycle

SDM Standard Data Protection Model

SIPOC Supplier-Input-Process-Output-Consumers

TOM Technical and Organizational Measures

TUM Technical University of Munich

WP Working Party

1 Introduction

The adaptation and implementation of General Data Protection Regulation (GDPR) norms in the European Union (EU) are actively accompanied by the advancement and dissemination of various methods and procedures aimed at safeguarding personal data, while at the same time, the number of various threats and variety of privacy attacks aimed at violating the integrity of personal data is growing. However, the complexity of interpretation, application and meeting compliance norms with various national legislative amendments, as well as the primary legislation as GDPR on protecting personal data in the EU, is demanding for many organizations for numerous reasons. Moreover, the demand and need to adapt and enhance compliance examination approaches for such methods as Privacy Enhancing Technologies (PET) aimed at protecting personal data are rising and mandated daily.

1.1 Research Context

Currently, various methods and processes related to meet the GDPR compliance and monitoring the level of personal data protection are some of the most active areas of research in the academia and various industries, the results of which are in great demand. The active development of additional methods aimed at protecting personal data, namely PET Differential Privacy (DP) [51, 50], and their subsequent active implementation and disposal require more careful investigation in the context of GDPR compliance. Undoubtedly, due to the widespread use of various PETs (as an alternative or supplementary method for protecting personal data) and the expectation that referred PETs will enable better compliance with current norms, various problems related to managing and meeting the GDPR needs are the highest priority for many organizations, data processors and stakeholders. The wide adaptation of PET methods could be traced in various domains and industries and is due to the following advantages: a) multifunctionality and flexibility in utilization, b) countering various compliance incidents such as Cambridge Analytica [18] and others [138] , and c) delivering an improved data protection level for sensitive data. It is also worth noting that intensive PET research in academia is primarily focused on topics of data utility and increasing the data subjects' privacy level; however, multiple issues in meeting GDPR needs and bridging the legal and technical domains remain open [135]. Moreover, as indicated by various studies [182, 147], the need for additional interdisciplinary research is in demand due to the need to interpret issues of personal data protection from different perspectives, considering a) both the technical and legal GDPR requirements b) the features of PET applied and c) the needs of various stakeholders.

As a result, data processors, third parties and stakeholders experience difficulties in accurately interpreting the GDPR requirements and ensuring the required level of personal data privacy [107, 145, 99, 117, 10]. Simultaneously, since the GDPR announcement, the privacy area in Information System (IS) has been more focused on describing the current state of data protection [17], although questions regarding compliance have been studied for a long time [1, 36, 191], nevertheless, with the introduction of PET, as well as the Privacy by Design (PbD) approach, made it possible to enhance not only the taxonomy but also qualitatively validate the methods of previously introduced ideas [108], but still many questions remains open [182].

1.2 GDPR

The GDPR is a set of regulations that have been in force since May 2018. It is considered one of the most serious and consequential contributions to the recent progress and transformations aimed at protecting

personal data. From 1995 until 2018 inclusive, before the entry into force of the GDPR in the EU, the so-called EU directives on protecting personal data were in force and widely used [48]. Simultaneously, in the context of the GDPR evolution and development and subsequent implementation, these stages were, to some extent, hasty but highly in demand. Simultaneously, in each iteration of norms updating and laws related to data protection, the delicate balance between privacy and limitations on processing personal data has been mostly maintained. However, it is worth noting that since the mid-20th century, the development approach and implementation of legal acts, as well as specific taxonomy and vocabulary related to safeguarding sensitive information, has not changed in many respects and remains the same, as a result of which similar consequences and complexity can be traced in the interpretation of requirements, terms and separate GDPR pieces partly addressed by author [139]. As a result, the public and experts have witnessed multiple scandals related to the leakage of personal data from social networks, which currently have extensive access to sensitive data and are also vulnerable to the data subjects' identification [64]. Even so, the implementation and meeting compliance needs delivered in the GDPR make it likely to achieve a certain balance between transparency and fairness in personal data processing [120]; however, still many IS and various services do not fully comply with GDPR, even while utilizing additional methods or approaches aimed at improving these shortcomings [41].

In other words, the requirements of regulatory authorities are ahead of the capabilities of data processors and third parties in the context of meeting compliance requirements and minding precise implementation, even considering the multiple additions and recommendations presented by data controllers and various privacy agencies. Put differently, the lack of methods or quality metrics for determining the degree of compliance with GDPR norms imposes limitations on the selection and use of a proper additional method of protecting personal data, such as PET DP. Also, it increases the privacy risks for data subjects and data processors, including reducing the efficiency of PET utilization.

1.3 Data protection and technology advancement

Historically, over the last decade, government authorities have been modifying and implementing various legislation norms to enhance the safety of personal data as one of the preventive approaches to facilitate privacy risks for data subjects due to the increased number of incidents with personal data. However, as practice shows, the implementation of legislative acts within the framework of present requirements in the data protection domain occurs with a slight lag [139] from various stakeholders' current needs or expectations. Thereby reducing performance and imposing certain restrictions on a) the adaptation of supplementary methods such as PET DP, b) the estimation of privacy risks for data subjects and c) the adaptation/development of various privacy frameworks and additions to GDPR norms. Among the many reasons for this imbalance between data processors, third parties and data controllers, there may be incomplete or differently interpreted GDPR terminology and requirements for assessing a) different levels of anonymization, b) privacy risks, c) compliance needs and d) Privacy Impact Asessment (PIA).

Despite the above difficulties in interpreting GDPR norms, existing gaps can be supplemented with the help of various auxiliary a) the International Organization for Standardization (ISO) standards 27k family, b) proposals of various privacy agencies Commission nationale de l'informatique et des libertés (CNIL) [76], and others [132] and c) privacy frameworks Standard Data Protection Model (SDM) [163], LINDDUN [207] and STRIDE [178]. Moreover, the further adopted SDM framework has reasonable advantages, such as comprehensive data protection (security) goals (DPG) [206], which could be addressed by various stakeholders [78]. We also provide an extended discussion of the SDM framework and DPG in Chapter 3. Nevertheless, despite the active development and adaptation of referred techniques and recommendations, effectiveness in evaluating GDPR compliance norms raises many questions among experts since it is complicated by insufficiently explicit recommendations of GDPR norms, as mentioned earlier. Therefore, data processors and third parties are forced to utilize other methods, such as PET DP and agree to a level of anonymization of personal data higher than required. In addition, evaluate compliance needs based on their own interpretations of GDPR norms, which could be biased based on the tasks performed, thereby reducing a) data utility and b) the safety level of personal data. On the one hand, the effectiveness

and transparency of a) monitoring, b) meeting compliance needs, and c) implementation of GDPR norms in the context of used PETs are partially defined and can be interpreted using different GDPR articles (Article 5 (2), Recital 29). On the other hand, it is also complemented by examination and monitoring by the designated Data Protection Officer (DPO) in the organization. Simultaneously, fulfilling Recital 26 GDPR, data processors and third parties can avoid many restrictions related to the processing of personal data "*The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.*" Put differently, in the case of complete anonymization, data processors and third parties are above the GDPR needs. However, estimating or determining the anonymization level is a complex and unregulated process, the description of which is not specified in the GDPR norms and is complicated by various interpretations, as mentioned earlier. Therefore, numerous data processors or third parties are attempting to achieve full anonymization or higher to avoid privacy risks for data subjects, even if it leads to loss of data utility. Therefore, the need for further research and examination of the GDPR norms in the context of PET DP utilization is highly demanded. Moreover, if the required level of compliance with the GDPR is not achieved, this can cause considerable financial difficulties for organizations; thereby, organizations need to utilize technological and human resources to achieve their goals and comply with the GDPR requirements [111]. Strictly speaking, taking into account the indicated difficulties, namely a) the complexity of interpreting various GDPR acts and articles, which is also twisted by many intersections with other GDPR articles b) the need for further legal interpretations, c) the absence controlled methods and approaches for the disposal of PET methods to protect the sensitive data, e) the complexity of interpreting the requirements, including the difference in the "privacy" taxonomy for different ecosystems (technical and legal). All these factors significantly impact data processors within the framework of PET utilization and fitting the compliance norm, thereby increasing the probability of privacy risks for data subjects.

In addition, GDPR compliance assessment and examination of privacy risks in the context of PET DP [149] utilization can be supplemented and carried out using a) indirect indicators and metrics (such as Data minimization, Availability, Integrity, Confidentiality, Unlinkability, Transparency, Internveability), which may depend on a) use cases, b) type of data, c) adaptation of various PbD or Data Protection by Design (DPbD) approaches and d) other DP parameters. To put it another way, an improved approach for assessing PET compliance in the context of GDPR requirements will provide a more transparent process for assessing compliance risks, b) effective disposal of PET DP, c) tight integration of the two ecosystems of technical and legal, and d) effective monitoring of compliance standards by data controllers. Although, over the past decade, there has been progress (improving approaches to meet compliance standards) noted by many experts and scientists, there are still many questions about how and in what capacity it is necessary to comply with compliance standards and how to monitor the effectiveness of the methods used to protect personal data. However, this issue has remained open for a long time [134] and requires additional attention [7] (*there is a need for a developed artifact, which not yet suggested or produced*).

In other words, although the existing GDPR norms and amendments to them have been available to the public since 2018, the demand for supplementary analysis is caused by the lack of a suitable approach to assessing GDPR compliance during PET DP utilization. Alternatively stated, the lack of a unified approach to assessing GDPR standards within the framework of PET may, in some cases, negatively affect the degree of protection of personal information, as well as the selection and tuning of an additional method such as PET DP. In addition, data processors are forced to comply with GDPR norms but do not have the necessary tools to assess compliance and are forced to sacrifice data utilities to ensure a higher degree of reliable protection of personal data. This addresses various questions related to assessing the level of anonymization since data controllers and available interpretations of the GDPR do not specify specific requirements or methods to ensure complete anonymization. At the same time, these restrictions are imposed on the PETs used, which also do not have tools and methods for assessing the degree of compliance with either the level of anonymization and GDPR standards [195]. Although, despite the constant updating and addition of legislative laws and regulations, the questions of how and by what means to provide these services to the proper extent remain open and the implementation of these issues

remains at the discretion of the data processors. To identify conflicts between various requirements for different ecosystems (technical and legal), a suitable approach or a single structure for modelling or a consistent approach is required to consider the requirements for previously specified stakeholders, which can be simultaneously merged and, if necessary, expanded or supplemented.

Therefore, the relevance of this work is in high demand, the outcomes of which will deliver data processors and third parties with the possibility to evaluate and comply with the GDPR provisions in the context of the specified PET. Accordingly, within the framework of this dissertation, a reference model (RM) (which can serve as a GDPR blueprint for various stakeholders) will be delivered and analyzed for considering data protection goals (DPG) while utilizing PET (using DP as an example, since DP is one of the approaches that allows one to mathematically substantiate and validate the degree of protection of personal data), in the context of existing GDPR norms.

The results of this work contribute to both the theoretical and practical application and analysis of compliance measures for various actions performed with personal data. In addition, the proposed reference model can be applied or adapted for various needs of a) PET, b) privacy frameworks, c) Data Protection Impact Assesment (DPIA), and d) stakeholders, and will also complement the current GDPR norms in assessing compliance level considering various anonymization levels.

1.4 Motivation

We argue that existing GDPR compliance assessments (methods) for PET (DP) are insufficient due to the lack of clear guidance or procedures in existing GDPR notations. Consequently, these problems impose restrictions on a) the degree of compliance and b) the limited use of PET DP. To respond to this gap, we will consider and propose an alternative approach (reference model) to enhance the examination process and meet compliance needs by utilizing the data protection goals established below. The suggested reference model will not only improve compliance standards for various organizations and stakeholders in the context of using PET DP, but can also complement the existing DPIA approach and can also be adapted for DPO duties. Also, the proposed solution has the ability to adapt and interpret results for other PET methods. Simultaneously, the results (developed reference model) of the current study and dissertation will make it feasible to incorporate the technical and legal ecosystems in the context of an alternative approach to assessing GDPR requirements for DP, thereby enhancing a) the data utility b) the transparency of transactions with personal data for all participants (stakeholders), c) a sufficient response to new kinds of privacy attacks, d) safeguard of personal data, e) sufficient compliance with anonymization means, and f) practical implementation of add-ons or ISO standards. Our research and the accompanying suggested reference model are designed to complete a critical gap in the implementation and adaptation of theoretical research (results) in the field of data protection, compliance and the PET DP method.

1.5 Challenges and Research questions

This section discusses open high-level research problems and challenges that are the focus of this dissertation.

As mentioned earlier, organizations and data processors are constantly challenged in the context of compliance and the use of PET DP. Simultaneously, failure to comply with compliance needs leads to a) a decrease in the degree of protection of personal data and b) significant or critical financial difficulties or fines for organizations. Although the GDPR prescribes and is designed in such a way that organizations fulfil and follow the necessary requirements to meet data protection needs, the lack of a uniform regulation or standard leads to the fact that organizations have to approach compliance standards in a non-trivial way and taking into account the specifics of the tasks performed in the organization.

Simultaneously, compliance issues have been addressed in IS studies for quite a long time. However, even after the GDPR was concluded in 2018, compliance issues remained for quite a long time without

much attention from researchers [111]. Although the situation has improved slightly since the introduction of GDPR, researchers have expanded their research focus and begun to study privacy and compliance issues for various stakeholders, including comparing and considering multiple technological approaches and methods. On the other hand, attempts to address compliance issues in IS were present earlier [52, 36, 1]; however, these studies and their contributions present largely preventive or detective solutions rather than corrective solutions. This is explained by the fact that the authors excluded legal examination, which is taken into account in corrective decisions. Thus, this indicates the absence of a suitable structure or approach that would take into account the compliance requirements on a full scale applied to the entire organization or in the context of the PET used, including allowing analysis and tracking of ways of their implementation. In other words, the lack of significant recommendations on how to meet the requirements of regulators and combine them with the needs of organizations cannot be achieved using only considering available regulator requirements and proposals of the [191]. At the same time, since the GDPR was announced and put into effect, a lot of studies has addressed compliance issues and has made significant contributions to improving compliance. However, due to the complexity of the interpretation of legal requirements, technical approaches suffer from a) compliance assessment methods [135], b) implementation and understanding of compliance requirements [111]. In addition, as the authors state [108], existing works in IS only discuss these gaps in the context of possible goals and approaches to improve compliance rather than offering artefacts to solve these problems. Therefore, our goal in this study is:

- **Research Goal:** To support the implementation of data protection goals in PET DP utilization from GDPR compliance perspective by developing a GDPR blueprint reference model.

The current research goal involves multiple and different steps as well as research contributions, including the research being interdisciplinary. The outcome of the current study could be applied in various domains and industries, on the one hand, enhancing better privacy protection for data subjects and, on the other hand, enabling more efficient data utility and more suitable transparency in meeting GDPR notations. Taking into account the previously described research gaps and according to our scientific research, we have identified several challenges in the field of assessing compliance with GDPR while utilizing the PET DP, which will be further discussed and analyzed in this work.

- **RQ1** How can the different levels of GDPR anonymization for PET (DP) be interpreted?
- **RQ2** Which stakeholders and related objectives influence compliance, monitoring and compliance with GDPR?

Initially, we require a ground on which to rely, put differently, to define, propose and validate metrics that can be utilized in the context of assessing the degree of compliance or anonymization levels and later analyze if it can be adapted for PET (DP); To achieve this goal, an analysis of both scientific studies (primary and grey) was carried out, and a survey of various stakeholders involved in data protection, fulfilling compliance needs was conducted. These insights will allow us to determine the core and structure of the problem topic discussed in this dissertation and structure the development of the proposed solution (reference model). The results of previous studies represent and form the necessary basis for our further research. Simultaneously, we complement our results with a critical analysis of existing literature related to our study area and compliance domain. Saying this, we analyze existing industrial approaches and recommendations of various privacy agencies, namely the European Union Agency for Cybersecurity (ENISA) and CNIL, and we also consider and incorporate additional ISO standards 27k family to validate and improve the development process as well the final result of the developed reference model. Applying and implementing various recommendations and practices enable better integration of reference models and benefit various organizations and industries while further utilization.

- **RQ3** What methods exist in the scientific literature/industry to achieve and comply with GDPR compliance standards (for reference models)?

The third research question addresses and validates the primary and supplementary specific conditions that the developed reference model must address in the context of GDPR compliance, as well as the suitable PET DP. Simultaneously, we are conducting a critical analysis of the diverse methods available, considering the knowledge in the adapted security reference model (RMIAS) for IS [34]. We use various techniques and approaches to achieve the required results, including adapting and critically analyzing the results obtained during various a) interviews with experts, b) review of existing recommendations for the industry, including DPIA, c) various research papers, and d) scientific workshops and seminars.

- **RQ4** How can the GDPR compliance reference model (for DPG adaptation) be defined and presented?

We are developing and offering a reference model, which is one of the main contributions to this work. Based on the results of previous research, including those supported by alternative sources and critical analysis of existing research studies and additions to various industry standards, as well as separate subsequent research (investigations) of GDPR compliance efforts. We argue that the previously mentioned approaches and the proposed reference model developed in this dissertation can satisfy the reference standards of compliance and vice versa.

- **RQ5**. How do experts and practitioners evaluate aspects of the proposed compliance reference model for PET from various perspectives (feasibility, necessity, importance, economic and technical components)?

One of the main goals of oriented research design is creating and developing a relevant artefact. In addition, one of the criteria for a relevant artefact is the result of validating the artefact and applying it to research in the area under study, followed by a critical analysis of the results. We achieved and validated the proposed solution through qualitative interviews with key stakeholders, which we supplemented during the research and interviews with suggestions for improving our solution. During the iterative process of developing and finalizing the reference model, additional qualitative interviews were conducted among key stakeholders to assess the future prospects of the presented artefact.

1.6 Contribution

The results of this work are threefold:

- a) analysis and recommendations were carried out to determine the levels of anonymization, which made it possible to interpret the levels of anonymization for PET DP (formulate requirements for assessing the level of anonymization of PET (DP));
- b) an analysis of additional metrics was carried out based on various privacy frameworks, PET, including the addition and systematization of the needs of various stakeholders involved in assessing the degree of compliance considering the DPbD and GDPR requirements; a privacy taxonomy was proposed, and an approach for benchmarking the level of anonymization based on the proposed taxonomy was proposed;
- c) a reference model is presented that ensures optimization of the critical gap in the implementation and adaptation of theoretical research (results) in the field of compliance and the use of PET DP in practice.

1.7 Methodology

In order to achieve the needed results, this work and research strategy has been carried out following a solid theoretical foundation, which allows the end user to achieve their goals. A research strategy is primarily a general plan and structured approach to conducting research, which contains both recommendations

for planning, conducting and monitoring, which can be supplemented using various methods presented below.

We utilize an adapted design science research method [84] and a design-oriented IS approach [141]. The configured and adapted approach is presented in **Figure 1.1**; this Figure introduces a high level overview of the elements included and also depicts the complexity of a research domain (In the following chapters, elements included in this Figure will be explained or discussed in detail). This approach has a flexible and constructive guideline and paradigm that allows to expand the boundaries of human and organizational capabilities by creating alternative or new artefacts (both jointly and separately), namely methods, models, theorems, algorithms, as well as system design methods [40, 82, 84, 141]. The chosen approach guarantees the accomplishment of both practical and theoretical results while simultaneously allowing for the accomplishment of results in such issues as novelty, usefulness, and advancement (including explained, modelled, and accessible for interpretation). In addition, design-oriented IS research treats four main phases [141]: 1) analysis, 2) design, 3) evaluation, and 4) diffusion. Simultaneously, in order to qualify as research (as opposed to solution development) as specified in the approach, four principles must be met:

- **Abstraction**: Each artifact must be applicable to a class of problems;
- **Originality**: Each artifact must substantially contribute to the advancement of the body of knowledge;
- **Justification**: Each artifact must be justified in a comprehensible manner and must allow for its validation;
- **Benefit**: Each artifact must yield benefit – either immediately or in the future – for the respective stakeholder groups;

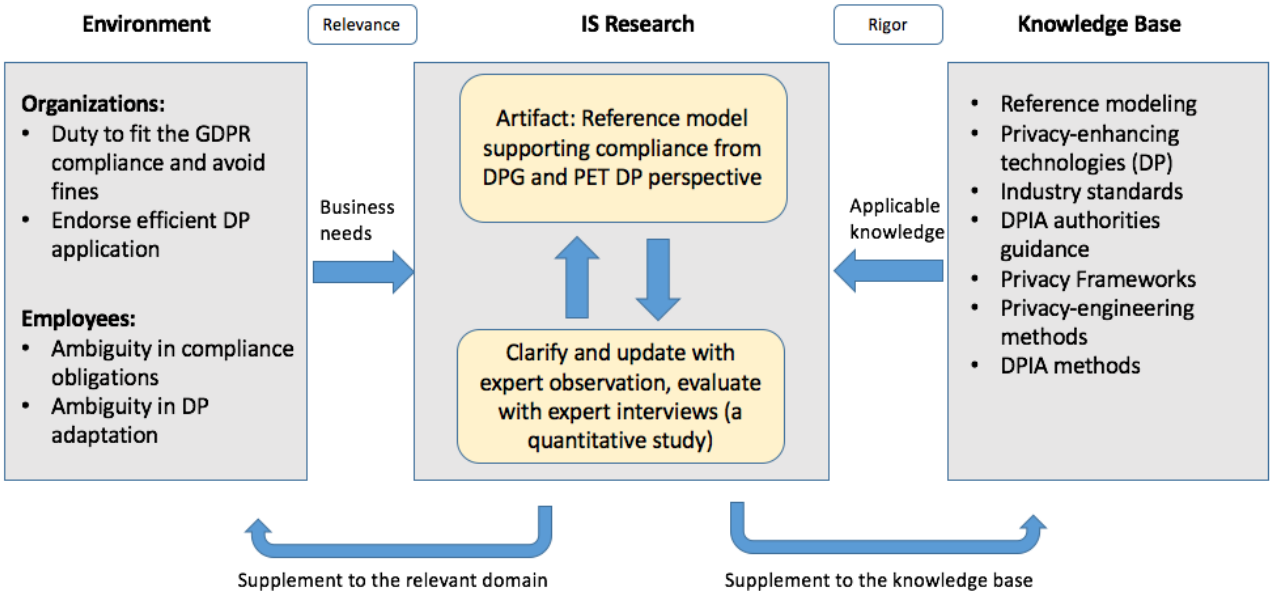


Figure 1.1 Adapted research framework structure for this dissertation [84]

The conceptual framework [84] developed by Hevner et al delivers seven recommendations and guidelines for performing and complying with the principles of good design science research.

- **Recommendation 1** Design science research must produce viable artifacts such as a model or a method (Design as an artifact).

  The chosen approach itself contains various methodologies, approaches, models and artifacts. The result of our search for a solution (reference model) is also a significant artifact for the IS field. As U Frank notes [60], one of the variants of an artifact can be both a conceptual model and a reference model.

- **Recommendation 2** A relevant problem has to be solved (Problem relevance).

  Earlier in this chapter, the importance and necessity of applied PET and GDPR compliance is presented.

- **Recommendation 3** The utility, quality, and efficacy of the designed artifacts have to be demonstrated via well-executed evaluation methods (Design evaluation).

  We use various methods and approaches to validate our work. The analysis presented below contains the analysis and examination of these methods.

- **Recommendation 4** Clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies have to be provided (Research contribution).

  The full contributions of this dissertation are presented and displayed in **Figure 1.2** and summarized in **Section 1.5**;

- **Recommendation 5** Design science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact (Research rigor).

  The presented reference model in this work is in some way complementary to a variety of existing studies, including validation through a structured research approach and methods, critical analysis and inclusion of thematically and extensively research papers, as well as discussions with researchers in the academia.

- **Recommendation 6** The search for an effective artifact requires utilizing available means to reach desired ends (Design as a search process).

  Finding and identifying a valid solution to the previously mentioned problem is the design of the solution search process. In our study, we first identified research problems and gaps, and for the proposed solution (artifact), we presented methodologies and approaches (we present a detailed description of our research path in **Chapter 4**). We demonstrated the capabilities and convenience of our concept using various techniques, confirming the validity of our approach. In conclusion, we argue that based on the limitations, future research directions are essential and vital in this topic.

- **Recommendation 7** The design science research must be presented effectively (Communication of research).

  In conclusion, the dissertation presented is a structured and organized document. Partial results of this work are presented in the context of published articles at international conferences/workshops, including an adaptation of various critical remarks received during presentations at diverse events in the academia.

Introduced and explained principles were utilized in the context of the current dissertation. In this work, we deliver a (reference model) to enhance compliance (safety and integrity) of personal data while utilizing PET DP.

Research Method: During the research and writing of dissertations/articles, various scientific methods can be used. The method chosen to validate a given work must be appropriate for the type of study being conducted. Applied researchers and/or engineers whose work results in the creation or adaptation of noticeable or tangible artifacts primarily choose design science research methodology. The research project

followed agile approaches and iterative processes, including collecting feedback on each identified gap, as well as a presented (reference model) depending on the results of the research and where the research led. This work uses the design science research method and the Best Evidence Synthesis (BES) method [181]. This method for developing RM was chosen because it enables and provides guidance during data synthesis and also allows one to maintain focus when selecting knowledge for inclusion in the synthesis. In addition, this method provides instructions for identifying and including studies in the analysis, where the primary contribution when using the BES method is the inclusion of concepts in the developed RM. This process was accompanied by multiple iterations (data collection, analysis, model development, testing and validation of the model with various experts, as well as presentation of results and improvement of the model and elements of the model in academia; the results of this research were integrated into a joint project between Bavarian Research Institute for Digital Transformation (BIDT) and Technical University of Munich (TUM) (Project name: Differential privacy for social networks[1]).

1.8 Reference modeling

Any field of knowledge may have both an explicit and an inferred conceptual model that allows one to describe or represent the phenomena being studied or hypothesized. Neglecting a detailed description of the technical implementation and using an abstract approach with a focus on the basic concepts of the subject area [60], these models are one of the most essential elements of computer science and business informatics [20], used in many applied and theoretical problems.

- **Definition: Model**

  A model is the abstraction of observations regarding the contents of a subject that serve a specific purpose. The design of models takes place in the construction phase. [201]

As the authors [94] state conceptual models "may and should map reality, guide research and systematize knowledge," as well as a reference model for descriptive knowledge may include methodological knowledge, and RMs are usually depicted graphically. The effectiveness of the graphical presentation of RM is critical so that it can be easily analyzed and understood by any stakeholder from diverse backgrounds. RMs generally do not use any formal modelling language, but ease of understanding and memorization is achieved through visual reflection. A conceptual model that has the ability to represent a problem in an industry and includes "domain knowledge", can be interpreted as a reference model [130].

- **Definition: Reference Model (RM)**

  A reference model is an abstract framework for understanding significant relationships among the entities of some environment. It enables the development of specific reference or concrete architectures using consistent standards or specifications supporting that environment. A reference model consists of a minimal set of unifying concepts, axioms and relationships within a particular problem domain, and is independent of specific standards, technologies, implementations, or other concrete details. [119]

Simultaneously, reference models are considered as a reusable (usable) artifact [54], which has the necessary and required characteristics (recommendation and generality) [190, 201], as a result of which, during the practical application of the specified reference model, significant efficiency is achieved by reducing the complexity of the issue or problem under study [196]. RM can be used in a variety of scenarios [56], including the description of organizational and business processes re- ingeniering or knowledge management. However, the most developed and widely used direction can be characterized in the context of IS development [130]. One of the significant advantages of RM in these scenarios is the high ability to reuse specified model, as well as high compatibility with IS [61], along with this, analysis of the research area and development of specialized models are the main goals in the development of RM [173].

[1] https://www.bidt.digital/forschungsprojekt/differential-privacy-ein-neuer-ansatz-zum-umgang-mit-social-big-data/

1.9 Thesis Outline

The rest of this dissertation is structured as follows:

In **Chapter 2**, we describe and present the workflow of the thesis and also the issues of compliance (privacy) in the context of levels of anonymization, PET and current gap issues. Alternatively stated, this chapter contains the necessary information to understand this dissertation; First of all, we provide an understanding of the basic concepts of compliance, different levels of anonymization, their main characteristics that lead to the challenges we address or study, including comparison with other approaches; Secondly, this chapter provides a broad understanding of PET (DP) in general;

Chapter 3 (related works) reviews work relevant to the general objectives of this dissertation, including State of the Arts current a) compliance approaches, b) adapting DP to compliance standards, and c) industrial approaches and methods. We also discuss the advantages and disadvantages of the chosen approaches; In particular, we analyze the requirements and needs of various stakeholders who are directly involved in the process of protecting personal data, including analyzing the possibility of supplementing existing GDPR regulations based on the data received.

Chapter 4. From a review of current work and scientific studies, we will extract and discuss the necessary requirements for the presented solution (reference model), which are critical for the utilized PET with respect to the GDPR, considering the DPG and the requirements for different levels of anonymization. Modelling of the reference model will be carried out in the context of the general requirements for designing reference models, supplemented by feedback from various privacy experts. In addition, we will present and critically evaluate the relevance of a) the metamodel and b) visual representations of the reference model.

Chapter 5 is devoted to the development of this reference model, which describes the process of constructing the model as well as the first iteration and the feedback received from experts in academia and industry. This chapter discusses in detail all the elements of the reference model, as well as the corresponding design decisions during the implementation of the creation of this model.

Chapter 6 is devoted to the validation of the presented reference model among various stakeholders and experts utilizing varied methods.

Chapter 7, we discuss future work and the direction of the following open research questions, including the prospects for the proposed reference model. We also discuss possible alternative application approaches for the proposed solutions, including the verification of results. Finally, we conclude the dissertation with a critical assessment of the results achieved, as well as highlighting and presenting for future research directions.

1.10 How to read this thesis

The diagram below, **Figure 1.2**, visualizes the chapters, research questions, as well as workflow and study strategy for the presented research gaps. The four conceptual chapters reflect the main contributions of this dissertation and answer the research questions. This dissertation is presented in such a way that the reader can read each chapter separately, but the greatest contribution and understanding can only be achieved by reading all the chapters in sequence.

Each of the chapters, as previously mentioned, is validated according to the presented principles and techniques presented in *Section 1.5* and *1.6*. In the addition in **Chapter 6**, we validated the presented solution in the context of an interview with a DPO.

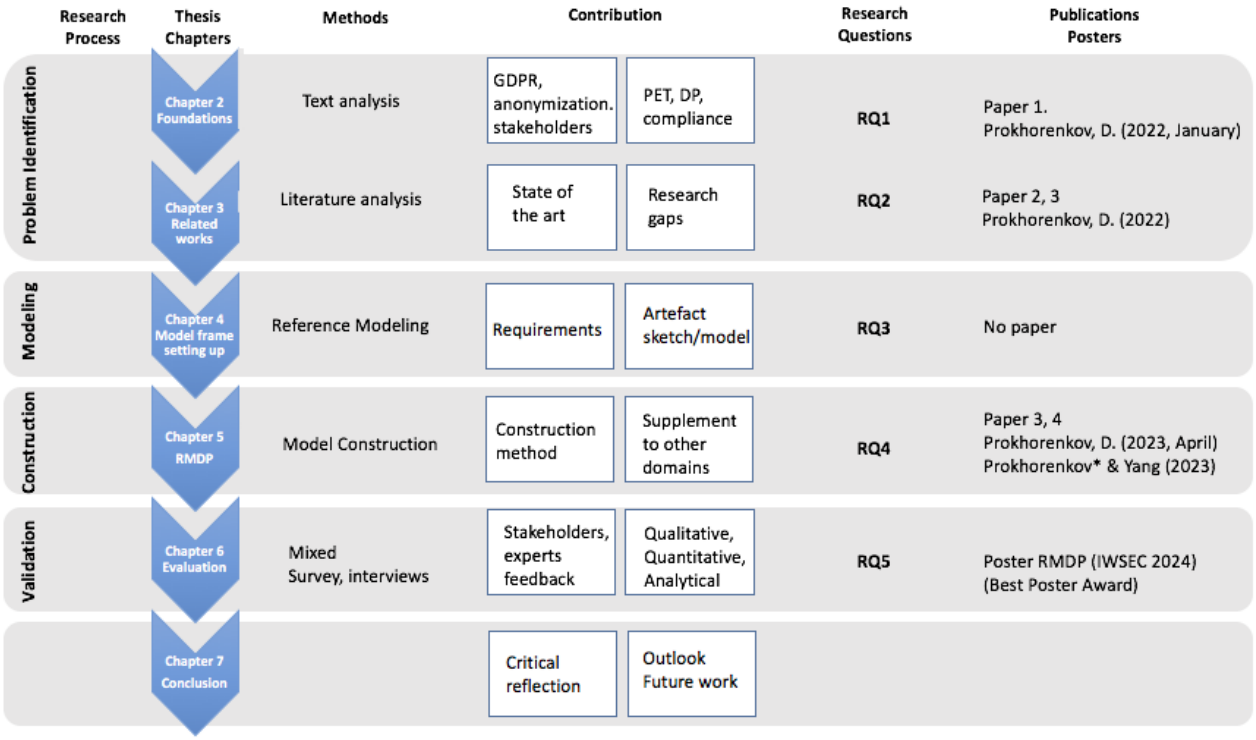


Figure 1.2 Research approach and overview of current disseration

1.11 Preliminary Publications and Talks

This work uses materials and results written by the author of this dissertation. The results of some chapters are currently being reviewed at research conferences and workshops. The list of already published or presented research contributions is presented below:

- Prokhorenkov, D. (2024, September). Best Poster Award[2]: Reference Model for Differential Privacy (RMDP). 19th International Workshop on Security (IWSEC).
- Prokhorenkov, D., Cao, Y. (2023, November). Towards Benchmarking Privacy Risk for Differential Privacy: A Survey. In Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (pp. 322-327).
- Prokhorenkov, D. (2023, April). Toward Compliance Implications and Security Objectives: A Qualitative Study. In 2023 IEEE 39th International Conference on Data Engineering Workshops (ICDEW) (pp. 138-145). IEEE.
- Prokhorenkov, D. (2022, January). Alternative methodology and framework for assessing differential privacy constraints and consequences from a gdpr perspective. In 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 0359-0364). IEEE.
- Prokhorenkov, D. (2022). Anonymization Level and Compliance for Differential Privacy: A Systematic Literature Review. 2022 International Wireless Communications and Mobile Computing (IWCMC), 1119-1124.

During the course of the research project and the writing of this dissertation, the results and process of the research were actively presented to various stakeholders, including among numerous thematic workshops and events in the academia and in industry (a more detailed description of the events is presented in **Chapter 6**).

[2]https://www.iwsec.org/2024/

1.12 Limitations of the study

Emphasizing the significance of the study, it provides a comprehensive interpretation of the PET DP compliance assessment process, including the data protection goals and an analysis of the GDPR requirements. However, it's crucial to acknowledge the encountered limitations during the thesis work and study.

Firstly, it is worth noting that the primary language of the current study is English, which can otherwise be called "lingua franca in the sciences" [141]. However, there is a lot of additional literature in other EU languages dedicated to the protection of personal information and issues compliance. In addition, some studies were read with the help of an uncertified translator or without an official translation, which may, therefore, indicate a loss of meaning in the context of an inaccurate translation.

Furthermore, certain materials, including those from data protection agencies or regulatory bodies, which could be crucial or recommended, were not included in the search index (or the final sample for questionnaires). This exclusion was due to technical or language barriers, underscoring the challenges faced during the study.

Thirdly, the selection of the research strategy and approach is influenced by the interdisciplinary nature of the topic, as well as time constraints.

Finally, as current research explores a new and interdisciplinary topic, we acknowledge that the findings may differ from existing practical solutions. We also demonstrate a more comprehensive explanation of the current study and the limitations of the results in *Section 7.3*.

2 Foundations

Putting privacy in Context, Compliance outlook

The current chapter is a descriptive part of the GDPR regulations and requirements. Initially, we present necessary terms from the GDPR to enhance understanding of issues related to the concept of personal data and terminology in the GDPR (**Section 2.2**) and related details regarding PET (**Section 2.3**), including accompanying knowledge regarding DPIA (**Section 2.4**) and data protection goals (**Section 2.5**). We also highlight and stress attention on relevant terms, marked as definitions or interpretations, to enhance the consistency and readability of the presented material due to various available interpretations of the same terms in academia and industry. Among other things, we present and describe the limitations and possibilities for interpreting levels of anonymization, including critically debating the likelihood of interpreting levels of anonymization in the context of PET, considering various stakeholders. We also consider the numerous academic studies in the personal data protection domain and PET utilization and the DPG interconnection with GDPR requirements and highlight interconnection with outcomes from our studies.

2.1 Introduction

Privacy issues have historically been studied from many perspectives, as mentioned earlier. Moreover, privacy and GDPR compliance issues are the most actively examined and in demand by many stakeholders and various industries.

2.2 The common Terms in this Thesis

In this section, we will familiarize ourselves with the terms that are most often used in this document. It also presents a number of terms and definitions that are utilized throughout the dissertation and are necessary to the reader's understanding. Some of these terms and definitions are based on GDPR, whereas others refer to technical standards or are explicitly specified for the purpose of this dissertation. **Appendix A** also demonstrates further essential sources, terms and contains a definition for them.

2.2.1 GDPR Anonymization levels

The GDPR norms represent a fundamental regulatory framework that delivers transparent control over the disposal of personal data in the EU. The two levels of anonymization introduced in the GDPR deliver full compliance with the GDPR. According to Article 4 (1) and Recital 26 GDPR, among the anonymization levels given, the most comprehensive description is presented for the pseudoanonymization level.

- **Definition: Anonymization**

  The principles of data protection should therefore not apply to anonymous information, namely, information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable (Recital 26 GDPR)

- **Definition: Pseudoanonymization**

  means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such

additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person (Article 4 GDPR)

Although reaching the complete anonymization level avoids numerous regulatory requirements and limitations, however the lack of method(s) to determine the specified anonymization level or threshold between given anonymization levels imposes restrictions on the data processor and third parties in achieving or estimating the specified anonymization level. Moreover, the limited representation and lack of explicit requirements for assessing the level of complete anonymization also impose considerable limitations on estimating the level of anonymization and decrease the ability to observe the rights of data subjects [184]. In addition, the delivered methods for assessing DPIA risks (Article 35(7) GDPR) do not provide the opportunity to estimate privacy risks considering various anonymization levels. Previously, in our study [147], as well as in other studies [165], the need to complement the definition of anonymization levels or the boundary between two levels has been emphasized. Also, some studies indicate the possibility of defining the boundary between complete anonymization and pseudonymization [19, 104, 85]. However, the proposed approaches are highly specialized and intended only for specific industries or domains. Despite this, according to our study [149] survey among various stakeholders, we identified that complete anonymization is preferable to pseudonymization.

2.2.2 Personal data

The concept of personal data is one of the essential definitions established in Article 4(1) GDPR; in the following chapters, we can use sensitive data, personal information, and data of individuals in addition to “personal data”. However, the term “sensitive data” or “individuals data” inherits the definition of personal data presented above, as well as additional interpretations of personal data declared in Article 9 GDPR (special categories of personal data), as well as the national identification number (Article 89 GDPR) and relevant information. Before the introduction of GDPR to the public, the concept of personal data was defined and enshrined in Article 29 Working Party (Art. 29 WP). This organization ceased to exist in 2018 year before the GDPR entered into force. Nowadays, its functions continue to be performed by the independent European organization for the protection of personal data, namely the European Data Protection Board (EDPB). Simultaneously, personal data in the interpretation of Art. 29 WP, considering (Art. 29 WP Opinion 4/2007 on the concept of personal data), also does not deliver a clear explanation; however, it can be interpreted as: "information can be considered to "relate" to an individual when it is about that individual" [152]. We also introduce other definitions, b) and c), to highlight the originally stated definition "a" of personal data is very widely utilized and functional.

- **Definition: Personal data**

  a) means any information relating to an identified or identifiable natural person (’data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person (Article 4 GDPR);

  b) refers to any information relating to an identified or identifiable natural person (i.e., who can be directly or indirectly identified from publicly available data) now or in the future [121];

  c) ‘any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly...’ [144]

Comment: In addition to the primary definition of personal data, we also introduce other definitions, b) and c), to highlight that the stated initially definition of personal data is widely utilized and functional.

- **Definition: Non personal data**

opposite from personal data demonstrated in Rectital 26 GDPR, “data that is always non-personal (because it never related to an identified or identifiable natural person) and there is also data that once was personal but no longer is (as linkage to a natural person has been removed)” [57]

Figure 2.1 depicts the approach for assessing if the data corresponds to personal data. In other words, if personal data cannot reasonably be related to an individual or has never been related to an individual, according to the GDPR, they qualify as “anonymous information” [57]. However, in the context of the analysis and various options for interpreting the risks of identification by the data controller, the presented approach (extracted from Recital 26 GDPR) has several limitations. It does not provide a clear understanding of the data classification, thereby affecting the assessment of the level of complete anonymization.

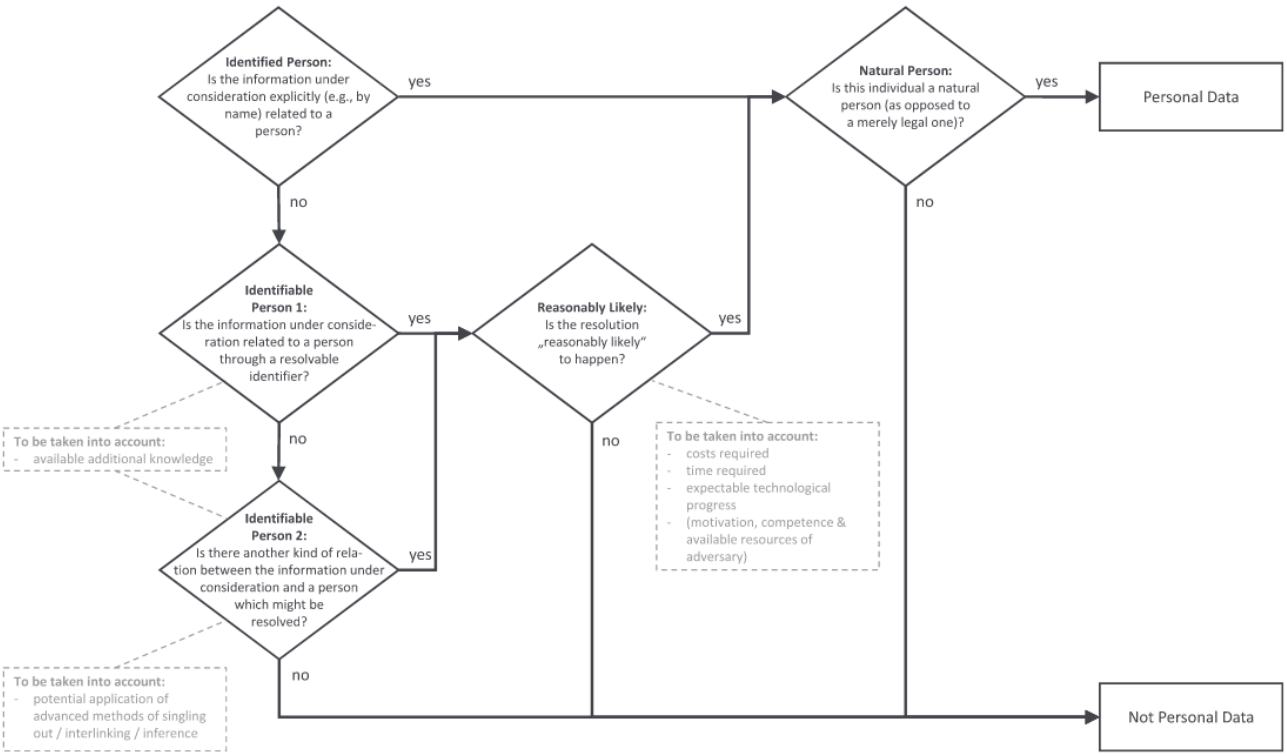


Figure 2.1 Defining Non-personal/personal data, assessment scheme [57]

- **Definition: Personally Identifiable Information (PII)**

  PII given by NIST [125] includes ‘information that can be used to distinguish or trace an individual’s identity’ and ‘any other information that is linked or linkable to an individual’ [144].

2.2.3 GDPR stakeholders

The GDPR contains various stakeholders, and **Figure 2.2** illustrates partly their relationships and connections, including how they associate with the different GDPR notations. In the presented visualization, interconnections between the data subject and DPO are excluded due to the study’s limitations [90]; however, the GDPR notations between the data subject and the Controller could be applied. In addition, stakeholders, namely third parties and Data protection authority (DPA), are not depicted; however, represented GDPR notations could be utilized; considering the definition of referred stakeholders, current stakeholders are not the focus of the present dissertation. However, we provide further a) the correspondent description and supplementary knowledge for a more reasonable overview and also an understanding of the domain-specifics and b) we utilize the proposals from DPAs and consider critical analysis with respect to the third parties. Simultaneously, it is worth considering that the number and variations of privacy risks can vary for different stakeholders [129] as well as methods or approaches aimed at reducing privacy risks.

Example	Category
Name and surname,	personal data
A home address,	personal data
Email address, such as name.surname@company.com,	personal data
An identification card number	personal data
Location data (for example the location data function on a mobile phone	personal data
An Internet Protocol (IP) address	personal data
A cookie ID	personal data
The advertising identifier of a phone	personal data
Data held by a hospital or doctor, which could be a symbol that uniquely identifies a person,	personal data
A company registration number,	not personal data
An email address such as info@company.com.	not personal data
Anonymized data,	not personal data

Table 2.1 Example of personal and nonpersonal data, (EU, 2020)

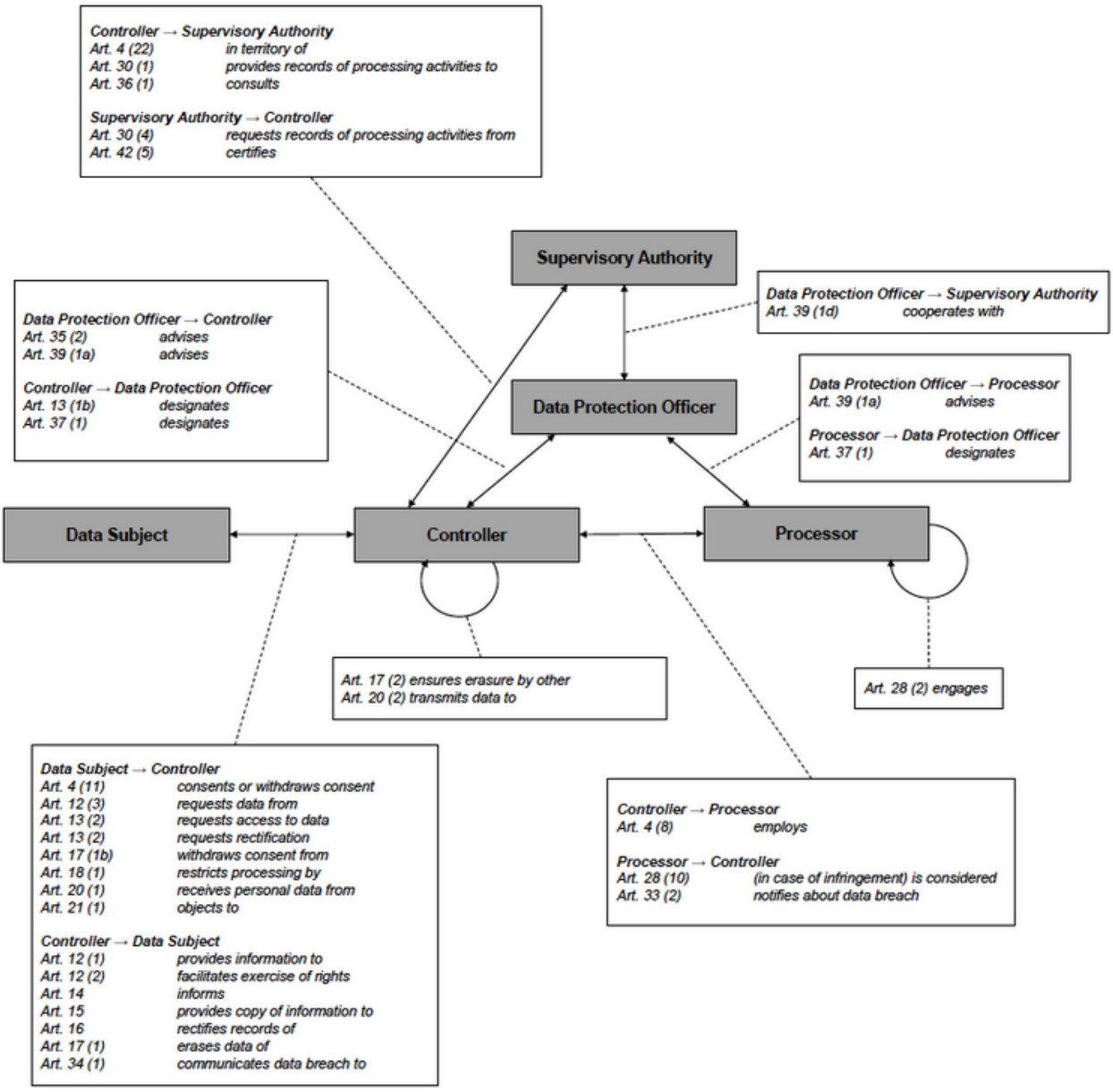


Figure 2.2 Stakeholders overview and interconnections considering GDPR notations [90]

2.2.4 Roles in the GDPR

As mentioned previously, the number and interconnections between various stakeholders, including their roles in the context of GDPR requirements, can vary dramatically. The roles presented below are a fundamental part of GDPR compliance, as they indicate who and how is responsible for the protection of personal data and also entitle data subjects to exercise their rights [16]. Simultaneously, the principal roles and accompanying terminology are presented below, and disadvantages and further discussions will be extended in **Chapter 3**.

- **Definition: Data subject**

  identified or identifiable natural person (Article 4(1) GDPR);

- **Definition: Data controller**

  means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law (Article 4(7) GDPR);

- **Definition: Data processor**

  means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller (Article 4(8) GDPR);

- **Definition: Data protection authorities (DPA)**

  DPAs are independent public authorities that supervise, through investigative and corrective powers, the application of the data protection law (Article 4(16), Article 51-59, Recitals 117 - 123 GDPR);

- **Definition: Third party**

  means a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data (Article 4(10) GDPR);

- **Interpretation: Data protection officer (DPO)**

  represents a role ensuring compliance and implementation of GDPR requirements in the context of personal data protection in organizations [155];

- **Definition: Supervisor Authority**

  independent public authorities to be responsible for monitoring the application of this Regulation, in order to protect the fundamental rights and freedoms of natural persons in relation to processing and to facilitate the free flow of personal data within the Union (Article 51 GDPR);

2.2.4.1 Roles in GDPR adaptation

Despite the variety of roles in the context of GDPR, numerous works [98, 180] point out the complexity of interpreting roles in the context of GDPR utilization and also following effective adoption of technical measures [100]. Put differently, the complexity of interpreting the roles, as well as their relationships, highlights the sophistication of implementing technical data privacy measures, which are also complicated by the interpretation of legal requirements. On the other hand, this imposes restrictions on organizations regarding compliance with GDPR norms, even while utilizing various PETs, which suffer from adequate usability or understandability knowledge [100]. In [149], we identified relevant groups and stakeholders which influence the adaptation of PET, as well as the transformation, implementation and monitoring of GDPR requirements; we also deliver more detailed results and critical discussion in the following **Chapters 3** and **4**.

2.2.5 Taxonomy in the GDPR

Historically, the development of a privacy taxonomy and matters of protection and assessment of personal data is one of the active areas to this day. Before the GDPR introduction, offered approaches to protecting data subjects were characterized and described in the document Art. 29 WP; after various revisions and adaptations, they were comprised concerning the GDPR, which took effect in 2018. However, despite various transformations of legislative notation in the EU, many questions still remain open in the context of compliance and fitting with the existing compliance standards imposed on data subjects. From the technical perspective, many questions are related to the data utility, as well as the relative methods used to anonymize personal data. Considering Article 32 GDPR the controller must apply and use the necessary technical and organizational methods to meet the security level of the relevant risks [121]. In the context of the legal ecosystem, the interpretation of the requirement is complicated by the lack of technical requirements, as well as by multiple disputes about how to assess risks [144] or the inefficiency and difficulty of applying the so-called hybrid approach (HA). To ensure effective implementation of referred requirements, many approaches have been suggested, where the most common is the risk assessment approach or the so-called risk management; however, the lack of a unified approach for risk assessment still remains one of the research gaps [144]. In addition, as the results of many studies indicate, the need for a unified approach to assessing compliance risks is also speculated by many privacy researchers, where one of the many directions is to determine a unified approach in analyzing compliance risks, as well as interpreting it consistently in both the technical and legal ecosystem [127].

- **Interpretation: Data minimization**

  Article (5) GDPR, which refers to limiting personal data collection,storage, and usage to data that are relevant, adequate, and more importantly necessary for carrying out the purpose for which the data are processed [73];

- **Interpretation: Identification**

  The term identification is not defined in the context of the GDPR [152], however, according to Recital 4(1) GDPR it refers to "identified and identifiable natural person";

- **Interpretation: Re-identification/ (sometimes deanonymization)**

  one of the possible terminologies used to describe privacy attacks on a dataset containing personal data; In other words, it is understood that the identity of the data subject may be (derived) extracted from the data set [144];

- **Interpretation: Privacy**

  is a dynamic concept that depends on changing technological and social aspects in the context of the utilization of various compliance standards [75], which is also influenced by the expansion of various attacks or technologies;

The GDPR regulations do not specify on what grounds it is possible to achieve complete anonymization [88], however, with the help of the additions Art. 29 WP Opinion 4/2007 four further terms can be supplemented:

- **Definition: any information**

  clearly signals the willingness of the legislator to design a broad concept of personal data. This wording calls for a wide interpretation. From the point of view of the nature of the information, the concept of personal data includes any sort of statements about a person. It covers "objective" information, such as the presence of a certain substance in one's blood. It also includes "subjective" information, opinions or assessments;

- **Definition: relating to**

  In general terms, information can be considered to “relate” to an individual when it is about that individual;

- **Definition: Identified or identifiable**

  In general terms, a natural person can be considered as “identified” when, within a group of persons, he or she is "distinguished" from all other members of the group. Accordingly, the natural person is “identifiable” when, although the person has not been identified yet, it is possible to do it (that is the meaning of the suffix "-able"). This second alternative is therefore in practice the threshold condition determining whether information is within the scope of the third element. Identification is normally achieved through particular pieces of information which we may call “identifiers” and which hold a particularly privileged and close relationship with the particular individual;

- **Definition: natural person:**

  the protection afforded by the rules of the Directive applies to natural persons, that is, to human beings. The right to the protection of personal data is, in that sense, a universal one that is not restricted to nationals or residents in a certain country. Recital 2 of the Directive explicitly makes this point by stating that “data processing systems are designed to serve man” and that they “must, whatever the nationality or residence of natural persons, respect their fundamental rights and freedoms”;

2.3 GDPR and PETs

Historically, Privacy-enhancing technologies (PET) have been a reliable tool for effectively protecting personal data and increasing the confidentiality of data subjects. According to the definition [197]: *"Privacy-Enhancing Technologies is a system of ICT measures protecting informational privacy by eliminating or minimising personal data thereby preventing unnecessary or unwanted processing of personal data, without the loss of the functionality of the information system*” which has been widely adapted and corresponds to some extent with the ENISA revisions. Simultaneously, PETs are associated with the so-called Technical and Organizational Measures (TOMs) “to ensure and to be able to demonstrate that [personal data] processing is performed in accordance with” the GDPR (Art. 24 GDPR) [192]. In addition, PET could be represented by various solutions (hardware or software) that aim to extract data without compromising security or privacy, and PETs also depend on the privacy use case [80].

Many scientists and experts point out the importance of using PET to enhance personal data protection for various endeavours and use cases. Moreover, since GDPR is aimed at protecting PII [124], utilization of PETs enables sufficient and promising guarantees of GDPR compliance, where PETs adaptation also fulfil various ISO standards (e.g. ISO/IEC 29100: 2011) in the data protection domain [69, 33]. In other words, PET implies numerous technologies aimed at safeguarding personal data, including ensuring control over the specified privacy level. In addition, in our study we identified [149] PET as one of the approaches that enable both improving the degree of protection of personal data and fulfilling the GDPR norms. However, as the study’s results [204] indicates, utilizing PET in safeguarding personal data has issues since not one of the PETs is guaranteed to solve all the confidentiality complexities and principles. In addition to conducting DPIA/PIA on IT systems with various PETs [195], it is not a trivial task and requires much effort. In the studies presented earlier [149, 65], various stakeholders responsible for GDPR execution were interviewed, where PET plays a significant role in fulfilling the GDPR norms, and the importance of issues w.r.t data subjects privacy is highlighted. The results [149] regarding PET adaptation and interconnection to the anonymization levels, will be introduced and explained in more detail partly in **Chapter 3** and **4**. A more detailed outline and examination of the selected PET DP is explained in *Section 4.1*. Also, introduced data protection goals *Section 2.5* will be explained in **Chapter 4** during the development process of a reference model.

2.3.1 Privacy by Design

Growing concerns regarding protecting sensitive data among organizations and academia have led to the development of the concept of Privacy by Design (PbD). PbD has different connotations and is represented in academia in different ways. However, it's summarized in seven principles [29, 32], which facilitate more reasonable privacy while embodied into technologies or systems:

- Proactive not Reactive; Preventative not Remedial
- Privacy as the Default
- Privacy Embedded into Design
- Full Functionality – Positive-Sum, not Zero-Sum
- End-to-End Security – Lifecycle Protection
- Visibility and Transparency
- Respect for User Privacy

The concept and philosophy of PbD consider privacy requirements in the context of developing new IT systems from the earlier (concept) stage to their implementation [29]. Put differently, following the PbD concept enables practitioners to improve the privacy of IT systems at different stages [87], enabling them to reach the critical advantages for business tasks [31], as well as meeting legal requirements in the context of the technical solutions used. Nevertheless, the presented high-level PbD principles were given for various types of systems, but this is not enough for multiple stakeholders to design and develop the above systems [143]; in other words, there is a lack of specific "how" guidelines to apply PbD guidelines are reduced to the fundamental question "what", leaving the "how" question to the discretion of the academia or other experts.

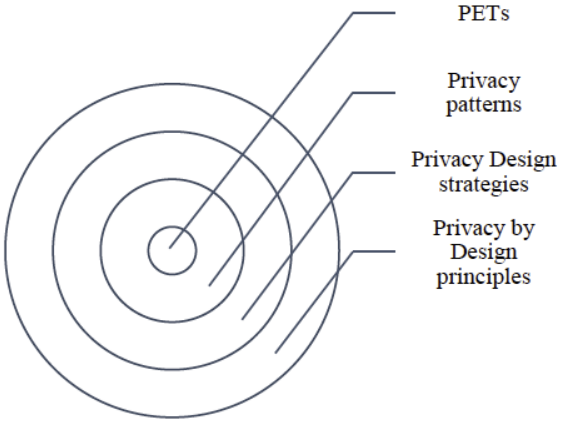


Figure 2.3 Privacy by Design layers [128]

In addition, as the authors note [87], referred principles have limitations, and not all of the principles presented have an impact on the processes of privacy design strategy. Current privacy standards or regulations [29] indicate that organizations face privacy issues and personal data confidentiality should proactively think about privacy issues (e.g. adapting the PbD) rather than in a reactive one [95]. As the authors point out [176], introduced complexity could be partly formulated and stem from the lack of the necessary level of expertise and various stakeholders in the privacy or security domain. In addition, as it claimed [12], the PbD approach does not provide the required explicitness for mapping legal data protection requirements into developed/maintained systems and components. As a response to the debated PbD

matters, privacy design strategies [87] was introduced, which in turn can be integrated to a large extent utilizing privacy patterns [113]; however, it is worth noting that referred privacy patterns depend primarily on the PETs implementation (**Figure 2.3**). In other words, the PbD complexity consists of many granular approaches, with PET at its core.

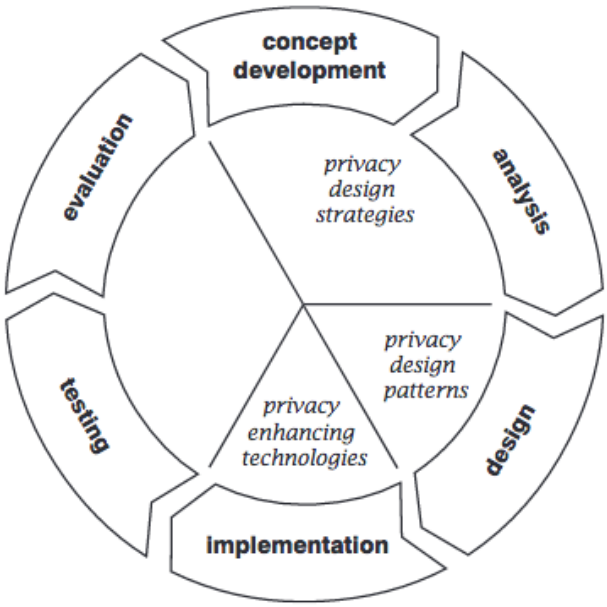


Figure 2.4 SDLC relationship to the strategies, patterns and technologies [87]

Also, according to Article 25 (1) GDPR, the need for a PbD approach is justified and recommended, and the need for a) technical and organizational measures (TOM), b) pseudo-anonymization and c) compliance with the DPG is suggested. On the one hand, Article 25 (1) GDPR suffers from many shortcomings [25]: a) vagueness and difficulty in interpreting the language and b) complexity of terminology, which imposes restrictions on adapting the PbD approach. On the other hand, Article 25 (1) GDPR and referred recommendations should be seen as the beginning of a possible dialogue between various stakeholders, which is aimed at ensuring interests (technical and organizational aspects) related to privacy [105]. In line with Article 25 (1) GDPR, Recital 78 GDPR suggests that "technical and organizational measures should be taken", and referred measures must comply with the Data Protection by Design or DPb Default requirements. However, DPbD is an approach built on PbD concepts, the principles of which are integrated into TOM on the data controllers' side [198]. With the introduction of the GDPR, the DPbD approach gained its significance and functionality, but many questions related to privacy issues still remain open [198]. While debating privacy issues, practitioners continue to look at DPbD less holistically. Moreover, the focus has shifted to PET as a technology that is utilized for more effective mitigation of privacy issues.

Despite possible previously highlighted limitations and mitigating referred issues, authors [87] offer another perspective for adopting PbD principles and concepts (**Figure 2.4)** through the Software Development Life Cycle (SDLC), which enables a more accurate presentation of recommendations for the artefact being developed (**Chapter 4**). The approach utilised by the authors is based on the framework [183], which represents 4 degrees (**Figure 2.5**) of privacy-friendliness interpreted via identifiability measure (namely identified, pseudonymous, anonymous). The first and second stages can be achieved using privacy-by-architecture and others using privacy-by-policy. However, it is worth noting that systems that fit the privacy-by-architecture description do not process sensitive information and, therefore, should not be privacy-by-policy [183, 87]. We demonstrate and discuss various frameworks and consider the referred limitations in **Chapter 3**.

Privacy stages	identifiability	Approach to privacy protection	Linkability of data to personal identifiers	System Characteristics
0	identified	privacy by policy (notice and choice)	linked	• unique identifiers across databases • contact information stored with profile information
1	pseudonymous		linkable with reasonable & automatable effort	• no unique identifies across databases • common attributes across databases • contact information stored separately from profile or transaction information
2		privacy by architecture	not linkable with reasonable effort	• no unique identifiers across databases • no common attributes across databases • random identifiers • contact information stored separately from profile or transaction information • collection of long term person characteristics on a low level of granularity • technically enforced deletion of profile details at regular intervals
3	anonymous		unlinkable	• no collection of contact information • no collection of long term person characteristics • *k*-anonymity with large value of *k*

Figure 2.5 Framework for Privacy-Friendly System Design [183]

2.3.2 Differential Privacy and GDPR

Differential Privacy (DP) has been proposed and is considered one of the reliable and reasonable PET methods for safeguarding personal data in the context of statistical analysis [49]. Simultaneously, the refereed method was also suggested and is considered as a method for advancing the issues with anonymization level considering GDPR norms [88]. Among the offered and various DP forms (central, shuffle or local), the most vulnerable in the context of GDPR compliance is the Central DP form [209]. However, as indicated by various studies [88, 38], considering additional analysis carried out in our research [149], the specified PET is associated with further issues and limitations such as (a) curator versus data holder, and b) accurate and false identification, considering various GDPR norms. Simultaneously, any approach that ensures complete anonymization implies a prohibition or impossibility of identification [88], as well as an adequate presentation under what conditions complete anonymization can be achieved or in which cases complete anonymization is violated. In other words, technically, this means that in the event of an attack, this privacy attack is aimed at revealing the identity. However, DP does not offer an adequate presentation of cases or settings in which complete anonymization may be violated; on the other hand, DP contains all the essential functionality (privacy parameters) to protect against identification.

2.4 Data Protection Impact Assessment (DPIA)

Although there are many opinions and approaches for assessing the privacy risks of personal data, the GDPR offers explicitly only one approach, namely the DPIA, partially presented and described in Article 35(7) GDPR. In addition, as demonstrated (**Figure 2.6**), this approach is partially limited since its application is limited only to assessing high risks for the security of personal data [171].

- Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single

assessment may address a set of similar processing operations that present similar high risks (Article 5(1) GDPR);

In addition according to the Article 35 GDPR, the assessment shall contain at least:

- systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;
- an assessment of the necessity and proportionality of the processing operations in relation to the purposes;
- an assessment of the risks to the rights and freedoms of data subjects referred to in paragraph 1;
- the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation taking into account the rights and legitimate interests of data subjects and other persons concerned;

At the same time, as we identified earlier [149] in addition to other works [175], the DPIA approach does not have the necessary flexibility in application and risk assessment for the data subjects, thereby provoking organizations and third parties to use additional PET approaches.

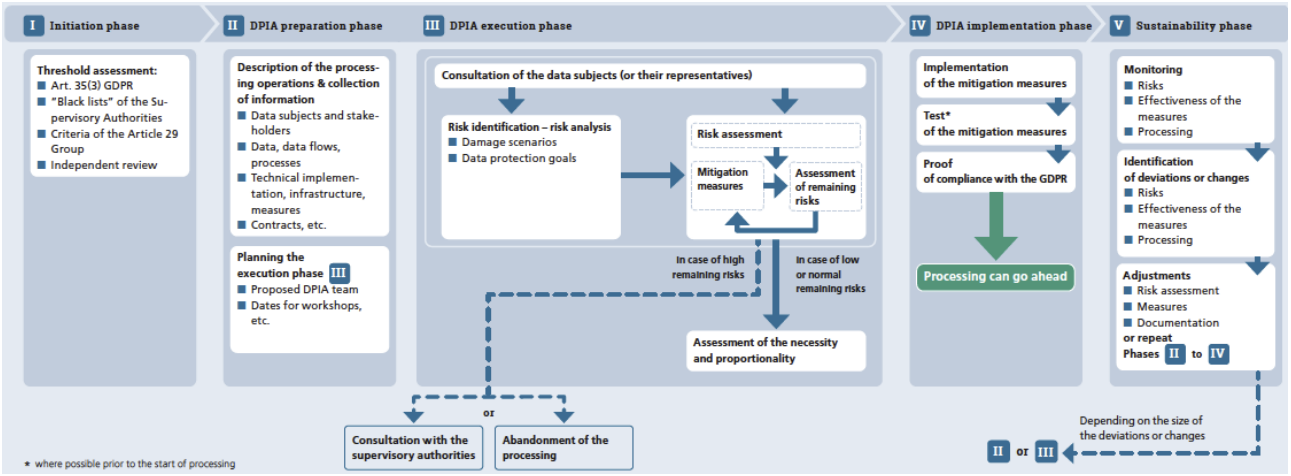


Figure 2.6 DPIA sequence [122]

2.5 Data Protection Goals (DPG)

As mentioned earlier, the GDPR contains DPG according to the following scheme such as (first priority): (1) availability, (2) integrity, (3) confidentiality and (second priority), (4) unlinkability, (5) transparency, and (6) intervenability, where such groundbase utilized and included in existing privacy frameworks such as CNIL [76] or others [132]. In the present dissertation, we concentrate on the DPGs introduced in the SDM framework, which, in addition to earlier presented DPG and stated in the GDPR, is complemented by supplementary DPG such as "Data Minimization" [162]. As demonstrated in **Table 2.1**, systematizes the legal requirements of the GDPR in the context of the studied DPGs. A more detailed and expanded description of DPGs and the complexity of DPG interpretations is presented in **Section 5.4**. Further, the demonstrated definitions (*Data Minimization, Availability, Integrity, Confidentiality, Unlinkability, Transparency and Intervenability*) in *Section 2.5* are extracted from the SDM framework[1] and explained to a better extent in **Appendix A**.

[1] https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methodology_V2.0b.pdf

- **Definition: Data Minimization**

  Data Minimisation covers the fundamental requirement under data protection law to limit the processing of personal data to what is appropriate, substantial and necessary for the purpose. The implementation of this minimisation requirement has a far-reaching influence on the scope and intensity of the protection concept determined by the other protection goals. Data minimisation specifies and operationalises the principle of necessity in the processing process, which requires of this process as a whole as well as each of its steps not to process more personal data than is needed to achieve the purpose of processing [...];

- **Definition: Availability**

  Availability refers to the requirement that access to personal data and their processing is possible without delay and that the data can be used properly in the intended process. For this purpose, the data must be accessible by authorised parties and the intended methods for processing must be applied to them. Availability includes the concrete retrievability of data, e. g. through data management systems, structured databases and search functions, and the ability of the technical systems used to present data appropriately for humans [...]. If, in exceptional cases, the protection of personal data with regard to availability is nevertheless violated, it must be ensured that measures are taken to rectify and mitigate the violation;

- **Definition: Integrity**

  Integrity refers, on the one hand, to the requirement that information technology processes and systems continuously comply with the specifications that were defined for them to perform their intended functions. On the other hand, integrity refers to the property that the data to be processed remain intact, complete, correct and up-to-date. [...];

- **Definition: Confidentiality**

  Confidentiality refers to the requirement that no unauthorised person can access or use personal data. Unauthorised persons are not only third parties outside the responsible body, but also employees of technical service providers who do not require access to personal data in order to provide the service, or persons in organisational units who have no connection whatsoever with the content of a processing activity or with the data subject. The confidentiality of personal data must also be ensured when the underlying systems and services are subject to unexpectedly high loads. [...];

- **Definition: Unlinkability**

  The protection goal of Unlinkability refers to the requirement that personal data shall not be merged, i. e. linked. It must be implemented in practice especially if the data to be merged were collected for different purposes. The larger and more meaningful the data base, the greater the potential greed may be to use the data beyond the original legal basis. Such further processing is only legally permissible under strictly defined circumstances. The unlinkability is to be ensured by means of technical and organisational measures. In addition to measures for pseudonymisation, other measures that allow further processing separately from the original processing are also suitable, both on the organisation side and on the system side. [...];

- **Definition: Transparency**

  Transparency refers to the requirement that both data subjects and system operators and competent supervisory bodies shall be able to identify to varying degrees which data are collected and processed when and for what purpose in a processing activity, which systems and processes are used to determine where the data are used and for what purpose, and who has legal responsibility for the data and systems in the various phases of data processing. Transparency is necessary for the monitoring and control of data, processes and systems from their creation to their erasure and a prerequisite for legally compliant data processing to which, where necessary, data subjects can give an informed consent [...];

Data Protection Goals	GDPR snippets
Data Minimization	Art. 5 para. 1 lit. c,e; Art. 25 para 2;
Availability	Art. 32 para 1 lit. b,c; Art. 33 para 3 lit. d; Art. 34 para 2;
Integrity	Art. 5 para. 1 lit. d, f; Art. 32 para 1 lit. b; Art. 22 para. 3,4; Recital 71; Art. 33 para 3 lit. d; Art. 34 para 2; Art. 32, 33, 34;
Confidentiality	Art. 5 para. 1 lit. f; Art. 28 para 3 lit. b; Art. 29; Art. 32 para 1 lit. b; Art. 32 para 4; Art. 38 para 5; Art. 33 para 3 lit. d; Art. 34 para 2;
Unlinkability	Art. 5 para. 1 lit. c;
Transparency	Art. 5 para. 1 lit. a; Art. 12 para. 1,3 - Art. 15; Art. 32, 33, 34; Art. 4 No. 11; Art. 7; Art. 4;
Intervenability	Art. 12 para. 2,6; Art. 5 lit. d; Art. 16; Art. 17 para. 1; Art. 18; Art. 20 para 1; Art. 22 para 3; Art. 25 para 2; Art. 33 para 3 lit. d; Art. 34 para 2; Art. 4 No. 11; Art. 7; Art. 4;

Table 2.2 Adapted interpretation of Data Protection Goals w.r.t GDPR notations [162]

- **Definition: Intervenability**

  Intervenability refers to the requirement that the data subjects' rights to notification, information, rectification, erasure, restriction, data portability, objection and obtaining the intervention in automated individual decisions are granted without undue delay and effectively if the legal requirements exist and data controller is obliged to implement the corresponding measures. Where the data controller has information enabling him to identify the data subjects, he must also take measures to identify and authenticate the data subjects who wish to exercise their rights [...];

In our recent study [148], the likelihood of interpreting the DPG and the DP concept was examined, and the outcomes from this study indicate the possibility of mapping the earlier suggested DPG w.r.t concept DP. We demonstrate and critically asses results from the study in the following **Chapter 4** while developing a reference model. More critical analysis w.r.t. DP and various DPG will be demonstrated in the following Chapters.

2.6 Conclusion

In this chapter, we presented key definitions related to personal and non-personal data with various supplementary knowledge, privacy taxonomy, various stakeholders, and roles, as well as activities related to implementation, monitoring and compliance with GDPR, including examples of personal information. We also present and discuss various anonymization levels and the DPIA approach declared in GDPR, supported by the auxiliary knowledge. In addition, we highlight and present DPG nominated in the GDPR and SDM framework and demonstrate and present PET and its advantages.

We also present and evaluate various studies and consider our interviews and survey results, which highlight the possibility of adapting DPG in the context of different levels of anonymization, as well as various DPG for different stakeholders, including the capability to adapt DPG for DP. Simultaneously, the survey results indicate the difficulties of interpreting anonymization measures and levels, including the need for additional methods such as PET for improvement or compliance with GDPR. In addition, the results of the study indicate the possibility of adapting DPG for DP in the context of different levels of anonymization, which can complement the existing DPIA approach and increase its effectiveness.

3 Related Works

The Kobbi Nissim and others [136, 116, 7] inspire this dissertation and work on issues of correct interpretation between legal requirements and technical performance in the context of DP method disposal and compliance requirements.

This chapter introduces various previously suggested and known approaches in academia and industry to address GDPR compliance and related works. First, in **Section 3.1**, we introduce the approaches in the academia that have been presented to comply with the GDPR recommendations and requirements. Further, we demonstrate and inspect given methods and approaches for compliance in the context of the examination area. Also, in **Section 3.2**, we present alternative approaches widely operated in the industry; we will also demonstrate the SDM framework developed in Germany, one of the most comprehensive and up-to-date operated in the industry. The referred framework is exhaustively developed and extracted from the GDPR norms. Thus, we can assume that the referred framework will likely recall all the requirements presented in the GDPR. In addition, we will discuss the results, approaches, and implementation methods, including the adaptation of previously obtained results from our works [147, 148, 149, 150]. Finally, we will consider alternative methods or approaches for the examined domain, followed by a critical analysis and summary.

3.1 Academia

Nowadays, academia offers a sufficient number of significant advancements in the privacy, data protection, or privacy engineering domains. However, not all comprehensively reflect and comply with the requirements and are entitled to the most practical integration of GDPR compliance. First, we present approaches that exhaustively facilitate GDPR compliance and subsequent integration. Later in this chapter, we critically debate other vital contributions to the studied domain.

3.1.1 PRIPARE

Preparing Industry to Privacy-by-design by supporting its Application in Research (PRIPARE), a project established in 2013 and finally delivered to the public in 2015, is considered one of the advanced methods for integrating existing privacy best practices (**Figure 3.1**) during the design process [39]. Nevertheless, the most prominent activity in the development of the project PRIPARE was conducted until GDPR came into force. Simultaneously, the current method and approach embrace and integrate existing privacy standards, various practices and research proposals and fully describe the seven methodological phases of privacy design (**Figure 3.2**).

The current method enables compliance with GDPR norms [137] while simultaneously playing a meaningful complement to Article 25 GDPR while utilizing various PET [164]. In addition, the PRIPARE approach has and prescribes 26 possible privacy patterns (Section **2.3.1**), which, as noted [26, 42] are practical in making design decisions, and can also serve as reusable solutions for frequently encountered problems. The suggested method is developed and describes six possible roles [39]:

System engineers: are responsible for enabling the realization of successful systems; they ensure that all aspects of a project or system are considered, and integrated into a whole, during the full system's lifecycle. They can be subdivided into: Business system analyst, System designer, System developer, UI designer and Tester.

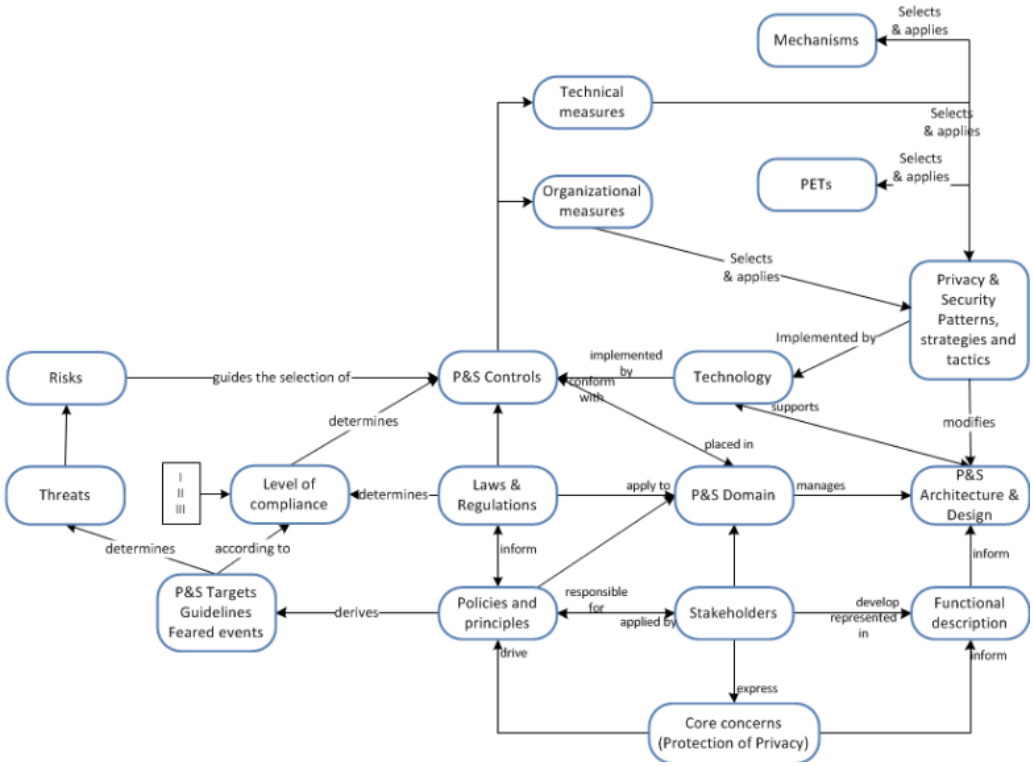


Figure 3.1 PRIPARE reference model methodology [39]
Note: Depicted model and integrated concepts in a single reference model highlight the complexity between concepts utilized and the interpretation between technical and organizational needs.

Privacy Security managers officers (PSMOs): the senior-level executives within organizations responsible for managing the establishment and maintenance of security and privacy procedures across the organization who address privacy and security issues and minimize their risks.

Data Protection Authorities: independent bodies which are in charge of monitoring the processing of personal data within their jurisdiction (country, region or international organization); providing advice to the competent bodies with regard to legislative and administrative measures relating to the processing of personal data; hearing complaints lodged by citizens with regard to the protection of their data protection rights.

Data subjects: people whose personal data are collected, held or processed.

Project managers: the senior-level executives within organizations responsible for making project-level decisions regarding scope, costs and schedule. They can be sub-divided depending on the role the organization has in relation to the system: System owners, System operators or System suppliers.

End users: people who make use of the engineered systems.

The PRIPARE method [39] is structured into seven phases that correspond to fundamental and generally accepted engineering systems design, which subsequently facilitates their subsequent integration with similar and generally accepted engineering practices:

1. **Analysis**: during this phase privacy and security issues or concerned must be discovered and reflected so they can be addressed during the design and implementation steps. The principal objective is to characterize the system from a security and privacy perspective.

2. **Design**: define the system's architecture, components, modules, interfaces in order to satisfy specified requirements Implementation: during this phase (transform the design into a built system) technology privacy and security principles and best practices must be followed.

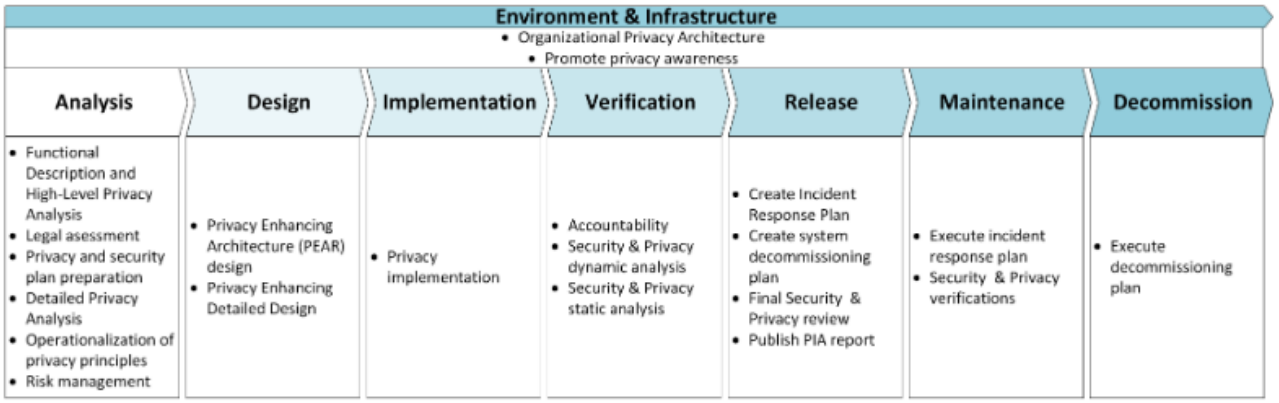


Figure 3.2 PRIPARE method (processes and phases) [39]

3. **Verification**: it must be ensured that the system meets privacy and security requirements through privacy and security testing, code reviews, audits.

4. **Release**: once the system is implemented and verified it is ready to be delivered or provided to the customer. During this phase it must be ensured that an action plan is present in order to respond to the discovery of privacy or security issues or breaches. Final security privacy reviews could be required prior to the release.

5. **Maintenance**: once the system is running and it is being used, responsibles must react to security privacy incidents ensure that privacy security processes are being enforced. Any evolution of the system during this phase must contemplate all previous steps and be incorporated into the system following the established security and data protection principles.

6. **Decommission**: it must be ensured that systems are correctly dismantled and that personal data is correctly treated accordingly to the current legislation and policies.

7. **Environment Infrastructure**: In order to support the proper application of PRIPARE methodology for the engineering of a system, the organization developing or operating the project must have an appropriate organizational structure, besides a reasonable level of privacy and security awareness, to support the various stakeholders during the methodology application. This phase is horizontal and it is not bound to any concrete system or project.

An additional phase (number 7, above) is contemplated to group privacy and security horizontal practices that should exist in any organization and should be independent of the engineering process itself.

Discussion The presented approach has a graphical visualization and is primarily aimed at developing processes associated with activity processes or process design (privacy and security domain). This approach includes various steps in the context of privacy design, while these requirements can be distributed simultaneously to technical and organizational needs; however, PET utilization and further integration are limited. It is worth noting that the approach considers various roles, but there are no transparent relationships between roles and following actions. Considering the requirements offered for both the technical and organizational areas and following the data subjects' requirements, the approach does not have or offer a more detailed and thorough analysis of these requirements and subsequent integration into given processes. In addition, considering the reference model methodology and presented phases, the approach needs to clearly define the possibility of incorporating DPIA or the capacity to develop multiple processes in parallel. Nevertheless, this issue can be supplemented with supplementary approaches, norms, standards, policies or methods. In other words, although the approach provides the opportunity to integrate the GDPR requirements, however, the interpretation and integration are limited.

3.1.2 Capability Model by Labadie, C., Legner, C.

The Capability model developed and presented to the public by the authors [109], is an approach based on the GDPR principles and is designed to help organizations implement GDPR requirements into their practices while providing a systematic approach to monitor the effectiveness of compliance with specified requirements. However, as the authors state, this model is not intended to "put out fires" caused by non-compliance with GDPR requirements. Put differently, this model utilizes various approaches (namely, data and resources) to provide capabilities that must be observed to comply with GDPR. Simultaneously, considering Article 24 GDPR, "appropriate technical and organizational measures to ensure and be able to demonstrate that processing is performed in accordance with this Regulation", this model has adopted two main capability groups: system and organizational groups (**Figure 3.3**). Accordingly, system capabilities mainly represent (Define Protected data scope, Manage Consent, Enable data processing rights) and are supplied by data processing systems, and organizational capabilities (Orchestrate data protection activities, Demonstrate complaint data processing, Disclose information) rely on the data protection processes and requirements. The offered model acts (represents) as a level of abstraction between normative aspects of regulation and specified compliance requirements [109]. Put differently, the authors attempted to explain "what" the organization needs to do rather than "how" it should be implemented (specific implementation). Earlier adoptation and validation of the referred model highlighted three shortcomings which the current approach can advance as declared and demonstrated by the authors [109]: a) lack of specific steps or instructions on what actions need to be taken considering compliance requirements, including an understanding of what problems they can solve, b) activities presented were self-isolated, and does not provide a clear indication "which" and on "what" level GDPR norms were achieved, which also was complex to track and forecast the future outcome, c) no documentation mechanism or systematization if reports, measures for utilized technical solution were put in place.

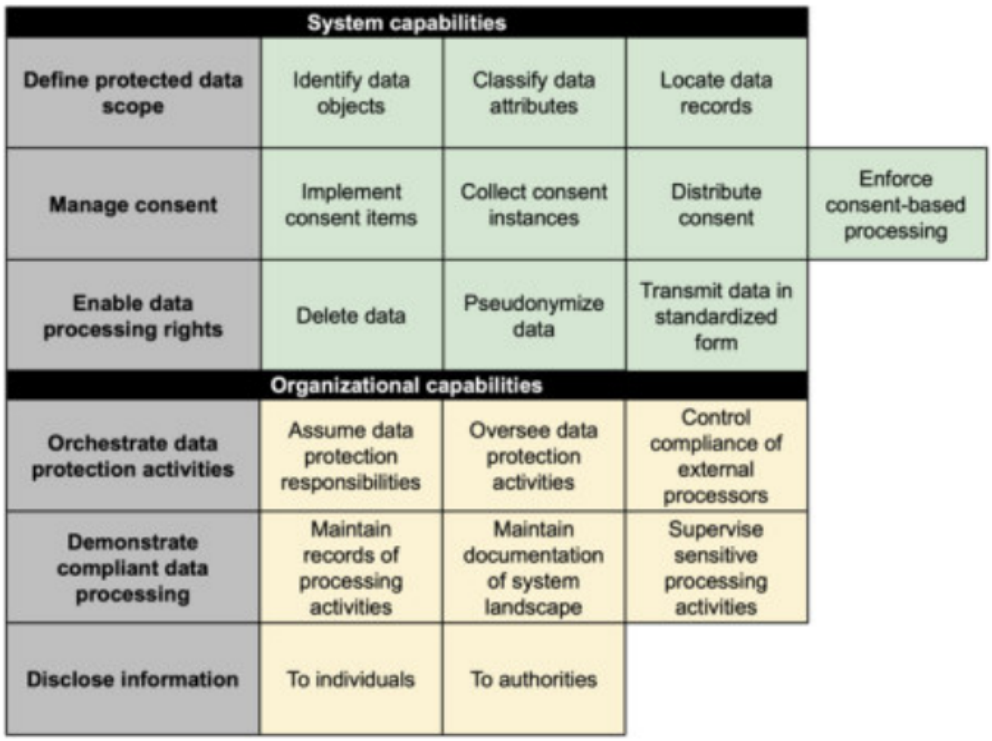


Figure 3.3 Capability model [109]

Define Protected Data Scope: Identify data objects: identify data domains and related data objects that fall within EU-GDPR's scope of applicability; Classify data attributes: assign levels of sensitivity

to data attributes contained within personal data objects; Locate data records: identify all storage instances of personal data objects and have the ability to access and retrieve them.

Manage Consent: Implement consent items: define and implement consent items that mirror data processing activities performed throughout business processes. Record consent instances: collect and record consent expressed by individuals Distribute consent: ensure consent items updates in all affected processing systems. Enforce consent-based processing: ensure that data processing activities are performed in accordance with consent expressed by individuals.

Enable Data Processing Rights: Delete data: permanently remove data records from their systems. Pseudonymize data: use pseudonymization techniques in order to adhere to the principle of minimization. Transmit data in standardized form: transmit personal data to external parties using standard formats and set up communication channels with other organizations.

Orchestrate Data Protection Activities: Assume data protection responsibilities: responsibilities for data protection-related tasks in all business functions that routinely process personal data. Oversee data protection activities: a leading role should oversee, organize, control and coordinate data protection activities. Control compliance of external processors: monitor that data processing conducted by third parties for EU-GDPR compliance.

Demonstrate Compliant Data Processing: Maintain records of processing activities: inventory and document personal data-related activities performed throughout business processes. Maintain documentation of system landscape: inventory and document systems that store and process personal data on a regular basis. Supervise sensitive processing activities: identify and evaluate sensitive data processing activities.

Disclose Information: Disclose information to individuals: provide individuals with complete and understandable information regarding the processing of their personal data and respond to their data protection-related requests. Disclose information to authorities: collaborate with designated data protection authorities and communicate relevant information upon request.

Discussion The presented model consists of six possible capabilities that clearly describe the studied area, forming a clear abstract concept for eliminating problems in complying with GDPR and identifying and prioritizing actions to achieve the established GDPR norms. The current model was introduced and has a graphical representation, highlighting impacted areas in four colour schemes, namely: red - not/insufficient covered, yellow - in progress, green - covered and white not applicable, which enhance understanding as a high-level picture for various stakeholders. Nevertheless, no specific notations were demonstrated regarding the connections between the roles and the following actions. However, the presented model can be largely inverted as an approach that represents an assessment of goals, with no clear step-by-step instructions or relationships between specific actions to achieve given goals. The supplement lacks any connection between alternative privacy approaches or the possibility of integrating this model into existing privacy engineering practices. In conclusion, as the authors state, this model is ineffective when used in so-called “fire fighting” situations; in other words, the application of the current model is limited and appropriate simply for monitoring the progress of measures applied.

3.1.3 Method by Koç H.

The method presented by the authors [102] is a method that contains specific activities that include a detailed set of actions (five phases) for GDPR compliance in the context of an EA perspective. It also incorporates the roles of the project team, the managing director, various business unit managers, and employees, followed by the IT operations. The phases of the method [102] include (**Figure 3.4**):

1. **Project preparation** (The aim of this phase is to define the scope of the GDPR project)

2. **Recording and categorizing the systems** (The aim of this phase is to identify the systems relevant to the GDPR project.)
3. **Definition of data categories** (The aim of phase 3 is to classify the data and data subjects)
4. **Data collection and validation** (In this phase, information about the processing activities in the organization should be collected and validated.)
5. **Creation of the register of processing activities** (VVT) (In this final phase, the main artifact of the method, the VVT, is created.)

Figure 3.4 Method by Koç phase's representation (in German) [102]

In other words, this method enables a more comprehensive a) structure of the organizational design considering the GDPR requirements in the form of an EA, b) a list of all applications that process personal data can be extracted from the architecture. In addition, the described method articulates organizational units, business objectives and processes, possible risks and control.

Discussion The presented method describes specific phases and actions for integrating GDPR requirements into IT tasks; in addition, it can be adapted for various stakeholders. Among the limitations, it can be noted that this method is no longer discussed in the academia, and practitioners also do not have a clear integration possibility with various ISO standards.

3.1.4 PRIAM

The authors [43] proposed a combined Privacy Risk Analysis Methodology (PRIAM) consisting of two main phases: (Information gathering and Risk assessment), however, shifting the focus additionally to individual

stakeholders, since assessing complex privacy requirements requires evaluate not only the tasks but also the requirements of various stakeholders or groups, including society as a whole.

1. **Information gathering** consists in gathering all the relevant information and the main difficulty for the analyst at this stage is not to overlook any factor that could have an impact on privacy risks.

2. **Risk assessment** uses the values of the attributes to compute the risk levels, which are pairs (severity, likelihood), for each privacy harm. This computation is based on harm trees which are used to establish a relationship among privacy weaknesses, feared events and harms.

The main advantage and subsequent limitation are that this approach is based on assessments of seven components (namely, the system, the stakeholders, data, risk sources, privacy weaknesses, the feared events and harms). A distinctive feature of this approach is that using various data attributes (in this case, used for private data) in combination with attributes of other components (such as the system) allows one to evaluate the categories of the most vulnerable among events and the corresponding confidential elements. By comparison, such relationships are poorly represented (in academia), and the ability to interpret them is limited, but PRIAM enables practitioners to overcome this limitation. In other words, this approach enables a clear connection between privacy harms and privacy weaknesses. Utilizing the referred approach enables systematically and trackably conducting privacy risk assessment tasks. However, this approach does not provide a list or clear description of countermeasures to counter privacy risks, leaving this for future work. This approach is applicable particularly to the following stakeholders, namely, the data subject, the data controller, the data processor or third parties.

Discussion Although the PRIAM approach can be applied with various limitations due to the absence of countermeasures, which is one of the GDPR requirements. The presented approach can be used as a supplementary extension for assessing system requirements and is also practical for privacy risk assessment, as it has more adaptable capabilities in the context of estimating critical factors in evaluating privacy risks.

3.1.5 Maturity Model for Data Protection

The reference and maturity models, as frameworks or templates used by organizations, can differ in both goals and characteristics. However, the Maturity model (MM), as a tool for assessing the effectiveness and capabilities of an organization in various processes, is more accurate than the reference model for specific data protection processes. We argue that the MM can also be utilized to assess the degree of compliance. Therefore, it is not just beneficial, but necessary to consider the MM presented further, as this will enhance the review of privacy practices more thoroughly and from another perspective.

The concept of the Maturity model has been around for a long time as a tool that provides/helps evaluate the activity of an enterprise's progress in a given period of time, including allowing the identification of reasonable measures to achieve the necessary goals or desired state [189]. Although, for a long time, MM did not have any durable definition; nevertheless, according to the results of work [103] assessing 16 different MMs, a conceptual representation was defined and demonstrated "phases of increasing quantitative or qualitative capability changes of a maturing element in order to assess its advances with respect to defined focus areas." In addition, the MM element can be an individual, an object or a system; MM also has the ability to explain changes descriptively or to adapt more practical solutions for specific stakeholders [101]. It is worth noting that further development of MM was introduced to the industry through the ISO/IEC 33001:2015 standard. However, for the data protection domain, MM, Certified Public Accountants (AICPA) and the Canadian Institute of Chartered Accountants (CICA) introduced MM under the name Privacy Maturity Model (PMM). A further development of the PMM was presented in the context of the Data Protection Management Maturity Self-Assessment model (Autoévaluation de maturité en gestion de la protection des données) of the CNIL.

In other words, "In data protection domain such models could enable a reliable indications to measure an organziationls data protection posture" [101], MM also enable to recognize the data protection activities

in an organization. Still, the main disadvantage is the lack of harmony in the examinations.

Discussion The presented approach is relatively innovative and does not have multiple feedback in academia; at the same time, it does not have a single taxonomy and definition of the concept. On the other hand, the ability to adapt the MM approach can be widely and practically adapted to the needs of compliance inspection due to the complexity of concepts in the data protection domain and GDPR. However, one of the takeyaways that defining a clear information or privacy taxonomy could benefit the utilization referred approach in academia. Simultaneously, while adapting the referred approach, it is necessary to carry out intensive work to supplement contextual interpretation for various use cases. Simultaneously, due to the lack of a clear definition of the concept and the relationship to the norms using the example of the GDPR, the assessment of the necessary actions or steps to fit the GDPR norms and vice versa will be problematic.

3.1.6 Other Contributions

Since the GDPR introduction, the academia has delivered multiple approaches aimed at monitoring, assessing or optimizing the needs of a company according to the requirements of the GDPR; however, despite numerous updates and recommendations, various organizations and stakeholders are still encountering problems in designing and maintaining various intensive systems that comply with the fundamental principles privacy data protection, including various extensions as PbD or DPbD.

Despite the previously demonstrated problems and multiple difficulties for compliance with DPbD, the authors [179] presented an architectural approach, which, according to the study, offers a systematic model of the driver approach and also contains a consistent action plan. The developed approach not only takes into account DFD notations [45] and privacy threats based on the LINDDUN framework [46] but can also incorporate and complement DPIA / DPbD exercises. In addition, in their review work devoted to the review of legal architectural requirements in the context of data protection design [47], the authors note “a clear dichotomy between architectural and legal approaches” in the context of personal data protection and not one of The considered approaches do not fully satisfy those specified in the GDPR. In addition, as the authors point out, the possibility of synergies and incentives for integration (technical and legal requirements) can be realized. Further, the authors [158], to enable better compliance and incorporate the technical solutions considering the GDPR norms, demonstrate well-established and so-called 13 patterns (Transparency and Traceability, Purpose Limitation, Data Minimization, Accuracy, Storage Limitation, Information Obligation, Right of Access by the Data Subject, Right of Erasure, Restriction of Processing, Data Portability, Notification Obligation, Privacy by Default) relevant to the IS as a set of guidelines. Moreover, as claimed by the authors, the number of patterns is insufficient due to the active development of the data protection domain, and it also does not provide sufficient details regarding the roles and their interconnections.

The authors [30] presented an alternative (step-by-step) approach in the context of interpretations of PbD and EA requirements. However, the presented approach, to some extent, contains the integration of the capability maturity model, but to a large extent, this the approach focuses on security rather than privacy issues. The authors’ alternative guidelines [30] focus on approaching the questions from the legal perspective; as indicated in the work, the authors presented a step-by-step process that includes various types of analyses, integration of DPMS, RoPA, and monitoring measures.

3.2 Industry

3.2.1 The Standard Data Protection Model (SDM)

SDM (version v2.0b) was introduced to the public by the German Data Protection Office in 2019 (since 2023, SDM is available in version three[1]). However, the SDM v.3 contains nominal changes to the pri-

[1] https://www.datenschutz-mv.de/static/DS/Dateien/Datenschutzmodell/SDM-Methode-V31.pdf

mary v2.0b version presented earlier. SDM is a so-called framework aimed for the protection of personal data and control of GDPR recommendations based on the DPG; SDM represents and provides the means to support, establish and evaluate TOMs to satisfy and demonstrate that personal data is processed by all GDPR requirements [163]. In other words, SDM offers the possibility and procedure for establishing and evaluating TOMs that enable DPbD, as presented and not limited to the application considering MM [101]. Simultaneously, as mentioned earlier, SDM identify and are based on seven DPGs (*Data minimization, Availability, Integrity, Confidentiality, Unlinkability, Transparency, and Intervenability*) presented earlier; moreover, complete DPG descriptions are given in **Appendix A**), but SDM as a whole framework is not limited to the DPG offered and represents the opportunity to offer further procedures for assessing the control and protection of personal data [101].

The seven DPGs presented in SDM not only systematize data protection requirements that comply with the GDPR norms but also enable practical application in both technical and organizational contexts. In addition, DPG presents reflective protection objectives in the field of information security, which have been utilized and validated for considerable time. In other words, the advantage of using and adapting DPGs shown in SDM is mainly due to the simplified representation and modelling of functional requirements for various practical implementations, as well as the visualization of possible conflicts. Simultaneously, these DPGs make it possible to systematically integrate legal requirements and GDPR into technical and organizational measures. In addition, it represents the ability to interpret requirements and measures between technical and organizational areas. SDM utilize three conceptual layers: 1) business layer, 2) application layer, and 3) IT infrastructure layer. To apply SDM in the context of DPM, SDM suggests using the adopted plan-do-check-act (PDCA) cycle (**Figure 3.5**). PDCA cycle consists of four generally accepted steps:

1. **Plan** (Plan / Specify / DPIA / Document).
2. **Do** (Implement / Log).
3. **Check** (Check / Validate / Evaluate).
4. **Act** (Improve).

SDM framework is relatively closely related to the German "Grundshutz of BSI" (Federal Office for Information Security). At the same time, SDM only discusses methods for IT security based on a) DPG confidentiality, integrity and availability and b) process modelling for processing activities, thereby representing a tangible connection between the two concepts and the practical linkage between the two standards and also complements in the context of efficiency. However, it is worth considering that SDM is more aimed at protecting individuals, although IT security is aimed at protecting organizations.

Discussion The SDM framework is structured and reflects the GDPR requirements based on the proposed DPGs. In addition, each of the presented DPGs, according to the SDM, can interpret a specific GDPR part or document (Article or Recital). Although SDM is a relatively new development, referred framework is not widespread among academia but is mentioned in some scientific studies, mainly in Germany. On the other hand, since the presented work was published in the academia and also continues to be supplemented according to the needs of stakeholders and the academic community, this framework, thereby allowing many companies to avoid various penalties (GDPR) in the context of personal data protection, this indicates significance in the area under discussion. At the same time, this framework is aimed only at a particular group of stakeholders, such as DPO, thereby providing limited support to other stakeholders associated with or having the ability to improve the degree of personal data protection, namely software and privacy engineers and privacy legal experts.

3.2.2 ISO 27001

ISO stands for “International Organization for Standardization”, which was founded in Geneva in 1946 year. Since its founding, more than 19,500 international standards have been published. To a large extent,

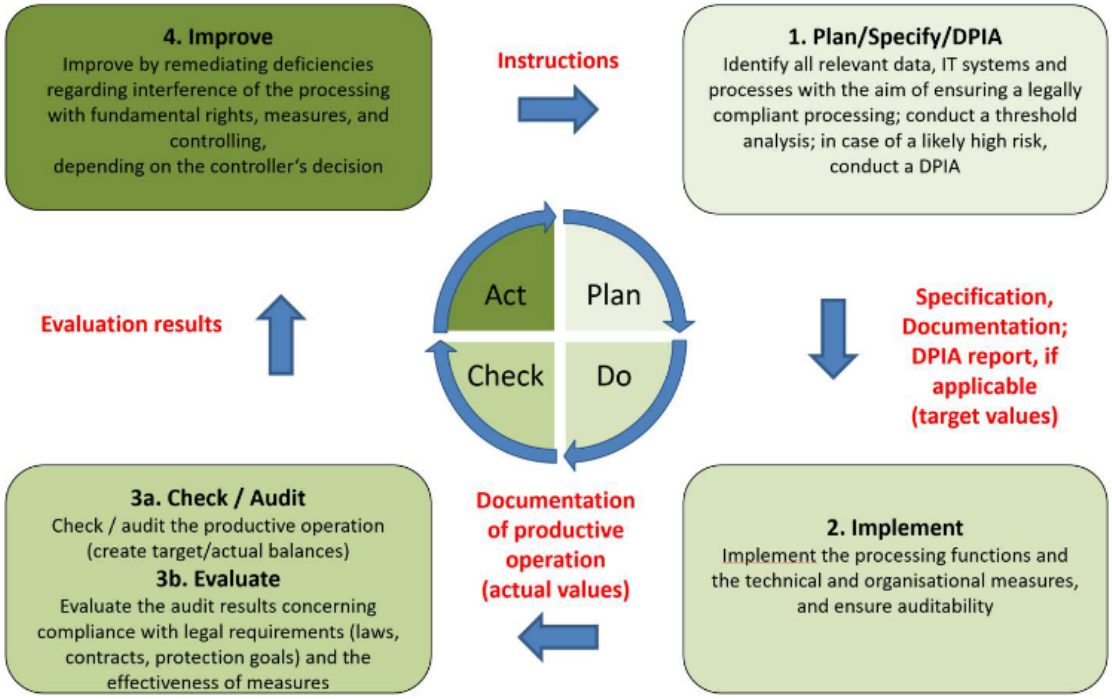


Figure 3.5 SDM framework v2.0b PDCA visualization [163]

the advantages of these standards benefit the industry and, to some extent, the consumers [44]. ISO/IEC family 2700x standards are a set of information security standards that represent current recommendations and the most popular, best practices in the context of information security management [35, 118]. ISO/IEC 27001 is a set of recommendations and is one document that allows organizations to achieve results in fulfilling the requirements (organizational and technical) to reduce the risk of hacking and establishing an information security management system (ISMS). In addition, ISO/IEC 27001 is an international framework with a long history indicating multiple additions and advancements, as well as the ability to interpret results with other ISOs and other frameworks (COBIT example Section **3.2.3**). The result of ISO/IEC 27001 (in the context of integration in a company) can be an adequate ISMS model (having the following properties: install, implement, operate, revise and manage). The ISO/IEC 27001 standard adapts the implementation PDCA cycle as a widely known and applicable approach [27] and proposes next steps/stages:

Plan: 1. Define the scope of the ISMS. 2. Define the information security policy. 3. Define a systematic approach to risk assessment and the risk acceptance criteria. 4. Carry out a risk assessment to identify, within the context of the policy and ISMS scope, the important information assets of the organization and the risks to them. This is where you assess the risks. 5. Identify and evaluate options for the treatment of these risks, selecting,where required, the control objectives and controls to be implemented. 6. Prepare a statement of applicability.

Do: 1. formulation of the risk treatment plan and its documentation, including planned processes and any required supporting documentation; 2. implementation of the risk treatment plan and planned controls; 3. appropriate training for affected staff, as well as awareness programmes; 4. managing operations and resources in line with the ISMS; and 5. implementation of procedures that enable prompt detection of, and response to, security incidents.).

Check: 1. monitoring, reviewing,testing and audit;

Act: 1. improvements to the ISMS should be identified, documented and implemented.

Also, ISO 27001 has a continuation in the context of the ISO/IEC 27701:2019 standard, which introduces Personal Information Management Systems (PIMS) as an extension of ISMS, allowing for expanded data protection requirements [112] by recalling and refines clauses from ISO/IEC 27001:2013 and ISO/IEC 27001:2013 and making them applicable for data protection domain. Since PIMS is based on and is a

continuation of ISMS, we will consider ISMS.

Discussion ISO 27001 serves as one of the most popular and practical documents that provide protection against security incidents and deliver the highest degree of data security for various types of organizations (without distinguishing whether the data is personal or not). On the other hand, this imposes restrictions on the rights of data subjects and, accordingly, compliance with GDPR standards. At the same time, the presented extension to ISO/IEC 27701 expands the concept of personal data protection. But as the authors note [3], in a practical context, full compliance can only be achieved using additional approaches or assessments inherent in the maturity model, which may be one of the disadvantages. Although further adaptation and integration of this standard, including that recommended by DPA CNIL (https://www.cnil.fr/en/iso-27701-international-standard-addressing-personal-data-protection), can serve as a development for more successful analysis and integration of data protection requirements in this context [112].

3.2.3 COBIT

Control Objectives for Information and Related Technology (COBIT) is one of the widely known and used frameworks by IT professionals in the field of IT governance to address issues related to security and privacy. Initially, the framework was introduced in 1996 and is designed in such a way that it can be independent of the progress of technology and industry and, at the same time, can be integrated with various ISO/IEC standards. The framework was proposed by ISACA (Information Systems Audit and Control Association), the latest version of the COBIT v.5 framework, which provides in-depth analysis and guidelines for IT governance depending on organizations' goals [91]. COBIT consists of 6 principles (**Figure 3.6**) introduced by author [67], and also introduced as a process reference model and developed in a way to be interpreted and applicable by various stakeholders. However, it does incorporate implementation guidance. In addition, the current framework is not well presented in academia, and only a few studies have investigated the COBIT utilization and benchmarked referred framework [53, 74, 193]; nevertheless, considerable attention has been given by practitioners [157]. We will introduce the framework and main concepts based on work [67].

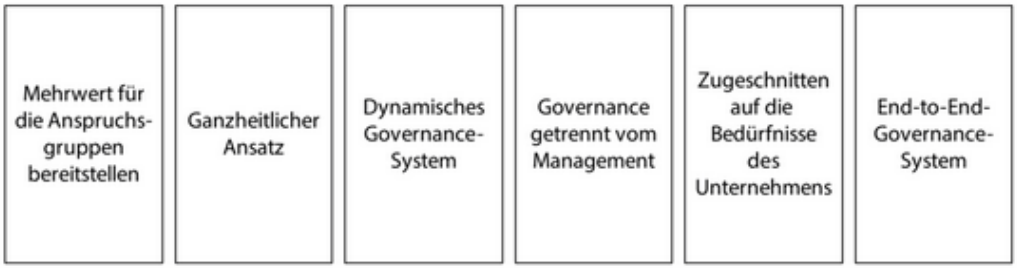


Figure 3.6 COBIT six principles (in German) [67]

COBIT, as a maturity model, is a tool for IT governance that enables measuring how managerial development processes are in accordance with the required goals [3]. Although the COBIT framework is comprehensive, as the academia points out, this complicates implementation for medium and small organizations [177]. As the authors note [186] COBIT 5 (latest version) retained the previous advantages of the framework but also expanded the possibility of interpretation by various stakeholders, holistic approach is now more organized and simplified, and the framework can be used by both internal and external stakeholders.

Discussion COBIT is designed and delivered to a large extent for managing and developing the IT governance model. The COBIT design enables it to be integrated with various ISO/IEC standards and supplementary methods. The COBIT delivers various roles and has the ability to precisely modify the requirements of each role. Nevertheless, the framework does not provide specific actions and solutions

Functional model	High level definition of all functional areas for IT Based on customer use case analysis
Service model	Based on ITIL, service fifecycle and high leve grouping of: Continuous Assessment, Continuous Integration, and Continuous Delivery -and later into Value Streams
Information model	Identification of key controlling IT artifacts Definition of artifact lifecycles according to lifecycle model
Foundational Integration Layer	Defines key control points for integration, based on artifact Link Information model with lifecycle model

Table 3.1 IT4IT Reference Architecture General approach [146]

for the data protection domain. However, the framework can address these issues by fulfilling the organization's needs in utilizing a COBIT conceptual framework. In other words, the COBIT structure is enabled to fulfil data protection tasks.

3.2.4 IT4IT

IT4IT, introduced to the public by the Open Group in October 2015 [97] - is an architectural approach that allows easy integration with IT solutions and performs a leadership role [9]. This approach is excellent for integrating various alternative approaches, such as ITIL and COBIT, and it includes significant flexibility in adapting new additions (services) or requirements to the current operating model. The approach's main novelty lies in presenting pre-existing processes based on a value chain perspective (**Figure 3.7**).

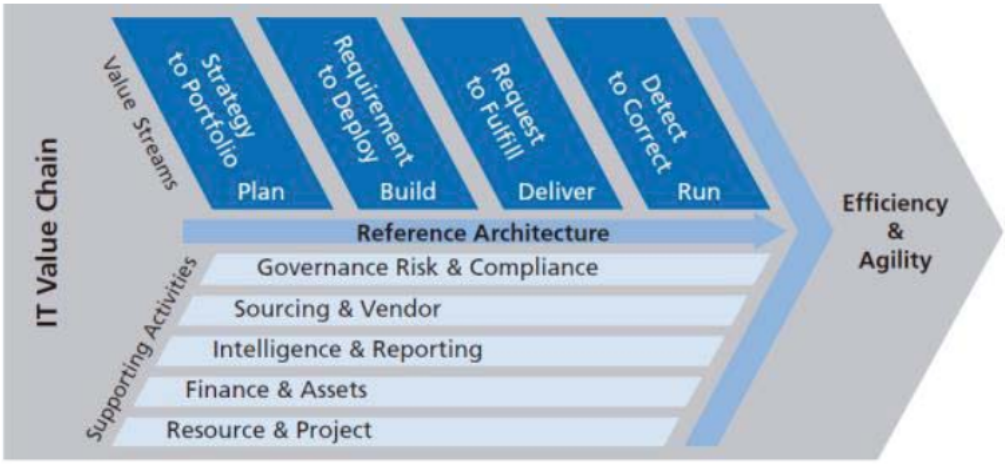


Figure 3.7 IT value chain [97]

The industry actively evaluated the presented solution, but interest can be described as limited [188]. According to [188, 8] combined, the architecture includes four functional models (**Table 3.1**), where each of the presented models is a self-sufficient element. However, the most effectiveness is achieved when utilized together.

It is important to note that IT4IT is considered as an independent approach, which differs from process and capability models, thereby being designed for different types of application scenarios [96]. It is also worth noting that the presented approach has five different levels of abstraction, which allows you to describe your own conceptual information models. In addition, IT4IT offers an entirely new approach; however, from an academic point of view, this approach lacks coverage or, in other words, insufficient implementations and lack of primary materials on this topic to carry out a complete analysis and integration cycle), including discrepancies and inaccuracies in the context of predicting results presents this approach as one of the most promising, but requiring additional research both in academia and use cases in industry [188].

Discussion Considering the GDPR norms, processing activities in most cases are related to or based on IT services; the reference architectural approach of IT4IT can contribute in the context of the GDPR recommendations through the use of abstract levels as well as various unique models presented earlier, allowing for high accuracy tracking and monitoring of development and implementation processes various services. The presented approach does not indicate obvious connections with various ISO standards or approaches, indicating the possibility of integrating DPbD or PbD. Although the presented elements in the IT4IT approach have a clear relationship and the possibility of interpreting recommendations or requirements for the organization, they do not relate to the field or domains of personal data protection. The next complication is that this approach is mainly aimed at tracking / (provisioning) IT services, where compliance criteria are unpleasant and are not presented or explained in the context of this approach. Therefore, IT4IT is clearly a valuable tool for organizations in the context of assessing IT infrastructure and services. However, it may be only partially a conceptual approach in the context of assessing compliance in considering GDPR norms.

3.2.5 Privacy Management Reference Model and Methodology (PMRM)

PMRM as a RM provides guidance for developing operational solutions to various privacy issues, including application to a variety of issues related to privacy management and compliance functionality. In other words, PMRM is a model or methodology that allows [166]:

- understand and analyze privacy policies and their privacy management requirements in defined Use Cases;
- select the technical Services, Functions and Mechanisms that must be implemented to support requisite Privacy Controls;

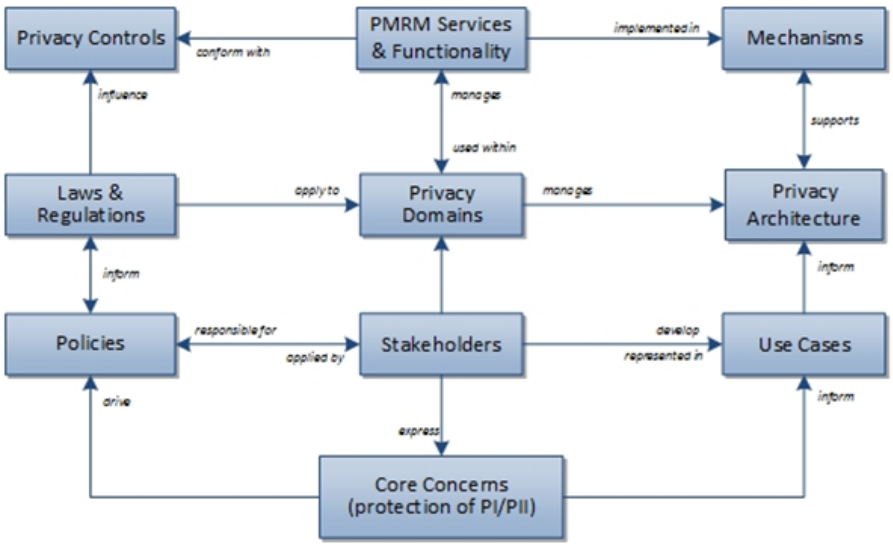


Figure 3.8 PRPM model and core concern of privacy protection and management utilized by various stakeholders [188]

The Privacy Management Reference Model and Methodology (**Figure 3.8**) considers multiple complex relationships between different systems, applications and devices, in the context of managing personal information PII according to various legal, regulatory or alternative requirements. In addition, the effectiveness of this model can be achieved in complex integrated information solutions and systems, applicable in various fields. The presented model is an effective tool for achieving privacy goals using the PbD approach and can be utilized by various stakeholders. Despite the fact that this model is not fixed (i.e. static, or a set

of rules or recommendations), the model has flexibility due to an individual approach to each user case. As the developers of this model indicate, this model is intended for a wide range of practitioners, and is also abstract.

Target stakeholders: a) Privacy and Risk Officers and Engineers, b) Systems/Business Architects, c) Software and Service Developers and d) Public policy makers and business owners.

PMRM includes a) a conceptual model of privacy management, and definitions of terms b) a methodology and c) a set of operational services, functions, including the relationships between the elements presented; The available methodology in the context of a high-level presentation (**Figure 3.9**), on the one hand, represents an organized approach to achieve the necessary goals, at the same time, this methodology and the subsequent steps presented are not absolute, accordingly, the choice of order is determined by the performer, however, the developers of the methodology indicate the need for iterative execution. In other cases, it is permissible to use part of the model; accordingly, some steps may be excluded in the context of developing and preparing the model. In addition, the presented methodology has already prescribed specifications, which includes a set of Services and Functions that are considered detailed and necessary to comply with the privacy policies and privacy controls being developed.

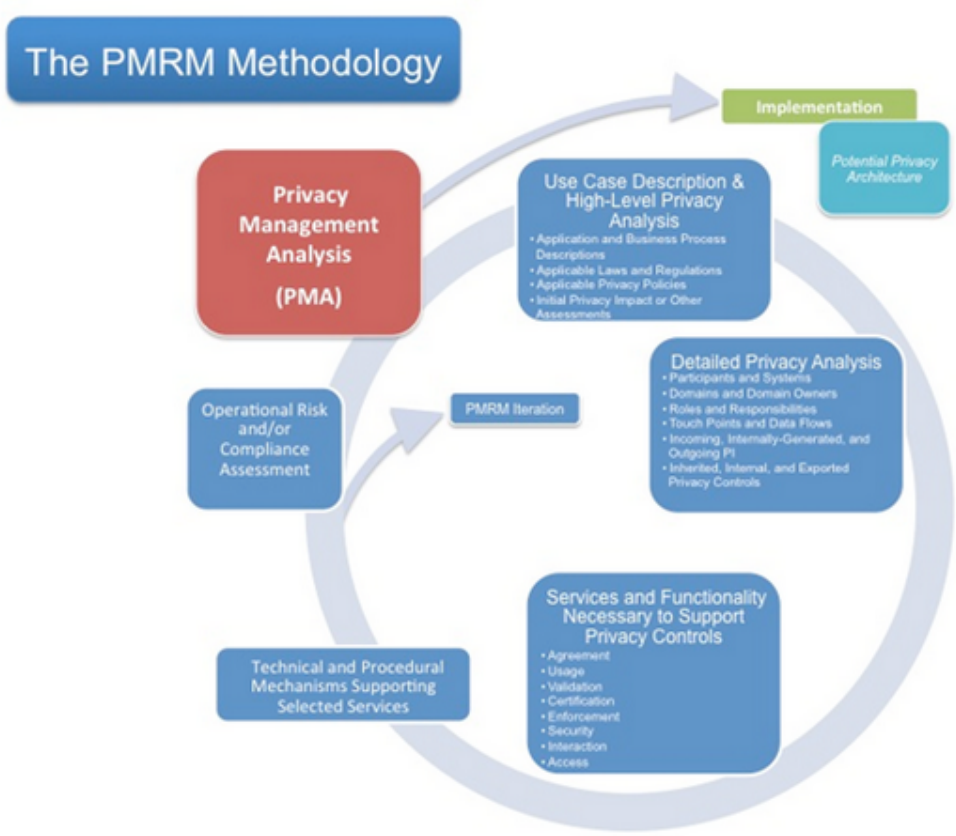


Figure 3.9 PMRM methodology [188]

In conclusion, it is worth noting that one of the most important advantages of PRPM is the ability to validate and control privacy services for both manual and automated mechanisms, which are necessary to implement any set of privacy control and compliance policies, including subsequent control of services and functions.

Discussion Although the presented PMRM has complete flexibility in the design and control of various privacy policies, the complexity of use or in other words the limitations of this model lies in the absence of a single established structure, thereby complicating the use of this RM for various stakeholders, including the interpretation of possible integrations. At the same time, the established relationships between vari-

ous elements and processes, including the presented set of functions, to some extent make it possible to optimize the previously mentioned shortcomings.

3.2.6 Other Contributions

Several studies also highlight the possibility and necessity of GDPR compliance considering data protection requirements in an organization. Among the following, we can consider various approaches that address these issues from the legal side. The authors [200] propose using a five-step approach that includes the following stages: gap and risk analysis, project planning, integration and following realization of national features and attributes. Approaches also include integration and monitoring of a) data protection management systems (DPMS) and b) RoPA. The complexity and difficulty of complying with the requirements of the GDPR leaves a wide field of possibilities for different approaches that can be presented by different consulting agencies, depending on the needs of companies.

3.3 Analysis Summary

The reviewed and analyzed studies have a broad scope of analyses considering the complexities of safeguarding personal data utilizing diverse approaches and methods, as well as the advantage of PET in the context of the area under study. **Table 3.2** supplies an overview of the approaches reviewed and highlights the design basis, evaluation and visual representation, as well as the purpose and contribution of each approach. **Table 3.3** presents and observes DPG in analyzed approaches.

Author (title), ref., year	**Design basis**	**Evaluation**	**Visual presentation**	**Purpose & Contribution**
PRIPARE, [39], 2015	Privacy best practices, privacy standards, academic publications	No evaluation, discussed in academia	yes	Integrating existing privacy best practices during the design process
Capability Model, [109], 2019	GDPR principles, guidelines, interviews with experts, stakeholders	Analytical evaluation, discussed in academia	yes	To help organizations implement GDPR requirements; as an approach that represents an assessment of privacy goals
Koç, [102], 2018	Synthesis set of actions, authors analysis and experience	No evaluation and no more discussed in academia	yes	Detailed set of actions for GDPR compliance (EA perspective)
Maturity Model, [103], 2009	Privacy standards, academic publications	No evaluation, discussed in academia	no	Phases of increasing quantitative or qualitative capability changes of a maturing element in order to assess its advances w.r.t defined focus areas

Continued on next page

Table 3.2 – continued from previous page

Author (title), ref., year	Foundation basis	Evaluation	Visual presentation	Purpose & Contribution
PRIAM, [43], 2016	Analysis of privacy risk guidelines and other privacy models	Analytical evaluation, discussed in academia, illustrated on the example	no	Conducting a privacy risk analysis which is both rigorous and systematic
SDM, [163], 2020	Observations of practices, interviews with experts, review of GDPR guidelines and best practices, gdpr principles	No evaluation, discussed in academia, industry	yes	Represents and provides the means to support, establish and evaluate TOMs to satisfy and demonstrate that personal data is processed by all GDPR requirements
ISO 27001	Information security standards, best practices	No evaluation	yes	Set of recommendations that allows organizations to achieve results in fulfilling the security & privacy requirements
COBIT, [67], 2019	Observations of practices, Information security standards, academic and industry publications	Analytical evaluation	yes, multiple versions	Framework to address issues related to security and privacy
IT4IT, [97], 2015	Industry publications, IT standards	Limited evaluation via industry examples	yes	An architectural approach, presenting pre-existing processes based on a value chain perspective
PMRM, [166], 2012	Observation of best practices, solutions, SOA principles, academic literature	Analytical evaluation	yes, multiple	Guidance for developing operational solutions to various privacy issues, including application to a variety of issues related to privacy management and compliance, valuable tool to achieve PbD

Table 3.2 Overview of the analyzed approaches

Various studies demonstrate the possibility of projecting and monitoring the GDPR requirement in multiple ways. The Capability Model from authors [109] addresses the GDPR requirements and objectives with a goals-based approach and can benefit from observing compliance considering the actual GDPR requirements. The PRIAM [43] outlines an approach to collecting and assessing privacy risks for various systems and data. MM for Data Protection [103] addresses data protection needs with comprehensive cri-

Author (title), ref.,year	Data Protection Goals						
	Data Minimization	Availability	Integrity	Confidentiality	Unlinkability	Transparency	Intervenability
PRIPARE, [34], 2015	P	P	X	X	P	X	
Capability Model, [99], 2019			P	P		P	X
Koç, [92], 2018		P	P				
Maturity Model, [93], 2009		P		P			
PRIAM, [38], 2016				P		P	
SDM, [150], 2020	X	X	X	X	X	X	X
ISO 27001, 2013	P	P	P	P	P	P	P
ISO 27701, 2019	X	P	X	X	P	X	P
COBIT, [60], 2019		P	P	P			
IT4IT, [87], 2015		P	P				
PMRM, [153], 2012	P	P	X		P		

Note: X - part of the approach, P - formally not a part of the approach but discussed;

Table 3.3 Data Protection Goals in the analyzed approaches

teria and measurements of compliance with GDPR requirements in a specific time interval. PMRN authors [166] outline an approach to accomplish privacy objectives utilizing PbD.

Among the analyzed approaches, various stakeholders accumulate added attention in the context of fulfilling or completing the GDPR norms. The method by Koç [102] introduced an approach for five different roles. PRIPARE [39] indicates the possibility of integrating this approach for the benefit of six different roles within the extended and detailed phases of the well-known and widely used design systems engineering. COBIT [67] includes more than 30 different roles in IT management and has documented functionality.

The GDPR requirements and the interactive nature of meeting/implementing the GDPR stated provisions are significantly traceable and are one of the essential elements in the SDM framework. Simultaneously, consistency and repeatability of GDPR compliance are included to a lesser extent in the PRIPARE project, as well as in the depicted approaches. The use of IT4IT can, to some extent, be adapted to the needs of personal data protection and related data protection measures due to the ability to link different levels of abstraction in IT. In addition, only two approaches [39, 163] consider PET utilization. Considering the research objectives "to support the implementation of data protection goals in PET DP utilization from GDPR compliance perspective by developing a GDPR blueprint reference model," we have identified various research gaps, which are presented below.

Some studies are presented in the context of position or white papers; simultaneously, not all approaches represent the methodologies used in the development of specific methods. Also, not all approaches are demonstrated and debated within the analyzed literature and references. In addition, most of the approaches reviewed are not present, accompanied by in-depth analytical evaluation, or not evaluated empirically. The visual representation and effectiveness of the considered methods are critical, as they are usually aimed at enhancing communication and understanding of complex abstractions and processes. Most of the approaches considered have a visual representation, with a few exceptions. To a large extent, the presented approaches are depicted as (complex) geometric objects (e.g. cubes, pyramids, circles or more complex). In many cases, approach concepts are illustrated utilizing blocks, and the relationships between blocks are represented using lines or arrows. Each concept or block presented in the approach visualization has a specific and distinct meaning, including a specialized application and purpose. In addition, it is worth noting that the experience and purpose of the developers of the presented methods can impact the method's quality and visualization. All approaches analyzed can be adapted or already developed considering the various levels of competence or experience of different stakeholders. In most of the presented approaches, countermeasures need to be included or clarified.

Also, many approaches require more clear step-by-step instructions (implementation strategy) and further detailed descriptions of interactions either between different levels (abstractions) or roles. However, these disadvantages in most approaches are minimized by presenting a various number of stages or phases. Only a few approaches utilize generally accepted implementation methodologies for performing step-by-step steps, namely PDCA (SDM) or SIPOC[2] (PRIPARE). Moreover, the PDCA implementation

[2]http://www.trialog.com/wp-content/uploads/2018/02/PRIPARE-Methodology-Handbook-Final-Feb-24-2016.pdf

strategy is the most widespread among the presented and analyzed. The different methods and accompanying privacy or data protection goals presented in studied approaches may be called by different names or interpreted differently. As the analysis indicates, only some of the methods satisfy, adequately represent or could be utilized in the data protection domain for PET DP. However, some approaches assessed with various extensions via various techniques can be adapted for stated earlier research goals and purposes.

3.4 Summary

Our overview recalls the main contributions suggested in academia and industry aimed at observation and fitting with the GDPR. Although some of the methods reviewed were presented before the GDPR came into force, various additions presented and released after the GDPR came into force have expanded the body of knowledge and data to be analyzed in the context of the current dissertation and research goals.

The analysis demonstrates that none of the presented and existing methods adequately represent the field of data protection in the context of the GDPR for PET DP and does not benefit research goals. This, therefore, requires a new revised conceptual model (in the context of the current dissertation, a reference model).

The next chapter will deliver, first, the development methodology and confirmation of choices and, further, the developed Reference Model for Differential Privacy (RMDP) model, which attempts to address some of the shortcomings identified earlier while also serving research objectives.

4 The RMDP Development and Base

This chapter presents the basis of the RMDP reference (meta)model. The RMDP development was carried out based on diverse sources in the domain of personal data protection and examination of existing privacy, compliance related approaches. RMDP is intended to overcome the previously demonstrated shortcomings and limitations in the analyzed approaches and meet current research objectives. Before moving on to the RMDP nomenclature, the development path for the above reference model is presented at the beginning of the chapter.

4.1 Development Methodology

The RMDP originated from various academic studies and industrial standards, as well as additional studies in the context of compliance and existing GDPR regulations, as well as the research objectives of this dissertation. As a result, the following sources were analyzed: 1) various data protection standards and practices in the field of personal data protection such as the ISO/IEC 27000 family, as well as the SDM framework published by the Office for Personal Data Protection in Germany (**Chapter 2, 3**); 2) various existing methods, approaches and frameworks for the studied domain (**Chapter 3**); 3) research results obtained during the joint research project "Differential Privacy: New approaches for managing social big data" [1] between BIDT and TUM , presented as research works at various conferences by the author of this dissertation [147, 148, 149, 150] and **Chapter 5, 6** results; 4) conducted surveys and the results [149] with various stakeholders involved in maintaining and monitoring GDPR requirements in the context of personal data processing, including discussions with various experts in the data protection domain, including various experts in the academia; 5) materials and additions from various data protection agencies and ICOs and 6) presentation of intermediate results and the concept of the developed reference model to various experts both among the academic and industrial domains; 7) various existing approaches and methods for the development of conceptual, reference, maturity and meta models and 8) frameworks for adapting privacy SDLC and engineering.

First, a careful selection and examination of various works devoted to protecting personal data in the context of the area under study was conducted. As a result, feasible definitions and interpretations were adapted for this dissertation, including the scope of the area under study, which was defined and supplemented, and possible dimensions were then outlined for the reference model being developed. In addition, during the examination of existing works, as well as results in the context of research studies delivered at international conferences. Interim results enabled to determine the importance/need for DPG (in the context of the studied PET), including feasible countermeasures in the context of the study area (**Sections 5.2 - 5.5**). Thus, this made it feasible to recognise and incorporate critical concepts in the personal data protection domain and data protection for PET DP in the formed intermediate reference model. In other words, as a result of the examination and analysis of research studies and examination of existing approaches to achieving the goals specified in **Sections 1.4** and **4.3**, the intermediate result of the developed RMDP reference model was introduced. Further adaptation of the research outcomes (including the results of publications [147, 148, 149, 150] and examination devoted to various dimensions of the developed model made it feasible to supplement and fill each block (dimension) of the formed RMDP reference model. The development of the RMDP concept was carried out in several phases; after each cycle, an analysis of shortcomings and comments was carried out, after which the concept was supplemented or, in other words, work was carried out to improve the developed concept (RMDP interim version), after which the

[1] https://www.bidt.digital/forschungsprojekt/differential-privacy-ein-neuer-ansatz-zum-umgang-mit-social-big-data/

relationships between the blocks (dimensions) (in this case depicted as arrows) were formulated (**Chapter 5**) and are depicted in **Figure 5.1**.

In the course of examination and fulfilling the specified goals, as well as organizing and systematizing available knowledge in the field of personal data protection, research was conducted using various approaches. The choice and utilization of a reference model that incorporates open gaps and enables us to arrange and systematize available knowledge in the field of personal data protection is given in **Sections 1.6** and **3.2**. This dissertation and the results and materials presented in this dissertation followed the selected methodology and recommendations (**Section 1.7**). The reference model being developed, named RMDP, is the result of an examination of various types of materials and academic studies, representing a synthesis of existing knowledge. The BES approach [181] was utilized during the study. This approach has significant advantages (**Section 1.7**), as it provides the necessary recommendations for synthesizing existing knowledge and enables us to maintain the specified research vector when selecting knowledge and subsequently including it in the synthesis of knowledge. This approach also suggests relying on quantitative characteristics of studies (if available) and including them in the analysis and review. Although this research and the sources presented in this study, in most cases, do not provide quantitative data. Despite this, the BES approach proposes determining a priori criteria for the inclusion/exclusion of (existing or significant) knowledge in subsequent synthesis. During the course of this dissertation, the following non-quantitative criteria [181] were included in the developed concept and reference model of RMDP (we also validate and comment on the presented criteria in parentheses) :

1. The concept(s) must be presented either in the academic domain or delivered in sources dedicated to industrial standards; simultaneously, the concept must be known in the domain under study (in order to be included in the reference model being developed, this concept must be significant for the domain under study, at the same time presented and recognized in scientific publications - "*germaneness to the issue at hand*" [181]);

2. The concept must be easily understandable and does not require significant technical, legal or privacy background or knowledge (the presented criterion is based on and predetermined in the context of the initial requirements when developing an RMDP, which is intended for certain categories of stakeholders or experts who do not necessarily have technical, legal or privacy competencies at the same time);

3. The concept must be universal, meaningful and valid for many or applied by most organizations (the RMDP reference model is an abstract model that, with the necessary and minor adaptation (or modernization), can be used by multiple organizations);

4. The concept must be essential and have added meaning for various organizations (personal data protection issues and the area of data privacy and GDPR do not exist only by choice and due to a wide area of research problems. For many organizations, issues of personal data protection and GDPR compliance are among the criteria for achieving business objectives and goals. In other words, RMDP can be distributed not only among the previously mentioned stakeholders but can also be expanded to business experts. Thus, the developed RMDP must convey thoughts and be relevant, including being interpreted in a publicly accessible language and accessible to various people: business experts, legal experts developing and updating various regulations (for example, GDPR, ISO privacy standards, EU AI Act, also articulated in **Chapter 7**) who, in the context of interpretation and RMDP utilization, have the opportunity to conclude strategic decisions (taking into account the specifics of various organizations (paragraph 3 presented above));

Thus, the choice of the abovementioned criteria is explained by the purpose and preference of specific stakeholder (roles) categories in the context of the RMDP and the dissertation objectives. As has already been demonstrated in previous chapters, DPGs play a significant role in monitoring and compliance; simultaneously, referred DPGs are directly explained in the GDPR. In other words, DPGs play an essential role in the context of privacy and compliance assessment (in the context of GDPR, it can also be applied

to different levels of anonymization [149]). Simultaneously, DPGs can be communicated between various stakeholders and interpreted without distorting understanding. Therefore, RMDP integrates DPG as one of the areas (dimensions) in the developed reference model. Although there is no specific decision in the scientific literature (academia) about the used or alternative DPGs, in the context of this work and research tasks, we concentrated on the DPGs already presented in the GDPR; in the course of added work, we adapted the results [149] which were also explored for PET DP. In other words, to determine and select DPGs, the following path was carried:

- the presented DPG in the GDPR (**Section 2.5**) and the analysis in the context of the various methods **Table 3.2** was carried out;
- each DPG was analyzed w.r.t DP [148, 149];
- DPG was included in the interim version of the reference model based on the following facts:
  - each goal has a unique name;
  - each goal has a unique definition and does not overlap with other goals;
  - each goal has a unique meaning and is significant for the data protection domain, including technical and legal areas, at the same time significant for the business context (including for various types of organizations) and can be interpreted without distorting the meaning;
  - each goal is presented in both academic and industrial sources, at the same time, each goal is discussed by various sources in the academia;
- during our survey [149], not only experts pointed out the need and opportunity to use goals data, including various privacy researchers pointed out the possibility and relevance of DPG data (although for researchers the interpretation and priority of DPG (in the long term) differs from other stakeholders);

Simultaneously, considering [37], in order for qualitative research to satisfy the requirements of rigor and credibility, the process of research and investigation of gaps a) must involve various people, not just researchers, and b) in addition, research results must be evaluated by people other than the researchers. During the RMDP development and evolution, including the interim version (**Figure B.3**), the RMDP was iteratively discussed with several practitioners, experts, researchers, and graduate students. As a result, the developed reference model was adjusted according to the comments and additions in order to enhance quality, accuracy and reliability. In addition, after concluding the design and refinement, the RMDP reference model was evaluated through interviews conducted with various industry experts, junior and senior researchers and partly doctoral, master's students (**Chapter 6** is devoted to RMDP evolution).

4.2 Reference Model

Considering the discussion and interpretations of various (conceptual and reference) models presented in (**Section 1.8**), any area of knowledge can be described using an explicit or implied conceptual model that describes the phenomenon under study. Simultaneously, a reference model can be characterized by a conceptual model if it attempts to represent the problem under study at the industry level and reflect knowledge about the domain or area under study [130]. Thus, we argue that the reference model being developed (in our case, RMDP) can clarify the phenomenon under study or "systematize knowledge, map reality and guide research" ("systematization of knowledge, guiding research, mapping an area of reality" [94]). According to the authors [54], the reference model has three main characteristics:

1. represents best practices (**best practices**),
2. has universality and may not be limited to a single or isolated case (**universal applicable**) and

3. reusable, that is, it can be used many times (**reusable**) (due to the solution template);

Furthermore, a reference model is a) the result of a design process in which the developer of the reference model defines critical elements and concepts [172] and b) is both descriptive and prescriptive due to the possibility of describing a subject area or provide design guidance [60]. Accordingly, when utilizing a reference model for a specific case (task), an organization needs to adapt the referred model (template) in the context of the required explicit tasks [55]. Since when developing a template or reference model, the capability to deliver for all potential requirements and application scenarios, possible mechanisms for adapting the reference model (including for reuse and recycling) are critically essential [13]. According to the authors [24] to accomplish two main goals (*universal applicability* and *reusability*), the following five adaptation mechanisms can be utilized (to adapt elements reference model for a situation-specific model (in other words, for a specific situation or task)):

- analogy (using a reference model as a guide for building a similar model);
- specialization (revision of the reference model and expansion of its elements);
- aggregation (combining various elements of the reference model);
- instantiation (specification of common elements of the reference model);
- configuration (selecting elements from many alternative options);

Thus, configured reference models can have mechanisms and be customized depending on the context, organization size, industry or other various configuration parameters (for example, different stakeholders, purposes of use) [161, 13]. A conceptual model (like a reference model) should contain a description of a) inherent (system) elements, b) relationships between elements, c) changes in elements and their relationships, and d) possible directions of research [94]. A conceptual model is generally visually depicted [56], and typically, the primary purpose of a conceptual model is to facilitate understanding and communication among stakeholders in the domain or system being modelled and studied [130].

The reference model represents and describes knowledge at a conceptual level and thereby enables and is intended to convey knowledge in an understandable and interpretable format for people. In the context of this dissertation and the goals of this study, it is to provide knowledge of the protection of personal information in the context of the disposal of PET DP and DPG in a form accessible to various stakeholders and a broad audience. Therefore, the developed reference model corresponds to and promotes practical mutual understanding and communication between various stakeholders and individuals actively involved in compliance with the GDPR requirements. The advantages and justification for choosing a reference model and its subsequent development, including the inclusion of knowledge in the developed reference model, is that the specified model does not contain special notations and is simplified for understanding by various individuals (stakeholders) due to neutral representations. Due to the lack of formal rules of what can be represented and what cannot, the reference model allows it to remain adaptable. In addition to eliminating the need for users to have significant knowledge and expertise in the field of privacy, it (reference model) enables expanding the user's audience and application domains.

In other words, we argue that utilizing the adaptable quality of the reference model enhances interpretations of non-specific and goals-based GDRP requirements. In addition, it reduces the complexity of interpretations of various DP forms and utility and privacy implications by broad applicability. On the other hand, we argue that the selected adaptability and abstraction level of the reference model developed neglects the complexity of a) implementation of various DP forms (to some degree raised and debated in authors analyses [148, 149], b) DP utility-privacy tradeoff for diverse use-cases and c) privacy taxonomy implications [150]. Additionally, previously reviewed and analyzed various models or approaches (**Sections 3.2** and **3.3**) ensure that the use of a reference model free from special notations is preferable. Thus, this is one of the factors influencing the development of the reference model in this dissertation, accumulating knowledge for compliance with GDPR within the framework of the use of DP and the previously discussed DPG and further presented RMDP (**Chapter 5**).

4.3 Reference Model Requirements

The requirements for the model being developed must align with the goals and purpose of the model. Additionally, these requirements are crucial components in the development of the specified reference model. In this section, we will explore and demonstrate the requirements that have guided us in the context of developing and refining the RMDP reference model, drawing partly from the results obtained:

- research studies from author of this dissertation [147, 148, 149], joint work where the author of this dissertation is the first author [150];
- generally accepted requirements for reference model design [54, 172, 55, 24];
- critical analysis of similar approaches and frameworks (**Chapter 3**);
- requirements PbD and DPbD (**Section 2.4.1**);

We have also identified separate requirements subsections (**4.3.1-4.3.3**) for a more transparent modeling process and design of the RMDP reference model, presented below.

4.3.1 General Requirements for Reference Model

As introduced earlier (**Sections 4.1** and **4.2**), a conceptual or reference model describes and strives to clearly represent the area under study and the accumulated knowledge in research domain, in part by combining various aspects into a single model. The transparency, clarity and intuitiveness of the reference model being developed must be clear to the identified stakeholders or groups who intend to utilize the reference model. As stated earlier, this model is being developed and aimed at stakeholders (various roles) who are experiencing compliance with the GDPR requirements and considering PET DP utilization. Simultaneously, visual presentation and simplified expression (excluding special notations) are some of the RMDP preferred attributes. Thus, we argue that the RMDP model will be more accessible to the understanding of numerous stakeholders and groups (roles). Further, we will demonstrate the main requirements starting with the following:

R1: Adequate and clear visualization and (step-by-step and consistent) compliance with GDPR requirements.

Interpreting and implementing the GDPR requirements, as well as minding various requirements and comments from various privacy organizations, is a labour-intensive and not strictly regulated process. Simultaneously, the ability to interpret and interconnect different concepts, processes or the overall privacy cycle must be adapted through the following requirement:

R2: The ability to trace the relationship and phasing between various stages/dimensions and/or concepts.

As we explained earlier (**Section 4.1**), the model should be applied to diverse types of organizations facing similar challenges and requirements in the context of GDPR compliance. Simultaneously, as we discussed (**Section 4.2**), the reference model must maintain the necessary elasticity to meet privacy organizations' various requirements and objectives. In other words, the model being developed must be able to synchronize with the tasks and requirements of various organizations in the context of various requirements (declared in GDPR) for personal data. Therefore, in accordance with the primary purpose of the reference model and the objectives of this work, we present the following requirement:

R3: Ability to adapt the model depending on the context, needs, objectives and requirements of organizations.

Considering imposed high-level requirements by GDPR, as well as based on the analysis of various similar models and approaches (**Chapter 3**), and including in line with the previous (**R3**) requirement, various dimensions (in the context of RMDP - four dimensions are available) or individual elements may have the possibility of progressive and added value when executing the stages of the developed model. Based on this, and also take into account that for different dimensions, the GDPR requirements may vary or have the same meaning, also taking into account the probability of a changing (or dynamic) DP privacy concept [150]. We argue that this demands the following requirement (presented below). In other words, following **R4** recommendation, the use of RMDP has significant flexibility in the context of the complexity of the presented requirements for the models.

R4: Possibility of progressive and repetitive execution.

4.3.2 Requirements to enhance GDPR compliance

Despite the requirement for the development of reference models [172, 55, 24] and the requirements for the reference model being developed (**R1-4**), one of the main tasks is to concentrate on the task of complying with the requirements of the GDPR in addition to **R1**. In other words, the elements and concepts used and integrated must have added value and be suitable for the matter under study, including having the required level of detail or otherwise be supplemented (in the context of the required level of detail) by other concepts or elements used in a given reference model. In addition to the above, the focus should also be on PET DP.

R5: The possibility of intersection with the GDPR requirements the process of integrating them in the context of PET DP utilization.

The developed (reference) model must be practical and combine both regulatory requirements and practical knowledge, including various recommendations from various projects and industrial recommendations and a wide range of academic results. Through the flexibility presented above and the adaptation of results from industry or academia, this will facilitate the reuse of various successful practices.

R6: Presentation of the possibility and practical information for fulfilling or following individual requirements or elements of individual parts (such as Article, Recital and others) of the GDPR.

During the examination of the GDPR requirements and various approaches as PbD / DPbD as well as various methods (**Chapter 2** and **3**), including additional examination among various stakeholders [149] as well as tolerating the need to conduct ongoing analysis and risk assessments for data subjects.

R7: Promote and encourage the reuse of existing concepts and elements.

One of the last requirements that we want to reasonably include in the developed model is the ability to apply this model to different stakeholders and roles while simultaneously operating with the necessary concepts and elements included in this model. This will minimize or reduce the dependence on previously identified problems when adapting the GDPR and its use and, most importantly, interpreting the results obtained for various stakeholders. In other words, this will bypass the loss or interpretation of various requirements received during diverse activities in the context of personal data protection between diverse groups of stakeholders. Simultaneously, it will ensure a more dynamic and logical flow, including exchanging opinions on the common task of GDPR compliance and minimizing privacy risks for data subjects. In other words, it is expressed to optimize the use of the various stakeholders involved in compliance with the GDPR requirements while at the same time disclosing the complexity of the requirements on the part of the data controller and presenting a clear picture and analytical tools to continue to implement ongoing compliance with the GDPR requirements and ensuring the highest level of competence while operating with personal data.

R8: Take into account the importance and significance of various stakeholders and interested parties in the context of GDPR compliance, thereby ensuring the ability to interpret the results and support compliance relationships between the various stakeholders directly involved in the continuous data protection and compliance process.

The following section provides a more detailed outlook on the metamodel of the developing RMDP reference model, followed by a section on additional parameters such as modelling language and visualization.

4.3.3 Metamodel

Generally, the reference model, function, and goals could be introduced on more detailed levels with the service of so-called metamodels while designing the reference models (**Figure 4.1**). In addition, considering various research goals or objectives, the reference model could be presented with diverse levels of detail supplemented with a complete description of the model due to the specific case or organization needs. Thus, modelling the reference model is complex and depends on the organization task or purpose of the reference model (we recall our critical analysis **Chapter 3**, which highlights the diversity of methods and approaches aimed to fit comparable data protection goals); however, considering the previously described and discussed requirements for the developed reference model (**Section 4.3**), we argue that for the research goal specified (**Section 1.5**), we utilize only specific elements from the generic way for developing a metamodel for RMDP.

Further, we explain our reasoning and logic for the metamodel development in terms of RMDP, including the choice of metamodel notation. We also want to point out that the decision to choose the further presented level of abstraction and details for metamodel development was raised through discussed implications related to the a) various DP forms, b) not explicitly defined GDPR notations and c) privacy-utility complexity issues for various use cases. In other words, we leave the choice to complete the detailed metamodel for organizations and actors since it enhances the adaptation of RMDP and a specific metamodel specifically for use cases and requires EMTs imposed by diverse organizations. On the other hand, such an approach delivers organizations with the flexibility (Requirement 3, **Section 4.3.1**) and does not limit the adaptation of reference models and only extends the universal applicable characteristics (**Section 4.2**).

Models and Metamodels are an integral part and are used to represent various systems. For the reference model being developed, the metamodel presented below describes RMDP. At the same time, according to OMG recommendations [2], the metamodel corresponds to one of four levels, namely M2 (**Figure 4.1**) or, in other words, levels of abstraction that can be applied to various OMG standards [79].

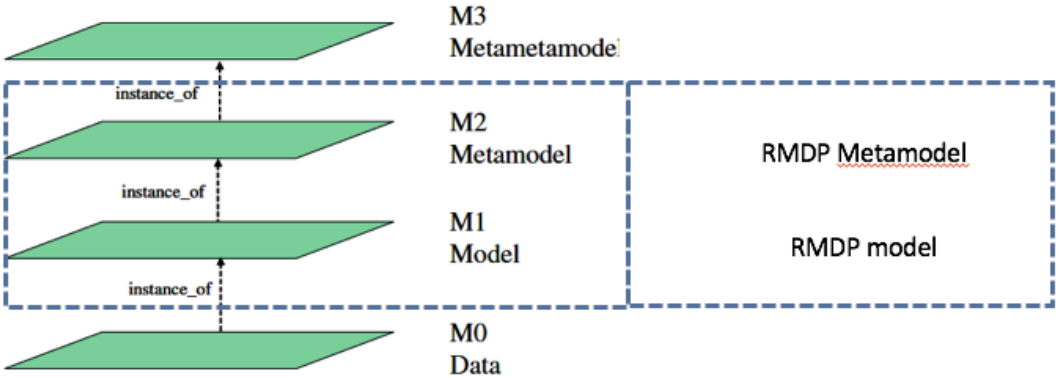


Figure 4.1 RMDP interpretation considering four levels of abstraction in various OMG standards [79]

.

According to [20], "the model should be well-formed according to specific well-formedness criteria" and "a model comes with its background, e.g., with paradigms, assumptions, postulates, language, thought

[2]https://www.omg.org/spec/SPEM/2.0/PDF

community, etc.". Thus, in agreement with the author [20], we address the following steps: 1) classification of (information) objects in the model and description of the master or metamodel and 2) selection of a suitable and representative modelling language. The metamodel presentation and logic overview (**Figure 4.2**), including the interpretation of the metamodel among different levels (**Figure 4.3**), creates a common understanding among experts from various expertise domains and deliver the foundation for the reference model being developed [83, 173]. In other words, a metamodel (or master reference model) represents the possibility of a blueprint for constructing a reference model [173] including this metamodel allows you to define elementary standard components (concepts) that can be used to build the reference model being developed. Additionally, a reference model (or metamodel) is a model that is used to describe and represent a reference model. In addition to the design and construction domain, RMDP, including the connections or relationships between designs/concepts, is defined using a metamodel.

In the process of modelling the metamodel, the author of this dissertation followed the recommendations of approach [79] and utilized the notation from study [123]. Commonly, the metamodel capabilities are not limited only to the reference model development; in some cases, a metamodel can also be utilized to supplement and develop a reference process model or adapted to fit the needs of an existing reference process model (however, this direction and the development of a corresponding artifact is not the goal of current dissertation). As the authors point out [167], such a need or opportunity satisfies one of the active areas of development of various information systems in the context of privacy engineering.

We also argue that organizations have the freedom to modify and adjust various notations for the metamodel depending on the organization's objectives and dynamic factors. We also want to note that the selected notation [123] could be modified or extended with details to fulfil the organization's goals and serve as a sample. We also want to draw attention to the fact that in adapting various notations (even those unrelated to privacy), organizations have the freedom to utilize notation at their own discretion, which is not mandatory.

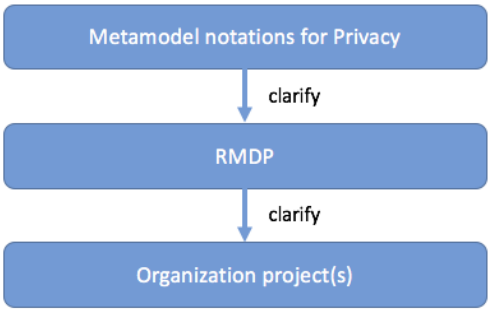


Figure 4.2 RMDP metamodel representation/logic overview considering privacy notations [123]

As previously explained, according to the authors [79], metamodels can be represented in different ways. In this work, we focused our attention and followed the recommendations of the OMG (SPEM 2.0) approach [140], namely, we followed a separate approach developed according to OMG recommendations specifically for privacy engineering (including using the notations specified by the author [123]). According to [123], the presented metamodel is an addition to the existing SEMDM [70], a metamodel for software and systems development presented in ISO/IEC 24744 and "provides a controlled vocabulary of privacy engineering methodological elements and a normalized set of connection points and relationships to organize those elements" [4]. In addition, the metamodel capabilities are not limited only to the reference model development; in some cases, a metamodel can also be utilized to supplement and develop a reference process model or adapted to fit the needs of an existing reference process model (however, such direction and the development of a corresponding artifact is not the goal of current dissertation). As the authors point out [167], such a need or opportunity satisfies one of the active areas of development of various IS in the context of privacy engineering.

Minding the given nomenclature and description of the elements [123], taking into account the previously presented requirements for the reference model (**Section 4.3.1**), we presented the first version of the metamodel (concepts and blocks) (**Figure 4.3**) which will be utilized to RMDP development. In other words, the presented structure is designed to ensure flexibility and extensibility. Decomposing various concepts or adding more diverse connections and hierarchies between roles/concepts will enable and ensure the built-in flexibility already at the metamodel phase. Thus, the chosen metamodel notation [123] and further adopted definitions not only provide a solid foundation for developing a metamodel based on existing software engineering metamodels for development methods (ISO/IEC 24744, also known as Software Engineering Metamodel for Development Methodologies (SEMDM) and also enable "a global perspective to define any method through integration of all methodological aspects related to the followed procedures" [77]) but also play a crucial role in enhancing privacy engineering approaches.

- **Definition: Role(s)** a collection of responsibilities that a producer can take [123];

According to [123], a producer is defined by the agent(s) that have the necessary qualities and describes the agents who are responsible for executing work units. Theoretically, this definition can be applied to not only system objects, or in the context of this dissertation we mean a person.

- **Definition: Work Unit(s)** describe things to be executed, such as a Process; Task(s): a small-grained work unit that focuses on what must be done [123];

In agreement with the author [123] a work unit can be divided into different classes of work that must be performed, including the ability to define additional recommendations (for example, mandatory, optional, possible).

- **Definition: Task(s)** a small-grained work unit that focuses on what must be done [123];

- **Definition: Recommendation** may recommend using specific techniques for a task (e.g. compulsory, optional, discouraged, etc.) [123];

Recommendations can also apply to or complement the definition of work units, namely the subcategory of work unit (Task).

- **Definition: Work Product(s)** are artefacts of interest for the project which can be used as inputs, intermediate results, or outputs of a work unit [123];

In other words, Work product(s) is an intermediate or final result obtained during the project, which can represent a model, document, etc.

- **Definition: Resource(s)** that is, methodology elements that are used ‘as is’ at the project level, without requiring any instantiation [123];

- **Definition: Stage(s)** subterm, which represent a managed time frame with duration such as a Phase [123];

Each of the presented elements/notations (Roles, Work Units, Work Products, Tasks, Recommendations, Resources, and Stages) can be interpreted in two ways [123]. On the one hand, these are instances of concepts defined in the metamodel; on the other hand, these are templates that will be created every time you follow or start. As presented in **Figure 4.4**, we utilized the three main concepts of Producer, Work Unit and Work Product (including their related concepts) in accordance with the notations and approaches indicated earlier. To depict the metamodel, we used a simple format to represent fragments and concepts, while a simplified format is not excluded for depiction in the context of ISO/IEC 24744. In other words, to define the RMDP metamodel, we used the definitions [123], including expanding the concept of Producer

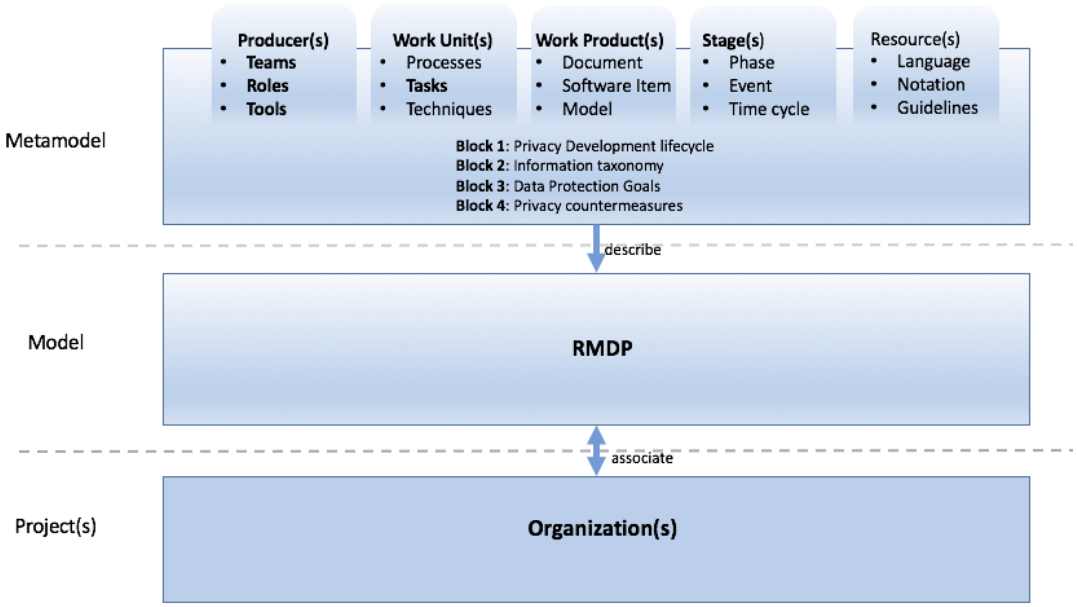


Figure 4.3 RMDP metamodel model representation in levels abstraction considering the privacy context
Note: RMDP metamodel representation/logic overview considering privacy notations; we present all categories (Producer, Work Unit, Work Product, Stage and Resources) for considered notations and we also highlight utilized and adopted notations particular for RMDP in bold.

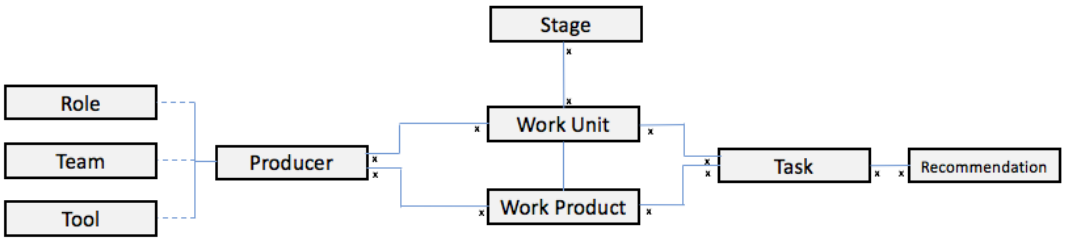


Figure 4.4 RMDP metamodel (simplified visualization)

and replacing the concept of Person with Stakeholder, which will enable us to focus on the issues of interpretation and execution of tasks and requirements presented earlier as part of the development of the RMDP reference model.

A separate essential aspect is the visualization of the reference model, including the selection of the accurate modelling language and structure of the model being developed [173]. The following facts and considerations drove the decision to present the developing RMDP reference model as a clear visual representation. Individuals are more receptive and better at remembering visually presented information than text, and at the same time, visualization increases interest in reading [72, 114]. RMDP is presented in a notation-free form, although some elements of the "knowledge representation" are like a mind map [205] can be displayed and traced (the mind-map approach is one of the recognized techniques for organizing knowledge). We also argue that using different colours in the reference model being developed improves understanding of the model's structure. At the same time, we assert that presenting the model in black and white does not distort the understanding or meaning of the model. A selection and explanation of the various colours are presented in **Section 5.8.** We also settled on choosing a written language as the most straightforward method (accessible to most individuals) and a universal modelling language.

4.4 Summary

This chapter meticulously presents the methodology and path for developing the reference model, providing a comprehensive understanding of the background for RMDP development. It includes the comment and assurance of the concepts intended in the developed RMDP model, ensuring a thorough and reliable process.

Particularly, we demonstrate in **Section 4.1** the way and through materials followed and accompanied the development process. This also highlights the implications regarding the various DP forms and implications in analyzing the GDPR requirements, which will be discussed and explained in **Chapter 5** in more detail. We also demonstrate and discuss the criticality of the integrated concepts and demonstrate the importance of the DPG in the reference model. **Section 4.2** is devoted to the basic requirements and characteristics of the reference model (applied for RMDP), which impacts the initial steps of developing the model reference model. It also presents the required description of integrated concepts and knowledge adaptation. Importantly, the discussed approach does not limit the developed reference models, ensuring their adaptability to various DP forms and utility-privacy issues for different use cases. Requirements, namely general requirements for the reference model and requirements to enhance GDPR compliance, are demonstrated in **Section 4.3**. This also supplied the development of metamodel, considering the following characteristics as best practices, universally applicable and reusable, and presented on a semi-abstract level, which satisfies various goals and tasks for diverse organizations. The presented metamodel utilizes the adaptation of privacy notations and serves as a template for further adjustment.

In conclusion, the RMDP metamodel and the elements and notations used to produce the first version of the RMDP metamodel were presented. The adapted concepts and elements, including the metamodel development logic, satisfy various recognized standards in academia and industry utilized in the context of privacy and personal data protection. The modelling approach, as well as the visualization and modelling language, are presented in a notation-free form to support understanding among various stakeholders, including elements of the mind map visualization that may be present in the reference model; this visualization technique is one of the most recognized in knowledge organization.

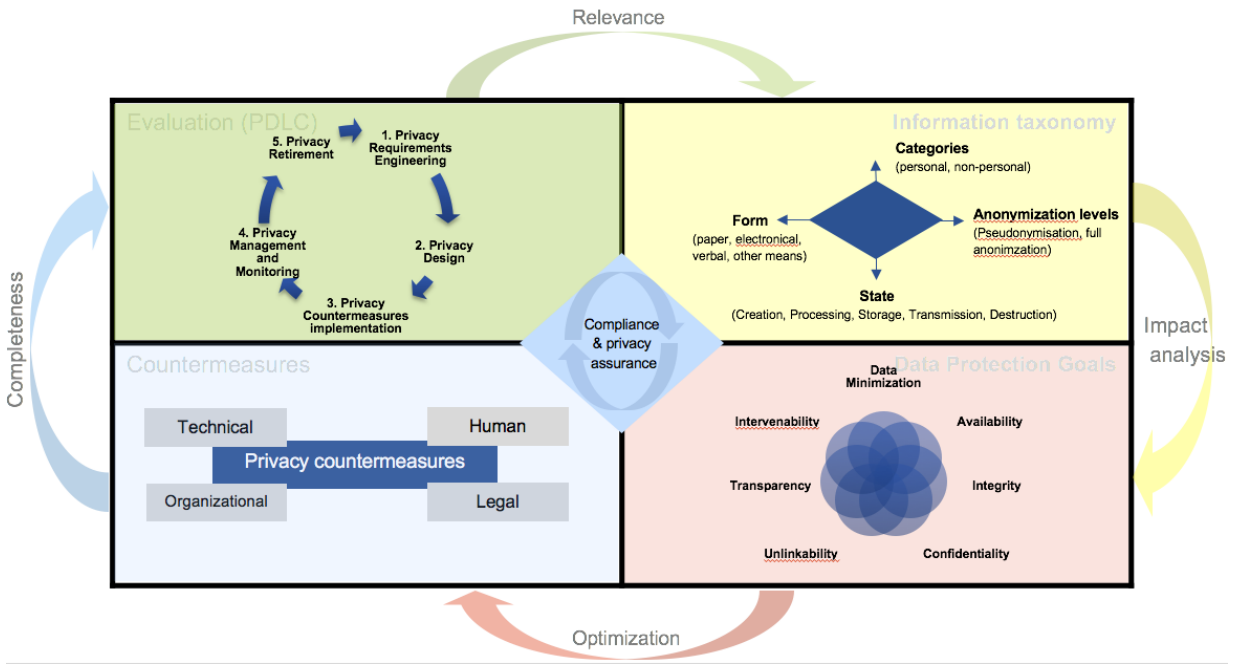
5 The RMDP

In this chapter, we will focus on developing the primary artefact of this dissertation. First, we will start by providing an overview of the reference model being developed (**Section 5.1**). Then, we will move on to a more careful consideration of each dimension of the developed model (**Section 5.2-5.5**). Further, we present the interaction between dimensions (**Section 5.6**) and the RMDP utilization (**Section 5.7**), followed by an introduction of RMDP visual interpretation (**Section 5.8**). Finally, we discuss the RMDP in **Section 5.9** with the following summary in **Section 5.10**.

5.1 Overview of RMDP

The developed RMDP reference model, as depicted in **Figure 5.1**, is illustrated using four dimensions:

- **Evaluation Dimension** (top left) demonstrates privacy progress according to the Privacy Development Lifecycle (PDLC);
- **Information taxonomy Dimension** (top right) emphasizes the characteristics of information assets that must comply with or be taken into account in the context of GDPR requirements;
- **Data Protection Goals Dimension** (bottom right) displays seven data protection goals extracted from those offered in the SDM framework and GDPR;
- **Countermeasures Dimension** (bottom left) depicts various countermeasures categories;



Note: The use of the model (enhanced with a specific colour scheme in Section 5.8) starts from the top left dimension (Evaluation), mainly explained in Section 5.6, including an explanation of arrows outside the model.

Figure 5.1 Reference Model RMDP

Four dimensions[1], which are the blocks of the developed RMDP reference model (Evaluation, Information Taxonomy, Data Protection Goals and Countermeasures), were incorporated in the model because they fulfil the criteria explained in **Section 4.1**. The suggested dimensions come from earlier results of various models/approaches that were examined; current results were also adapted to the necessities of this dissertation and demonstrated further. The developed RMDP model generalizes the studied models/approaches discussed earlier (**Chapter 3**) and at the same time complements the approaches SDM [163] and PRIAM [43]. The origin of each dimension, including how each block complements the RMDP reference model, will be presented in the following sections (**Section 5.2-5.5**) corresponding to each dimension. Additionally, the model delivered with a specific metaknowledge addressed from the determined color scheme (green, yellow, red and blue color), which enables a more satisfactory performance of the two-dimensional structure of the model and further execution, explained in **Section 5.8**. An examination of academic studies and industry standards (**Chapters 2 - 4**) implies and indicates that each concept included in the RMDP has been previously presented and is known to a broad range of experts for various critiques and suggestions. The mentioned studies and recommendations are discussed and explained in the following sections corresponding to each dimension, including a discussion of the importance of each concept in the focus of compliance and data protection and PET DP tasks. As we present and articulate each RMDP dimension further, it is important to comment that the four formed and shown RMDP dimensions do not intersect, overlap, or duplicate each other, or in other words, are independent. The use of RMDP starts from the upper left quadrant and follows the following presented steps: 1) what kind of (personal) data should meet compliance standards (Information taxonomy) 2) which/what data protection goals should be applied or reached (Data protection goals) 3) how DPG can be accomplished (Countermeasures) 4) and a cycle of privacy activities that ensures that the presented three blocks (Information taxonomy, Data protection goals and Countermeasures) can be achieved at each stage of the PDLC cycle.

Further, our comparison and examination demonstrate that in addition to the reference model RMDP being developed, only two models, SDM [163] and PRIAM [43], partially incorporate and adapt comparable blocks, the remaining approaches or models to some extent overlook some essential concepts of data protection, privacy and utilization of PET (DP). However, such interpretations could depend on the model designer's goals and purposes of previously analyzed approaches.

The differences between RMDP and SDM [163] and PRIAM [43], with which the developed reference model RMDP has some similarities, lie in the internal structure of dimensions, as well as the various concepts utilized: 1) RMDP represents more broad opportunities for assessing data protection goals (according to **Table 3.2**), 2) RMDP expands the capabilities of assessing compliance due to an additional data protection goal (Data Minimization) not explicitly specified in the GDPR, 3) RMDP represents a more expanded taxonomy for private information, 4) SDM [163], PRIAM [43], unlike RMDP, do not use decomposition and additional analysis through various PDLC phases. Additional comparisons of the developed RMDP model will be provided subsequently in this chapter after describing and explaining each of the four RMDP dimensions. Considering various research studies and our analysis, it was hypothesized that the previously mentioned RMDP measurements could be mandatory and sufficient for the data protection and compliance domain and are easily accessible and understandable to the target audience (with different expertise levels) for the selected level of abstraction. These assumptions were validated and argued in various ways through further interviews with various experts, stakeholders and researchers (also with (Master's) students) and offered in **Chapter 6**.

The RMDP reference model is a general abstraction, so when utilized in specific organizations, individual elements of RMDP can be adapted according to the organization's recommendations/requirements, namely:

- the general abstraction PDLC (Evaluation Dimension) can be supplemented or modified to suit the needs of the organization (**Section 5.2)**;

[1] synonyms of the dimensions can be used with the support of the auxiliary terms quadrant or block, which is a geometric description of one of the four dimensions of the developed reference model

- the Information Taxonomy Dimension can be supplemented or modified taking into account the various requirements/needs of the organization;
- DPGs Dimension can be supplemented or modified (priority is determined by considering the needs of the organization/performers;

Subsequent **Sections 5.2-5.5** describe each dimension (four dimensions in total for the RMDP reference model), including emphasizing the importance of the demonstrated dimension(s) for data protection and compliance domain, as well as the motivation for inclusion in RMDP model. **Sections 5.6**, **5.7** and **5.8** describe, explain, and comment on the relationships between dimensions and the graphical design of the RMDP reference model, respectively.

5.2 Dimension 1 (Evaluation)

In agreement with the SDM framework [163] and PRIAM [43], considering recommendations of various privacy agencies and the PbD principles (**Section 2.3.1**), the importance of time and reiterated examination of privacy risks for Evaluation block is irreversible. In other words, the importance of time or the possibility of repeated (iterative) use requires additional comments and explanations. This is also stated in Article 32(1)(d) GDPR: "*a procedure for **regularly** testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing*."

To begin with, the concept of "time" or "iterations" can be viewed from diverse perspectives. The primary idea can be seen as "time" (or iteration) or as a "time continuum" in the data protection domain and compliance matters following a) GDPR or other regulatory necessities or b) the current interpretation of the requirements for different models/approaches developed to meet the necessities. Considering previously examined models, approaches (**Chapter 3**), we can conclude that any model or approach depicts a snapshot of an area at a certain period of time and emphasizes or reflects its current environment or elements (including requirements). In other words, we can express that any model or approach contains a) a variety of existing privacy countermeasures, b) knowledge of privacy and data protection acquired at the moment or period of time in question, c) various interpretations and coherence of concepts, threats and techniques that can be associated among themselves, d) privacy expectations and awareness, e) current standards, requirements and best practices for the data protection and privacy domain, f) the level of expertise of various experts and model developers in privacy domain. This means each model or approach has a time limit (or expiration date); or, put another way, no conceptual or reference model (including even the RMDP) can be disposed of or operated forever (this is also implied to the concepts utilized or knowledge already applied in various models and approaches). In addition, nowadays, we can observe the constant evolution of various models/approaches, as well as various legislative acts in refered domain. Therefore, any changes in the data protection and privacy domain are reflected in conceptual models' evolution, revision or expansion. To this end, in the context of the development of conceptual/reference models, understanding "time" or the "time continuum" is essential for the expert or model designer, as it contributes to a) understanding the domain of privacy and provides possible means for interpreting privacy requirements and b) emphasizes the need for regular revision of the model and prevents the model from becoming entrenched or falling behind the requirements in the area under debate. Simultaneously, interpreting "time" in the first matter does not play the most significant role in practical application in model development.

An alternative view of the concept of time in the area under study and the reference model being developed can be presented as the life cycle of the development and use of a reference model in a specific organization or task. This perspective, on the one hand, is pragmatic and more consistent with the context and objectives under discussion (dissertation goals). In other words, time in the context of the presented task can be delivered in the context of the privacy lifecycle in the context of the established concept of SDLC. This alternative view is not just a theoretical concept but a practical and applicable approach. Simultaneously, pragmatism and definitions of time are most suitable for SDLC (and subsequent IS). In addition, an alternative view of the time concept does not contradict the first presented interpretation of time and, to some extent, can be explained as an addition with certain conditions.

The importance of integration and considering privacy issues is appropriate from the very early stages of the development of various systems [29]; simultaneously, it is also suitable and documented by analysis for maturity models [101]. In addition, privacy and security issues (which are not only one piece of various systems) can be extrapolated both vertically and horizontally, thereby indicating the broad impact and need to consider privacy issues at all stages. Simultaneously, considering various privacy requirements helps to achieve the most reasonable balance between privacy and security issues. However, it is worth considering that the concepts of privacy and security differ and can also be expanded depending on various internal and external requirements. Given the above, we adhere to the basic SDLC approach, based on the fact that many existing similar approaches are primarily developed and adhere to the primary phase SDLC layout [185, 154] (which in most cases includes the following phases: Governance, Design, Implementation, Verification and Operations). Thus, this allows the opportunity to interpret and adapt/synchronize the developed model for various (including complex) organizational tasks. We suppose that a similar approach to the concept of privacy will deliver the same (or equal) advantages in addition to many already existing ones. We argue that the potential benefits of integrating privacy into the SDLC layout, including considering the approach only from the Privacy perspective (hereinafter referred to as the Privacy Development Life Cycle (PDLC)), reflects the need and capability to adapt multinational technical standards and methodologies and allows organizations to adapt various mechanisms and approaches to improve privacy matters in mixed application areas. We acknowledge that integrating privacy into the existing SDLC, including by analogy with SDLC (Security), should be carried out according to similar principles. Simultaneously, we recognize that there is no generally accepted concept of PDLC at the moment of writing this dissertation. However, attempts are being made to present this concept in industry or academia, among which we can note the importance of the data protection domain in the context of maturity models [101]. Therefore, we argue about the need to introduce and incorporate PDLC into RMDP and focus below on the rationality of integration (supplemented by our comments).

The rationale for incorporating PDLC into the RMDP is threefold:

- Initially, this highlights the demand and need to address privacy issues at all SDLC phases [29];
- Secondly, this will enable fulfilling and integrating PbD guidance and will also minimize the complexity of interpretation between different roles and the integration of PET into the assigned tasks [106]. This is also in line with the PRIPARE approach [137], which reduces the following complexity: a) “providing a systematic methodology aimed at the complex ecosystem of all the stakeholders involved in the production of privacy-friendly systems, and which addresses the whole personal data and system development lifecycle (SDLC) of projects and systems, with the total disregard of their size and domain,” b) “detailing the engineering processes that allow a move from abstract principles to technical requirements, designs and actual implementation;” and c) “merging and connecting existing best practices in the area of PbD into a single methodology, while providing different alternatives and criteria to choose the most adequate for each context and at each stage of the SDLC.”;
- Thirdly and finally, various standards and practices (ISO, NIST, ICO UK) suggest taking into account privacy requirements from the first steps of the SDLC [137], which is also reflected in the PRIPARE approach [39]. Yet, privacy is mainly adapted at various stages but concentrates mainly on integration or support.;

The Evaluation dimension contains several phases, each with its own meaning and deliverables. There are many various SDLC types [154] that have advantages and disadvantages; in addition, organizations have the right to choose the most suitable for their purposes, as well as the freedom to adapt the recommendations of various privacy agencies (e.g. ENISA, ICO UK). In other words, in the context of interpreting privacy issues, organizations are free to choose or adapt their own path or version of the SDLC that suits the organization’s requirements or goals. Therefore, based on the above facts, we assert and argue that organizations have the right to adapt the generic SDLC, and therefore *Evaluation Dimension* can be adapted according to the organization’s requirements, which focuses on privacy issues. In other words, in the context of adapting the Evaluation dimension to the organization’s needs, this dimension will correspond to

the current compliance and privacy requirements, that is, dynamically comply with various changes or requirements in the subject area. Simultaneously, the need and importance of a) leaving the opportunity to adapt the SDLC in the context of privacy and b) inclusion in the Evaluation dimension is confirmed and recommended by the NIST framework [2]: *"Stakeholder security and privacy requirements constitute a formal expression of stakeholder protection needs across all system development life cycle phases, the associated life cycle processes, and protections for the assets associated with the system. Security and privacy requirements can be obtained from laws, Executive Orders, directives, regulations, policies, standards, or organizational mission and business requirements. Stakeholder security and privacy requirements are a part of the formal expression of required quality characteristics of the system—encompassing security, privacy, and assurance. The security and privacy requirements guide and inform the selection of controls for a system and the tailoring activities associated with those controls."*

At the same time, RMDP incorporates the overall security development lifecycle (introduced as PDLC), which is adapted into the concept of Evaluation Dimension (top left) as presented in the **Figure 5.1** and contains the following five stages (phases):

- privacy requirements engineering;
- privacy design;
- privacy countermeasures implementation;
- privacy management and monitoring;
- privacy retirement;

As we declared earlier, considering the specifics of different organizations and requirements, PDLC stages can be adapted to the needs of organizations and tasks. The complete catalogue and description of the various stages can be adapted, taking into account various recommended NIST documents [71, 22, 153]. We also understand that the list of possible documents may be supplemented due to the active development of both technologies and their supervision (from various perspectives), and the examples of documents presented are of a recommendatory nature. In addition, the remainder of this section will be devoted only to the first two stages, namely a) privacy requirements engineering and b) privacy design, since this dissertation handles concerns addressed to the privacy/compliance domain to a greater extent for these stages. We also acknowledge that further stages highly depend on the first two, and providing specific recommendations could differ from the organizational needs or objectives. Nevertheless, providing an overview of all privacy-related stages and supplemented by RMDP utilization demonstration (**Section 5.7**) will systematically lead to the desired results that fulfil the organization's goals. In other words, it leaves an unrestricted option for various organizations, including adjusting the developed reference model to suit their needs and in line with studies [142, 39].

In the context of the first stage (privacy requirements engineering) or in other words, the initiation phase, in accordance with PRIPARE [39], a risk analysis is carried out using (as an recommendation) the LINDDUN framework [46], and thereby the requirements of the policy, strategy, privacy management according to the current compliance standards according (for example PbD, Data protection by Design or by Default). Simultaneously, the opportunity for the organization to assign and modify the approach to assessing privacy requirements [137] either (a) risk-based or b) goal-oriented remains the responsibility of the organization itself in the context of the intended tasks. In other words, the proposed framework remains quite flexible depending on the assigned tasks and at the design stage. The typical output for this stage will be a privacy document, which may be supplemented with further comments and details (**Section 5.3.5** and **5.7**). The developed RMDP model plays a significant role at this stage. This helps to analyze and inventory a) data assets in the organization ((Information Taxonomy Dimension) **Section 5.3**), b) provides an overview of all possible threats and situations in the context of compliance ((Data Protection Goals Dimension) **Section 5.4**), c) allows and helps to identify countermeasures that minimize threats or difficulties

[2] https://csrc.nist.gov/CSRC/media/Publications/sp/800-37/rev-2/draft/documents/sp800-37r2-discussion-draft.pdf

(Countermeasures Dimension) **Section 5.5**)). **Section 5.6** and **5.7** interpretation between measurements and RMDP usage respectively. In the context of the second stage (privacy design), organizations require to define or update requirements in accordance with PbD or [92, 58] "*...The security and privacy requirements guide and inform the selection of controls for a system and the tailoring activities associated with those controls...*." to the subsequent design stage and after that go through the entire cycle of the reference model again and proceed to the next ones if the results satisfy the requirements of the organization and the GDPR recommendations (interpreted using DPG). As expressed before, following further privacy stages, countermeasures implementation, management, monitoring, and retirement should be completed according to the specified earlier PDLC and represented order (**Section 5.6**).

5.3 Dimension 2 (Information Taxonomy)

As declared in the GDPR, together with various recommendations from privacy agencies (ICO UK, ENISA) and the SDM [163], in order to determine the correct actions in the context of DPG and Contorumesures: one of the decisive elements is to have an understanding of "Information taxonomy" and further information attributes (namely types of personal data, level of anonymization and auxiliary parameters), this enables more accurately comply with compliance standards, as well as monitor the level of privacy for various types of (personal) data.

In the context of developing the RMDP reference model, we argue that to define countermeasures and the relationship with DPG, the concepts of personal information are not sufficiently defined in the GDPR (**Chapter 2**). We argue that extending the taxonomy and including possible categories related to sensitive information will enhance the interpretation of the offered dimensions in the RMDP reference model and the implementation of further privacy countermeasures concerning the DPGs. Simultaneously, further advantages include: a) with the service of the presented taxonomy, organizations can more effectively and sensitively prioritize privacy or compliance requirements for specific types of personal information, b) a well-defined taxonomy can demonstrate compliance requirements or how they can be applied, c) knowing where and what (personal) information is located allows simplify various tasks in terms of complying with various requests from data subjects or data controllers.

Therefore, aligning with the GDPR and the SDM framework [163] and considering the analysis from the previous section, the following attributes have been included:

- Form (Format);
- Categories: (personal, non-personal)[3];
- Anonymization levels (Recital 26, Article 4 GDPR);
- State: various forms of *processing of personal data*[4], demonstrated in Figure 5.2;

The previously introduced attributes will be discussed and explained in more detail in subsequent sections of this chapter. Simultaneously, considering ISO 29100:2011 and GDPR, the primary processing activities (collection, storage, use, transfer and disposal) can be applied to interpret privacy requirements for a more detailed interpretation of information taxonomy [169].

5.3.1 Form (Format)

Since in the main body of the GDPR, there is no detailed specification on the type or format of personal data, we shift to Article 12 GDPR, which regulates the methods and types of submission of information

[3] demonstrated in (**Table 2.1**), special categories of information (Article 9 GDPR): personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric and sexual orientation [202]

[4] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/personal-information-what-is-it/what-is-personal-data/what-is-personal-data/

to requests from data subjects from data processors. According to Article 12 GDPR, "*...shall be provided in writing, or by other means, including, where appropriate, by electronic means...*" At the same time, the default presentation of information is "written". On the other hand, the form of information presentation should be "easy accessible". Therefore, in compliance with the GDPR, we have used the following attributes below:

- **paper** (in writing implies that basic information can be presented on paper or other media; in most cases, this refers to information printed on paper),
- **electronical** (in addition to the traditional written form, any other means in electronic form, in this case by any electronic means: a document or part of the information in a specific electronic document, database, record in a database, or on other media containing an electronic record or part records),
- **verbal** (in another verbal manner where the identity of the subject is identified) or
- **other means** ("not necessarily electronic] means" may include infographics, flow charts or infographics);

The submitted forms may be convertible, and private information may be submitted on any of these forms. Simultaneously, possible privacy measures depend on the form of information. None of the solutions discussed in **Chapter 3**, including the SDM framework [163], provides or takes into account the ability to evaluate the format of private data in the proposed way. Only the Capability model [109] offers to identify possible date attributes in the context of System capabilities.

5.3.2 Categories (Information Sensitivity)

Earlier in **Chapter 2**, we introduce the definition of personal data and non-personal data, considering the interpretations presented in GDPR, which are particularly utilized in the current dissertation. We want to note that information sensitivity is applied to personal data. However, we are explaining further the choice of including the non-personal data under the referred category and how the organizations or model users can benefit from this. As we previously stated (**Section 2.2.2**), the GDPR refers to or applies only to personal data; in other words, if the data is non-personal, this excludes the GDPR application. Therefore, we argue that clearly distinguishing what qualifies as "personal data" or "non-personal data" determines the scope or norms applied to the data set. In addition, this makes it possible to determine whether the entity processing personal data is subject to the various restrictions presented in the GDPR. In the context of restrictions, we mean the regulatory requirements of the GDPR applicable in this case. In other words, in the context of different categories of personal data and the possible requirements presented in the GDPR for data processors or controllers leaves the organization room for manoeuvre based on the description of data category, considering the regulatory approach [93], as well as further activities to comply with GDPR. Nevertheless, due to the development of various technologies, including additions to various privacy standards and even expansion of the concept of representing data protection and data privacy, or privacy taxonomy [93], the concept of personal data can also be considered dynamic. Therefore, in accordance with [57], we adopt (**Figure 2.1**) an assessment scheme to assess the category of data that organizations can utilize. On the other hand, this will allow organizations to provide a) the ability to determine the category of personal data, b) integrate various recommendations when working with personal data, c) supplement the scheme for assessing the category of data (depending on the tasks, or taking into account various recommendations of privacy agencies) and d) more promptly monitor possible risks for data subjects (while respecting the rights of data subjects during various requests). We also want to note that the suggested assessment scheme could be modified or changed to fulfil the organization's goals. During the analysis in **Chapter 3**, only the method by Kos [102] presents a partial opportunity to determine the data category, including the data flow. In comparison with the SDM framework, the non-personal category is not included. However, the concept of personal data is used in practice to evaluate DPG and the process of compliance with GDPR requirements.

5.3.3 Level of anonymization

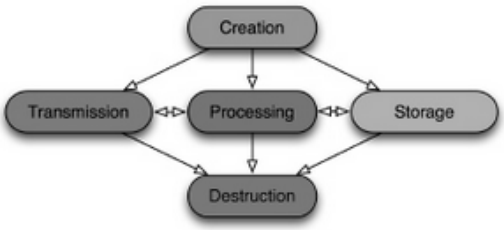
As we have already introduced and debated the concepts and levels of anonymization (full anonymization and pseudonymization) in **Section 2.2.1**, GDPR norms do not suggest any other. Simultaneously, according to our results [149], the importance of the presented anonymization levels in the GDPR is argued to be necessary, where the most significant interest and preferences are also expressed in accordance with full anonymization while utilizing PET. We also recognize that diverse organizations or data processors may interpret various anonymization levels differently, or in other words, relying on their own goals, objectives and ways to gauge the referred levels. Therefore, we acknowledge that for the developed reference model to be fully operable without limitations, it is necessary to consider both anonymization levels. Simultaneously, we believe that due to the fact that defining the boundaries between different levels of anonymization, firstly, there is no straightforward approach that would be presented either in the GDPR or other recommendations, therefore, responsibility is fully distributed to the data processor or third parties. Secondly, the issue of assessing the anonymization level still remains a relatively dynamic and actively researched area. We have proposed an alternative approach to assessing the anonymization level for DP [150], which can be utilized by organizations as a potential alternative (we discuss the proposed approach in more detail in **Section 5.5**). Thus, the validity or ability to take into account various anonymization levels expands the possibility and functionality of the reference model and also complies with the PbD principles (**Section 2.3.1**) and the requirements of the GDPR for PI and the responsibility and obligations of various parties involved. In other words, we want to remark that the impossibility of interpreting various anonymization levels in the context of the goals of this work should not limit the reference model being developed. We also think that the capability to assess anonymization levels can be complemented by other concepts/measurements suggested in the RMDP. We also argue that taking into account the dependence of various use cases and the presented blocks makes it feasible to evaluate the effect in the context of the DPGs discussed below and evaluate the significance of the applied privacy measures using other blocks (the “Evaluation” and “Countermeasures” Dimensions). On the other hand, this does not contradict the only one presented DPIA approach (**Section 2.4**) in the GDPR, which indicates the possibility of assessment only in case of high risks associated with the processing of personal data, which is one of the significant limitations [171]. In other words, we accept that the importance of representing all available levels of anonymization in the reference model being developed does not limit the use of the model and does not contradict the principles of BES [181], the requirements of RMDP (**Section 4.3**) and allows minimizing the risks of eliminating the likelihood of non-compliance with the requirement even under exceptional conditions if the information corresponds to the level of complete anonymization.

5.3.4 State

As demonstrated earlier, the Information Taxonomy dimension consists of various information attributes, where such attributes enable and enhance more accurate interpretations of GDPR requirements (**Section 5.3**). Additionally, as presented earlier, the form (**Section 5.3.1**), categories and status of personal data (**Section 5.3.2**), level of anonymization (**Section 5.3.3**), determine and participate in measures to enhance privacy, including compliance with DPG and measures aimed at fitting with DPG. Together with previously mentioned attributes, data also can be in different states or, in other words, in one of the states presented and explained further through the data life cycle depicted in **Figure 5.2**. Simultaneously, the information state particularly in the current study, suggests fulfilling the requirements[5] imposed in the context of processing personal data considering GDPR and we argue that considering diverse information states leads to the enabling fulfilling the GDPR requirements and enhancing privacy explained further. In order to reflect on various changes in diverse practices (data management) in the context of data protection regulations frameworks, we refer to (or utilize) the data life cycle as a frame for reference. According to [81], there are different variants of the life cycle, as well as different stages associated with them. The principles of the data life cycle are applied in various data-related domains [110], to which different requirements may

[5] www.edps.europa.eu/sites/default/files/publication/it-governance-management-en.pdf

apply according to diverse GDPR principles and conditions. In other words, the data, systems and various service components presently mind the GDPR requirements. Simultaneously, according to (Collins English Dictionary 2019), in an abstract form, "*the life cycle of something [...] is the series of developments that take place in its beginning until the end of its usefulness*." On the other hand, according to the GDPR, the "process" is not clearly defined and can be represented by sub-processes and according to (Art.4 No.2 GDPR): for example, "collection, recording, sorting or storage until erasure or destruction", thereby also reflecting the main stages of information are presented in **Figure 5.2**. On the other hand, the need to take into account the state of information is dictated by the GDPR requirements, and simultaneously, the state of information must be appropriately classified to determine the level of sensitivity or criticality, which enables various organizations to meet the required level of privacy [89]. Therefore, various organizations, in accordance with the GDPR requirements, need to track various process activities with personal data and also meet the seven further principles described in Art. 5 GDPR (lawfulness, loyalty and transparency; limitation of purposes; data minimization; accuracy; limitation of retention; integrity and confidentiality; responsibility), in order to be able to respond to various requests from data subjects (the right to be informed Art. 13, 14 GDPR; the right to rectification Art. 17 GDPR; the right to data portability Art. 20 GDPR; the right to object to data processing activities Art. 21 GDPR and rights in relation to automated decision-making and profiling Art. 22 GDPR). In other words, we represent the state of information using the following states: Creation, Transmission, Processing, Storage and Destruction. Simultaneously, we recognize that alternatives can supplement or modify these conditions. For example, from generic abstract data lifecycle (ADPL) [5], and also has the following advantages among the available data lifecycle (in the context of privacy and data protection): a) allows the integration of PbD principles, b) has a more flexible structure, or in other words flexibility in the context of adaptation for different roles and tasks and types of organizations), c) the ability to interpret different privacy recommendations or requirements (that is, not only GDPR), d) understand PIA or DPIA and e) synchronize with the requirements specified earlier for the reference model being developed. However, we believe that allowing flexibility and providing the ability to integrate or supplement baseline information states is highly consistent with the goals of the RMDP reference model being developed.



Note: contains further states considering [34] (Creation, Processing, Storage, Transmission, Destruction) and GDPR, adopted in the Information Taxonomy Dimension of the developed Reference Model.

Figure 5.2 A simplified representation of the Information state

5.3.5 Information Categories

Further, we highlight the relationship between the attributes previously presented for the dimension of Information Taxonomy, which could be utilized for further development of specific scoring functions (elaborated as future work in **Chapter 7**), which enhances the capabilities of the developed reference model and further efficient performance in a diverse organization. The examples of attributes presented earlier (form, categories, anonymization levels and state) represent the dimension of Information taxonomy or, in other words, a group of information assets with similar characteristics. We argue that the presented taxonomy enables the most transparent display of the characteristics of data assets and also provides the basis for specification and compliance with DPG (thus GDPR requirements) and countermeasures. Simultaneously,

we believe that the presented attributes have the capability to complement each of the presented attributes and thereby satisfy the requirements of the developed RMDP reference model. Thus, the information is represented using various attribute combinations above. We also want to note that studying and developing a specific scoring function [6] is not the primary goal of the current dissertation. Nevertheless, we assume that elaborating and providing organizations with the opportunity to acknowledge and highlight the example based on the Information Taxonomy Dimension leads to significant or further improvements in implementing the developed reference model.

$$I = \{F, C, A, S\}$$

where, I - information categories; F is a form and F belongs to paper, electronical, verbal, other means; C is category and C belongs to personal, non-personal; A is the level of anonymization and A belongs to pseudoanonymization, anonymization; S - state and S belongs to creation, transmission, processing, storage, destruction;.

5.4 Dimension 3 (Data Protection Goals)

DPGs represent desired compliance goals considering GDPR. The RMDP includes a group of seven widely practical DPGs (**Figure 5.3**) adopted from SDM framework[7], which include Data Minimization, Availability, Integrity, Confidentiality, Unlinkability, Transparency, and Intervenability (**Appendix A**). The presented group of goals was determined according to the methodology presented in **Section 4.1**. Simultaneously, considering our results [148], the capability to interpret DPG in the context of DP is feasible, but it is worth considering that it may depend on various DP forms, including various use cases. In addition, our results [148] demonstrate the capability to interpret the anonymization levels and DPG. We argue that each organization should establish its priorities or determine the importance of each DPG depending on its own objectives and goals. For example, organizations that do not consider the DPG of data minimization or confidentiality most likely, do not deal with personal data. In other words, each DPG maintains a different significance for diverse organizations or roles involved in assessing the degree of compliance with GDPR. This is also observed from the results of our work [149], where various experts (from diverse organizations) took part and more than 50 participants were questioned. We also consider the importance of reconsidering the assessment of DPG priority due to the importance of the concept of "time" or "time interval" (**Section 5.2**), which also follows from our results [149] where various researchers indicated different preferences from other groups surveyed.

Considering the importance of the DPG and according to the ISO/IEC 27005:2011 specification [8], we will assign or enable each goal to appropriate criticality levels. The available set of values (low-medium-high) is widely accepted and used in the risk privacy assessment domain [151]. According to the objectives of this dissertation, DPG prioritization implies that each DPG is assigned or given a so-called range criticality (low-medium-high). The choice of the range also does not contradict the requirements demonstrated in Art. 24 GDPR that compliance with codes of conduct and approved certificates (in this case, ISO/IEC 27005:2011) can be used as an element of compliance. Simultaneously, we point out that the methodology for assigning range criticality to DPGs may differ for different roles and is beyond this dissertation's scope. Simultaneously, in agreement with the work [62], we acknowledge that DPGs (offered by the SDM framework) are the most suitable option for trainees, as they more clearly summarise the abstract GDPR requirements for various organizations.

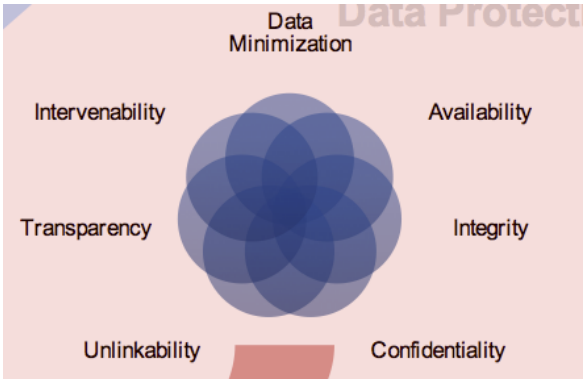
We also accept that the assignment or selection of a criticality assignment methodology for each DPG is a decision for each organization and thereby allows for flexibility in the RMDP model being developed

[6] Local scoring function applicable for the specific dimension or Global scoring function applied to the whole reference model, also dependencies on the referred functions

[7] https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methodology$_V2.0b.pdf$

[8] https://www.iso.org/obp/ui/iso:std:iso-iec:27005:ed-2:v1:en

according to the requirements specified in **Section 4.3**. Considering that the methodology for assigning criticality is beyond the scope of this dissertation, we leave the opportunity for organizations to choose a methodology guided by their own goals or one existing at a given time when utilizing the RMDP reference model. We also argue that using the presented criticality range in the case of the organization's chosen methodology, in the case of high risks, allows signalling to the DPO or other roles about the need to conduct a DPIA (**Section 2.4**).



Note: Fragment from Figure 5.1

Figure 5.3 Data Protection goals depicted in RMDP (simplified representation)

5.5 Dimension 4 (Countermeasures)

Demonstrating compliance with GDPR nowadays is a reasonably complex undertaking and presents a process with many challenges for various organizations [28]. In other words, the GDPR, being a normative document, does not provide a notable or specific solution for implementing its own requirements that apply to TOM. However, concentrating on Art. 25 GDPR, referring to the principles of Data Protection by Design (DPbD) and the principles specified in Art. 5 GDPR (**Section 5.3.4**), allows (the Data Controller/Processor) to determine reasonable and necessary countermeasures. Simultaneously, as the analysis of comparable approaches (**Chapter 3**) demonstrates, in nearly every approach or framework, mixed countermeasures appear in various interpretations, but to one degree or another, they are applied similarly to TOMs (as specified, in the GDPR) or their parts. However, despite similar approaches, there is no single pattern or specific solution (examined earlier) for enforcing GDPR requirements (in other words, they are diverse). Also, according to the analysis in **Section 2.3.1** and considering Recital 78, the GDPR indicates that "*technical and organizational measures should be taken*", and referred measures must comply with the Data Protection by Design or DPb Default Requirements. Thus, we can conclude that the complex design shown above, following various GDPR principles, does not provide a ubiquitous approach and leaves the choice of countermeasures to the data controller/processor. This, in turn, must suggest and further develop/integrate an approach that satisfies Recital 78 GDPR so that "*technical and organizational measures should be taken*" and referred measures must comply with the Data Protection by Design or DPb Default requirements.

In addition, considering our results [149] - we examined the possibility of an alternative approach, which can further be characterized as a specialized solution that enables to address this issue. However, as the results of our study indicate, the presented technical possibility is possible only by determining the privacy taxonomy either in the proposed way [149] or another possible alternative. Simultaneously, considering the dynamic structure of privacy, the utilization of only a technical solution may limit the assessment of compliance in a broad sense for PET DP. In addition, it is worth considering that only the technical and/or suggested DPbD requirements [198] and, accordingly, the GDPR recommendations do not provide the chance to fully address privacy issues due to the abstract nature of the requirements. Similar conclusions can be drawn from the variety of approaches and their objectives analyzed in **Chapter 3**. However, despite

the above issues, many organizations address privacy issues for TOM together with the ISO/IEC 27k family of standards (**Section 3.2.2**). We also want to draw attention to the fact that in adapting various ISO standards, it is worth considering that organizations have the right to utilize various ISOs at their discretion, which is not a mandatory condition. However, compliance with GDPR is a legal requirement. Nevertheless, despite the significant addition with the help of the ISO/IEC 27k family or other standards and recommendations, the issues of integration of these recommendations are in practical focus: ISO/IEC standards, as well as their combination with various technical solutions, can be improved [3]. An alternative approach for addressing privacy issues in addition to GDPR requirements can also be achieved using the NIST security control framework [28]. In conclusion, we would like to note that the expansion of the privacy countermeasures concept does not contradict the GDPR principles and only complements the capability of organizations to apply privacy countermeasures or their feasible combinations.

Therefore, according to the objectives of this dissertation, it is necessary to take into account and apply an integrated approach to privacy countermeasures in the developed RMDP reference model. We believe that considering the variety of various countermeasures in the context of the developed RMDP leads to the so-called holistic approach to assessing privacy. Based on an analysis of various studies (**Chapter 2**), including comparable solutions (**Chapter 3**), including taking into account the lack of a universal approach for assessing privacy. We interpret and adapt the fourth dimension of RMDP (lower left quadrant), which represents four (high-level, abstract) types of privacy Countermeasures:

- **Technical** - technical means which are utilized to reach the DPG;
- **Organizational** - organization means (administrative activities which enhance the environment to enable privacy countermeasures to be executed according to the DPG);
- **Legal** - legislation means and contract agreements (service level agreements, job contracts, employer and third-party non-disclosure agreements, laws);
- **Human** - means of minimizing risk from human-factor (education, training, creating awareness and motivation, ethics);

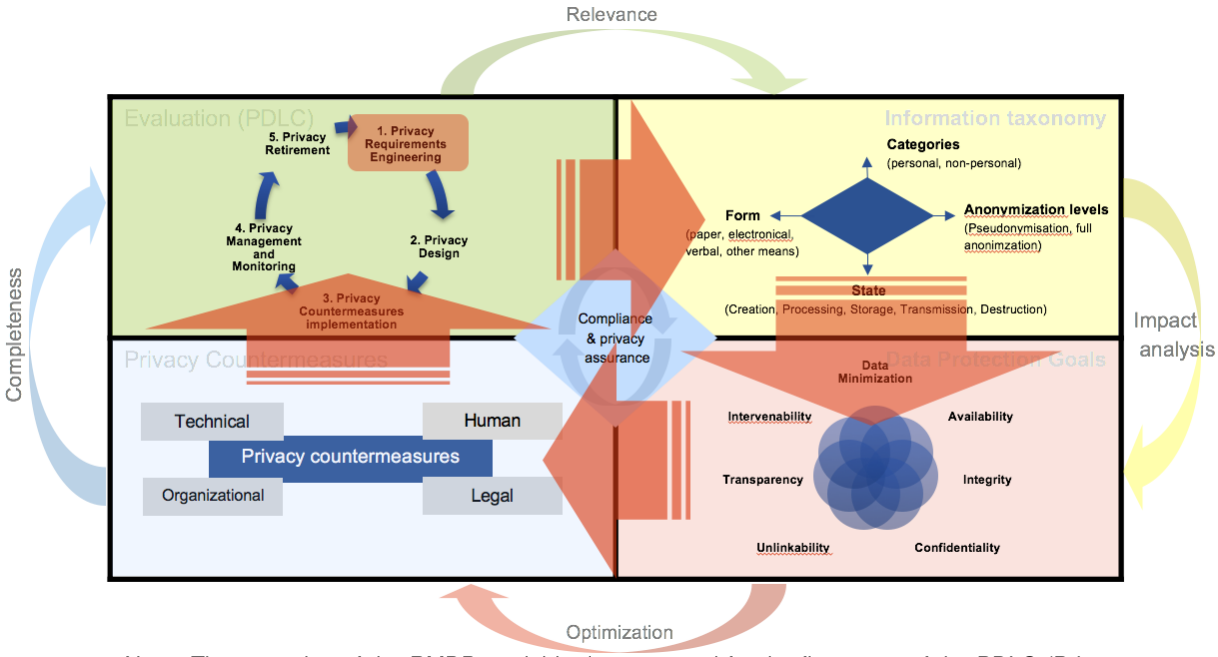
We will introduce a more detailed presentation of the four types of countermeasures presented and examples in **Appendix D**. None of the analyzed approaches or methods (**Chapter 3**) supports all four privacy countermeasures simultaneously. In addition, there is no single standard or established taxonomy among academia or industry, including various industrial standards, for the presented privacy countermeasures. Simultaneously, referred countermeasures are offered in various industrial standards (an example of the ISO/IEC 27k family) and are mostly seen together in the context of security and privacy measures. The developed RMDP model integrates only high-level abstractions of the presented categories and, at the same time, places greater emphasis on the privacy domain.

We argue for the inclusion of high-level abstractions on the basis that this field still a) remains dynamic (multiple studies are being conducted in academia and industry), b) reflects the essence and dynamics of the privacy concept, c) this level of presentation remains accessible to the audience for which this model is being developed and d) leaves the opportunity for organizations to modify or supplement the presented categories depending on the objectives and various requirements. We also discuss the validity of these categories in a later chapter. The suggested categories are not exhaustive or complete; we introduced these categories to represent the nature of the privacy countermeasures since a deeper and more expanded interpretation of privacy countermeasures with the development of a separate taxonomy is beyond the goals of this dissertation and is not a developed artifact of this work. Simultaneously, the abstract representation of privacy countermeasures satisfies the requirements of the developed RMDP reference model, considering the broad audience that can utilize and apply the referred model.

5.6 Interrelation between Dimensions

This section is dedicated to the interaction between the previously presented four dimensions, and it also illustrates the execution of RMDP, including the concepts offered for each dimension. As we noted earlier and considering our analysis (**Section 3.4**), the interaction and execution of models can be carried out in various ways. We further explain and demonstrate the execution of the model, emphasising the point of start and supplementary details, including the four arrows around the model. We also extend and highlight further details about the RMDP model regarding usage and visual interpretations, specifically in **Section 5.7** and **5.8**.

The developed reference model RMDP (**Figure 5.4**) provides a visual representation of the sequence and relationships between the four dimensions. The arrows around the model serve as a clear guide, naturally displaying the sequence of use and relationships between the four dimensions. We believe that the presented mapping and consistency are, to some extent, in synergy with the standards (ISO/IEC 27001[9], 27701[10] for privacy management within the context of the organizations) and generally accepted principles of the Article 32 GDPR and DPIA[11] (**Section 2.4**). This model is not just a theoretical construct but a practical tool that can be applied in real-world data protection scenarios.



Note: The execution of the RMDP model is demonstrated for the first stage of the PDLC (Privacy Requirments Engineering); the current PDLC stage and Evaluation dimension illustrate the beginning of the model utilization. Execution illustrated in red demonstrates the step-by-step process of performing the RMDP model through all four dimensions. Detailed model execution is explained further in the current section.

Figure 5.4 The sequence of execution and use of the developed reference model RMDP

The description and execution of the developed model (including the relationships between dimensions) begins with the upper left quadrant of the RMDP reference model. Further movement and interpretation of each block are carried out clockwise. The organization (performer), according to the proposed PDLC (Evaluation dimension), determines the current phase of the PDLC and then moves step-by-step clockwise along each quadrant (dimension) and after passing each quadrant (in this case, Information taxonomy, DPG and Countermeasures) returns to the quadrant from which started moving (i.e. Evolution). Thus, having completed a full circle or movement through all quadrants (dimensions), the organization

[9] https://www.iso.org/obp/uiiso:std:iso-iec:27001:ed-3:v1:en
[10] https://www.iso.org/obp/uiiso:std:iso-iec:27701:ed-1:v1:en
[11] "should follow as structured an approach as possible and comprehend the documentations produced during the DPIA"[122]

(or performer) has the opportunity to move to the next stage of PDLC (in the Evaluation dimension) and continue to use and implement the RMDP reference model further. We also want to comment that running the entire PDLC in the Evaluation dimension does not have to be completed entirely before proceeding on to another dimension.

As an example, if a performer or organization plans to utilize the RMDP reference model to evaluate the degree of compliance with GDPR in the context of using DP and developing a document in the prescribed form (privacy policy document). Before applying the RMDP reference model, the organization or performer should adjust or modify the model according to the interpretation of each dimension (or concept enclosed) according to the remarks in **Section 5.1** (also taking into account the possible interpretations presented in **Sections 5.2-5.5**). The RMDP reference model plays a significant role in this process. We will introduce a more thorough example of adapting the RMDP reference model in **Section 5.7**.

The organization (performer) identifies and/or determines what is at the privacy requirements engineering stage (Evaluation dimension, upper left). Afterwards, the organization (performer) moves clockwise to the Information taxonomy quadrant (upper right quadrant). At the (privacy requirements engineering) phase, the organization (performer) utilizes the previously presented taxonomy (**Section 5.3**) to conduct a complete so-called information/data inventory. Thus, a privacy document will be produced, listing the categories of information that is present. Then, after the list of information has been compiled, the transition to the next quadrant (DPG) is carried out. This quadrant enables to set or specify data protection goals for each category based on various attributes (**Section 5.3**). After completing the DPG stage, the transition is made to the Countermeasures quadrant (lower left quadrant), where countermeasures are identified or selected (including prioritization) that can help or facilitate the achievement of the DPG (previously defined) for each category of information. This completes the entire cycle for a full circle of all measurements (blocks) for one PDLC stage. Once this entire cycle has been completed, a Privacy document is produced (**Section 5.7**). The next stage returns to the Evolution quadrant and the next stage of privacy design, for which a privacy document has already been created. After this, movements are made clockwise, and movement occurs through all RMDP blocks, where these blocks ensure that previously classified information assets, as well as approved and defined DPGs in the previous stage, are consistently included and fully taken into account. After returning to the Evaluation quadrant, the second cycle of the reference model is completed. In subsequent phases for the PDLC (namely implementation and monitoring), the organization ensures that the early a) established or defined DPGs are carried out for each category of information (**Section 5.3**) and b) countermeasures are functioning and implemented as planned (i.e. in accordance with the results of previous phases and developed by privacy document). In other words, we presented the relationship between the loops, including how each loop addresses each dimension (quadrant) required for each RMDP loop/PDLC phase. Next, we comment on each arrow that directs the execution cycle for the RMDP reference model (**Figure 5.4**), including commenting on the idea followed in each arrow direction's context.

By analogy, with the description of the relationships between dimensions (commenting on the reference model (upper left quadrant), we also begin commenting on each arrow from the upper left quadrant.

The top arrow (Relevance) determines that based on the first stage of the PDLC, the organization must identify each information asset according to the information taxonomy (upper right quadrant). In other words, the organization audits and identifies information assets according to attributes (**Section 5.3**), we also show how this register can be created in **Section 5.6.1**. We believe and argue that the Completeness and high precision of filling out the specified register (hereinafter referred to as the privacy document) meets and contributes to the provision and implementation of the presented DPG (for each category of Information taxonomy). We also express that (in some cases) the priority of various DPGs may be different for various categories of information (**Section 5.3**). The right arrow (named Impact analysis) or the arrow from Information taxonomy to DPG indicates that the organization assigns a particular priority to each category of information by specifying a priority for each DPG. The organization determines the priority for each category of information utilizing a risk assessment. The developed RMDP reference model is neither a risk identification tool nor a risk analysis or assessment methodology. However, RMDP points to the importance of risk analysis in the context of an organization's data protection objectives. In addition,

the RMDP sets out requirements for a risk analysis methodology that facilitates the prioritization of data protection goals, the creation of a (detailed) information registry, and the interpretation of DPG priorities between different roles. The bottom arrow (Optimization) from DPG to Countermeasures illustrates the transition between DPGs and their priorities in the Countermeasures block. It indicates the effectiveness or implies the most optimal way to achieve the specified DPGs and their priorities. We also want to note that the organization attempts to eliminate possible risks and ensures that the selected countermeasures do not interfere with each other or reduce the effectiveness of achieving DPGs and their priority. By considering or analyzing the full range of DPGs, we believe organizations can extract the most practical combination and perform the best outcome in the context of GDPR compliance (which integrates the best possible response to privacy risks). The arrow (Completeness) from Countermeasures in Evaluation dimensions shows and is interpreted as the organization's selected privacy measures that can be tracked and consistent across all PDLC phases. The presented circle (or complete cycle), which is formed from four arrows, demonstrates that the model should be applied 1) clockwise, 2) sequentially a) moving from each quadrant to another, b) from each PDLC stage, performing an entire loop for each quadrant and c) a cycle or fulfilment of a full circle demonstrates the proper use of the developed RMDP reference model.

We also argue that organizations have the freedom to modify and adapt the execution of the model and its dependent components depending on the organization's objectives and dynamic factors. However, it is crucial to understand that if one of the RMDP settings or elements is adapted or modified, it is necessary for organizations to complete the entire PDLC cycle and all cycles of the RMDP model after the changes or modifications are made. This understanding is critical to successful implementation.

5.7 RMDP Usage

In addition to the description of each block (descriptive data in **Sections 5.2-5.5**) and the concepts presented in each block, the developed RMDP model also adapts methodological knowledge. This section explains how a privacy document is compiled and the RMDP's role in the referred process. The concept of a privacy policy has a broad meaning. In the GDPR context for organizations, it is a (broad) privacy document that brings together the various privacy-related requirements that systems must fulfil and enables data subjects to take informed actions in accordance with the GDPR [203]. On the other hand, privacy policies are complex and massive documents that are challenging to interpret or understand [115, 117, 126]. However, there are various approaches to reduce complexity in interpretation [194]. In other words, and in the context of the current dissertation, the concept of a privacy document should represent the various requirements discussed in **Section 5.3.5**, reflecting the GDPR privacy policy for diverse organizations. Further, in this dissertation, a privacy document refers to documents that define what DPGs must be accomplished and what privacy countermeasures must be incorporated (at a high level of abstraction). We believe that a high level of abstraction a) corresponds to the dissertation goals and conducted research, b) privacy policies do not contain complete information (which PET or other hardware, software solutions to utilize to accomplish privacy goals in the GDPR context), c) corresponds to the goals of utilizing the developed reference RMDP model for a broad audience and d) enable incorporating additional explicit recommendations and technical, legal or proper documents and various guidelines. We also want to note that we previously partially debated that each RMDP component can complement other concepts or dimensions. Moreover, we accept that leaving the presented level of abstraction for the privacy document, in combination with other dimensions, will enable the necessary efficiency in developing and supplementing the further presented privacy document for various organizations. The use of RMDP is explained subsequently in this section, including remarks.

The table format is the most optimal for outlining a privacy document and enables the combination of the concepts utilized for each quadrant (dimension). **Table 5.1** is a table with X columns and N rows. We demonstrate all the offered attributes and combinations (information attributes **Section 5.4**) in that we comment on the suggested sample filling. Simultaneously, leave the following columns (list) blank. The maximum number of possible category combinations, which is proportional to the number of rows, is the following formula:

	1. Form	2. Categories	3. Anonymization Level	4. State	5. DPGs	5. DPGs crtiticality	6. Countermeasures
1	paper	non-personal	none	processing	Data minimization	low	not required
2	paper	non-personal	none	transfer	Confidentiality	low	not required
3	paper	personal	none	processing	Data minimization	high	Technical: no or as little as possible personal data processed; data fields which enablethe identification of data subjects may be erased or transformed (anonymization or pseudonamization).
4	electronical	personal	pseudonymization	transfer	Transperency	medium	Technical: data subjects, processors and controlles could be able to identify at verying degrees which data is collected and processed when and at what purpose; Organizational: who has legal responsobility for the data and systems in the various phases of data processing;
5	electronical	personal	none	processing	Confidentiality	high	Technical: no unathorised pesrosn can access or use personal data, confidentiality of PI should be ensured that measures are take to remedy and mitigiate the accompaning vialotaion of the protection of PI

Table 5.1 Development of Privacy document utilizing RMDP

Note: Table 5.1 illustrates an example and a simulation of the Privacy Document (high-level abstraction) during the utilization of the RMDP, which aims to document the results of RMDP use. Creating a Privacy Document starts together with the first utilization of the RMDP. Current simulation examples demonstrated the results of the execution of the first PDLC and served as a sample stressing the utilization of the concepts introduced for each dimension of RMDP.

$$N = N_C * N_A * N_S * N_F * N_D$$

where, N - total number of combinations, N_C - number of information **c**ategories, N_A- number of **a**nonymization levels, N_S - number of information **s**tates, N_F - number of information **f**orms, N_D - number of **D**PGs;

A privacy document is generated in this way, where each line (including various categories of DPGs and Countermeasures) reflects a scenario or possible scenario groups in which information categories may require action in the context of observation with GDPR requirements, including data protection according to the suggested DPGs. For each line, the organization sets (in this case, for each category of information) the importance (or high-medium-low criticality level **Section 5.4**) for each DPG. The possible amount of DPG category can be calculated using the formula shown above. We also want to note that not every line or possible combination may be appropriate for every organization. We also want to note that not every combination requires countermeasures, and, as a result, some lines for the countermeasures column may be empty (in some cases, this depends on the organization's goals and objectives). We also emphasize that the accuracy of filling out the privacy document is relevant for more accurate compliance with the GDPR requirements. In particular, we advise that even in the event of the absence of a relevant countermeasure, we consider it necessary to include them in the privacy document. Thus, this will enable not to ignore a single compliance violation and enable more careful monitoring and taking the necessary countermeasures. An example and recommendations are presented above, including those presented using **Table 5.1**, which we consider sufficient and convey the necessary logic of how privacy documents can be compiled, supplemented or created using RMDP. This section demonstrates that the developed RMDP reference model has all the necessary capabilities for the study area. We believe that following the presented logic, the use of RMDP will enable to achieve the most accurate drafting of a privacy document and the highest level of compliance with GDPR requirements.

5.8 RMDP Visual

As we declared earlier, one of the goals of developing the RMDP was to simplify the representation of the study area in a simple, clear and accessible visual interpretation. These goals were one of the drivers in the development of the RMDP reference model based on the following facts: a) visual presentation (in relation to the current dissertation, this is the reference model) motivates people to read [72] and b) people are more receptive and remember visually presented information more effectively [114], c) mind map approaches are familiar to many and easily digestible [205]. Although the developed reference model was conceived in the form of notation form free, as we noted earlier, some elements may resemble the mind map approach. One of the significant advantages of the mind map approach [205] is the practical

assistance with the help and understanding of complex structures and knowledge, which allows various participants, regardless of the level of expertise and knowledge of the subject area, to participate equally when discussing a phenomenon or decision. We also believe that the additional colours utilized in the development process of the RMDP can enhance the perception and interpretation of each block shown. Simultaneously, we acknowledge and argue that the model does not lose its functionality or information even in a black-and-white interpretation.

Further, we will provide comments about each of the colours associated with each RMDP block (the motivation was drawn from the analogy with the colour form and meaning represented in the well-known traffic light concept):

- **Green color** (upper left quadrant) means the beginning of a movement or the start command (we associate this color with the beginning of work and adaptation of RMDP to the needs of the organization);
- **Yellow color** (top right quadrant) means attention or be prepared; in the case of RMDP, it reflects and promotes attention to the information assets and its categories;
- **Red color** (lower right quadrant) means the stop command, in other words, in the RMDP interpretation, it represents special awareness to data protection goals;
- **Blue color** (down left quadrant) is not a traffic light color, but we were motivated by the fact that blue colours signal a "must do" or "mandatory" action (ISO/IEC 7010:2019[12]);

The developed RMDP reference model also contains various geometric shapes to support and comply with the following interpretations and associations:

- Rhombus (upper right quadrant) - associations with four attributes of the information taxonomy dimension;
- PDLC is similar to the visual representation of a loop process, as it explains the nature, gradualism and necessity of alternating phases;
- DPG - presented in the form of seven overlapping circles, indicating the number of DPGs, where each DPG is presented similarly in the form of a circle with seven connections points, confirming the granularity of DPGs and their interconnection;
- Countermeasures (lower left quadrant) - a form of rectangular the edges of which indicate four possible privacy countermeasures;

5.8.1 Not evident concepts in RMDP

As we have noted previously, through the RMDP model development process, some elements or pieces were indicated and not part of the current dissertation. In other words, although the reference model RMDP emphasizes specific knowledge in the context of privacy and data protection to accomplish the goals of the current dissertation, some elements or concepts are abstracted or blurred. Next, we discuss and present concepts that need to be more clearly defined in the developed RMDP model utilizing Ockham's razor approach [21], which describes the option to focus on frugality, brevity and economy in translating a problem. In other words, the entity (in this case, the RMDP model) should not be overloaded, and it is preferable to use more straightforward explanations, interpreting the Razor approach. This section also tries to justify the not-included concepts presented below.

Risk is one of the critical concepts in the data protection domain, and it can be applied in various interpretations to both technologies and legal notations. Within the RMDP, risk is presented for the first quadrant (Evolution) for the first stage of privacy requirements. Also, the concept of risk is one of the

[12]https://www.iso.org/obp/uiiso:std:iso:7010:ed-3:v2:en

drivers in the assessment of DPG, which further determines the priority of the presented DPG and their interpretation for various areas in the organization in the context of achieving the organization's goals. Thus, risk in the RMDP reference model is represented and conveyed through DPGs, as well as the criticality of each DPG. Thus, the concept of risk is conveyed through the criticality of each DPG and compliance with the requirements of the GDPR. However, we argue organizations has flexibility and could utilize various industry interpretations or specifications and also enable a crosswalks between various approaches [13].

Threats - there are two approaches: goal-based and threats-based (**Section 5.1**); in this case, we define goal-based since the RMDP reference model being developed is intended for a broad audience. Therefore, we speculate that the goal-based approach is more optimal for these purposes.

5.9 RMDP Discussion

This section is devoted to discussion in the context of the developed reference model and concepts utilized.

5.9.1 Abstraction RMDP

The developed reference model illustrates a separate opinion on and perhaps the only one in the data protection and privacy domain for PET (DP) considering GDPR. We also suppose that the suggested reference model can be not only a single opinion that summarizes multiple pieces of knowledge in the data protection domain. We also argue that the proposed RMDP reference model is the most suitable for the purposes of this dissertation and research objectives, especially for the prospective audience who intends to utilize the RMDP model (including the chosen form of presentation).

The data protection domain itself is a dynamic ecosystem and, simultaneously, a complex and heterogeneous area. The RMDP model, however, is designed to be comprehensive, aiming to cover the wealth of knowledge presented in this complex and heterogeneous domain. In addition, using PET and adherence to the data protection domain is an added complication in the already presented complex and heterogeneous domain. At the same time, while endeavouring to systematize and cover the wealth of knowledge presented in the current domain, in the context of the RMDP model and others analyzed in **Chapter 3**, no one model has and cannot be absolutely accurate and reflect all existing knowledge. In other words, the reference model being developed or another model may include a) possible omissions to achieve the goals of the model designer and cope with the complexities in the subject area, b) the possibility of bias towards the competence, set of knowledge and view of the model designer's problem under study and c) designed for a specific task, and not be functional for other duties.

Like other reference models, the developed model may sacrifice the accuracy of some details or the precision of the concepts applied to display the entire breadth and complexity of the modelled and studied area. We also recognize that when attempting to model a complex domain and the solutions utilized to develop and design the RMDP model, they may cause controversy and criticism. Considering possible criticism and controversy, in this chapter, we support the choice of decisions made in the context of the RMDP development (including the use and form of the concepts used in RMDP). Firstly, one of the goals of RMDP is the ability to be utilized in various roles with diverse background levels or lack thereof. In other words, the use of this reference model extends to a broad audience. As a consequence, the inclusion of different concepts and the use of different levels of detail are among the reasons for the use of this model by different roles and with diverse expertise. Therefore, to suit the purposes (for a broad audience), the use and adaptation of high-level abstractions have been utilized in RMDP (in some cases excluding the exact application of the presented concepts in proper, precise use for various fields, namely technical, legal or other possible). In other words, this made it feasible to extend the audience that can operate or apply RMDP without sufficient expertise in various domains (technical, legal and privacy engineering).

We are convinced that the developed RMDP model is universal, like other reference models available and delivered in a way as, in general, reference models should be presented. RMDP strives to

[13] https://www.bitkom.org/sites/main/files/file/import/170919-lf-risk-assessment-eng-online-final.pdf

accommodate various interdisciplinary knowledge and domains of data protection and GDPR compliance/recommendations in the context of PET, which are critical for various organizations. Simultaneously, RMDP may require adaptation or synchronization with the organization's processes or goals in the context of customization or modification of the concepts used in RMDP. The structure of the proposed RMDP model is comprehensive, providing for and considering possible changes in concepts, technologies, and processes related to protecting personal information and fulfilling GDRP requirements. This comprehensive structure ensures that the model remains flexible in the context of various shifts in various ecosystems (both technical and legal). In other words, for each quadrant, it is possible to supplement or change the concepts used (for example, DPG to add or change those presented in this model; Information Taxonomy can also be supplemented, and Countermeasures could be expanded). We also argue that the various organizations/users of the developed RMDP reference model are not just required but motivated to think substantively about various aspects of data protection, privacy domain and GDPR compliance that may be critical for the activities of various organizations. Simultaneously, we argue that while utilizing RMDP not only highlights privacy issues in the subject area but also has the necessary functionality to ensure an understanding of how to address them to a broad audience. In addition, the RMDP illustrates a logical, complete structure and platform for the various roles involved in assessing the degree of compliance with GDPR requirements, allowing for the interpretation and construction of effective communication and transfer of results between various domains, roles both technical and legal. However, it is worth considering that while developing a reference model, RMDP does not provide specific direct answers or solutions in the context of compliance with GDPR provisions (for the specified subject area). To a large extent, RMDP advises and assists in the context of compliance with GDPR means and directs (indicates questions, factors) that must be addressed, taken into account and thought through regarding GDPR compliance.

5.9.2 Implications RMDP

In the following chapter of the dissertation, possible implications of RMDP will be shown, including an assessment of the compliance of the stated requirements when developing a reference model. In the RMDP developing flow, the interdisciplinary domain of data protection was considered from various perspectives in a broad sense. It was not limited to just a technical or legal perspective. We acknowledge that an attempt has been made to interpret, understand and cover the domain under study from a variety of perspectives, including the opinions of experts and professionals, including academics, in this interdisciplinary field. We also accept and argue that the achieved results enable us to predict the utilization of RMDP in various areas. We comment on the implications of RMDP for areas such as education, research and industry below:

Research:

- structured privacy/compliance knowledge drawn from various academic and industrial knowledge;
- can be used as a template for cataloging privacy/compliance gaps or used as mapping to identify open questions in a research subject area;
- can serve as a template or foundation to complement various academic literature and works to identify open gaps and questions;
- can serve as a concept for developing crosswalks document between ISO/IEC and GDPR norms;
- can be used as a template for cataloging privacy/compliance issues and mapping results among various PETs;

Education:

- can be used as an abstract framework for teaching privacy/compliance for various roles, groups of students or areas (has the ability to complement and focus on certain concepts) to develop a more substantive course or program;

- provides a basis for studying issues and compliance in the field of personal data protection;
- due to the abstractness of the model, they allow and provide the opportunity to see the “big picture”;
- can be used for substantive communication for audiences with different levels of expertise;

Industry:

- provides a universal, privacy and compliance-oriented view on the implementation of existing compliance practices in the organization;
- allows you to catalog and identify the weakest or most vulnerable points in the context of compliance;
- improves communication and interpretation of results for the various roles involved in the assessments and compliance process;

5.10 Summary

One of the aspects and main contributions of this work, as noted earlier, is the development of a reference model. We argue that in this chapter we have presented the main artifact to meet the objectives of this dissertation. The problems addressed in this dissertation represent an interdisciplinary study aimed at solving the assigned problems. In other words, to achieve the objectives of this dissertation, the first step was to establish a single platform and systematize the concepts and interpretations used by the various roles to achieve GDPR compliance goals. The RMDP model, a practical solution, captures the entities and represents the domain understanding for data protection and privacy in the context of the GDPR (as presented in various academic and industry sources). We believe that this model can be used not only to harmonize different roles but also the team as a whole to achieve the necessary balance while complying with GDPR requirements, but also as a platform to provide the necessary modeling, monitoring and understanding of the PbD and data protection domain. We also believe that the developed RMDP model is prevailed at some level in the process of model development and design, since it includes domain knowledge and research results presented during development and reasoning. The novelty of RMDP can be interpreted by an alternative and unique view and combination of discrete knowledge (in an interdisciplinary topic) in order to combine the necessary knowledge and ideas of the general (holistic) picture of the field under study. In the next chapter, the RMDP model is assessed based on qualitative criteria related to the conceptual model, including issues of understanding of data protection, privacy and GDPR by various experts, practitioners and researchers. Among other things, we evaluate the pragmatism of RMDP in the next chapter.

6 Reference Model Evaluation

This chapter presents the results of the validation of the primary artefact and followed research activity in this dissertation, or in other words, the developed reference model RMDP. Considering Design Science Research (DSR) guidelines [83], one of the main goals is to evaluate artefacts that solve previously identified gaps or issues. DSR plays a crucial role in this process, serving as a methodology that allows for the creation and validation of artefacts that address a specific problem in the studied domain. The chapter begins with the presentation of the multi-phase evaluation methodology and the confirmation of the chosen evolution criteria in **Section 6.1**. This is followed by the qualitative results obtained from in-depth interviews with various experts and users of RMDP (**Section 6.2**) and the outcomes of a semi-structured survey (**Section 6.3**). An analytical evaluation and evolution of RMDP, based on the identified gaps, is then conducted in **Section 6.4**, with a summary provided in the final **Section 6.5**. The chapter concludes with a comprehensive discussion of the evaluation results and the evolution of the RMDP.

6.1 Evaluation approach

Evaluation of a conceptual or reference model is a complex and challenging process due to the lack of a rigorous framework or approach for evaluation [130] including the difficulties associated with issues of practicality validation [170]. The use of various methods for evaluating reference models remains, in most cases, subjective and is also challenging to formalize [130]. In addition to the above difficulties, one can add the variety of requirements and all sorts of restrictions applied to the area under study [60]. Despite the difficulties and complexity mentioned above, evaluating the developed reference model is possible analytically and empirically.

We also want to note that the possibility of assessing RMDP as an artefact in the context of real-world cases, we consider impractical and, at the same time, challenging to implement. Since the process of coordination and implementation projects, compliance with the GDPR requirements in various organizations can be calculated for relatively long periods of time (months or years) and affect the entire organization (since the developed reference model includes a wide range of both requirements and knowledge, thus the range of problems that RMDP solves is quite extensive). Thus, in order to minimize the various consequences and limitations presented earlier, the assessment of the RMDP model should be a multi-phase approach. This will enable the demonstration of the advantages and disadvantages of the developed RMDP reference model. In addition, the choice and use of the assessment method have their own limitations, but we believe that using a multi-phase approach (a combination of different methods) will avoid the limitations of each of the methods. In addition, we want to note the advantages of using a multi-phase approach is intended to confirm the merits among academia and industry, including the pragmatism of RMDP:

- Scientific value (solid and profound theoretical basis, interviews with various industry and academic experts, critical analysis by the developer and designer of the reference model);
- Practical relevance (how the stated requirements and recommendations were completed) has been fulfilled;

While there are many diverse strategies for evaluating DSR artefacts [199], we believe that the chosen strategy “a) formal evaluation, including prepared and open questions for surveys including the analysis at which point design requirements are met and b) comparison with similar/competing artifacts;” in the context of this thesis and artefact evaluation is consistent with the various strategies for evaluating DSR artefacts

and is also consistent with the objectives and available resources given the design of the artefact presented in **Chapter 5**. Moreover, further evaluation ensures that the RMDP fulfils all the initial requirements presented in **Chapter 4**, including being superior to similar approaches and addressing the shortcomings of the approaches presented in **Chapter 3**.

6.1.1 Evaluation Criteria

As we noted earlier, there are a variety of approaches, methodologies, strategies and proposals (both in industry and academia) for evaluating conceptual and reference models, including their quality. Moreover, as a result, different approaches and methodologies have their own advantages and disadvantages [130]. However, despite the multiple diversity of approaches, as the authors point out [130, 60], various approaches support and agree on the use of a multi-criteria approach for evaluating reference models. In work [131] six different criteria are proposed, which include such concepts as *Simplicity, Understandability, Flexibility, Completeness, Integration and Implementability*. The author [60] proposed a framework that includes four different perspectives for evaluation, which include: *Economic perspective, Deployment perspective, Engineering Perspective and Epistemological perspective*. The authors [160, 15] argue and propose that three main guidelines (correctness, relevance and economic efficiency) and three additional ones (clarity, comparability, systematic design) may be sufficient for evaluating models.

Taking into account that the evaluation of the reference model is a complex task and does not have a clearly defined approach [130], therefore, in accordance with [60], we adapt various evaluation methods (multi-criteria approach) and adapt the criteria presented by the authors [94], since they allow minimizing the limitations in evaluating the reference model and allow a more objective evaluation of RMDP from different sides. As we noted earlier, for the evaluation of RMDP and the selection of criteria, we adapted the criteria (*Simplicity, Accuracy, Scope, Systematic Power, Explanatory Power, Reliability, Validity and Fruitfulness*) for the conceptual model [94], which are applicable also for reference models. At the same time, the specified criteria [94] can also be adapted for the IS domain [6]. The list and comments for the adapted criteria are presented below.

- **Simplicity**: The model and elements incorporated should be simple and also comparable to other models;
- **Accuracy**: Model as concepts included (information taxonomy, data protection goals, privacy countermeasures, privacy development life cycle) aimed to be clear and explicit;
- **Scope**: Model should cover a) the broad scope of the domains represented and b) not overlook the essential concepts/elements; We rename the criteria from Scope to Completeness;
- **Systematic Power**: Model, concepts and definitions incorporated should be easy to grasp, interpreted with simplified relationships and meaningful way;
- **Explanatory Power**: Model should enhance the understanding and forecasting of phenomenon under study;
- **Validity**: Model should enable and provide a clear understanding, representation of findings applicable to various relevant situations;
- **Reliability**: Model should lead to the simirial understanding, observations while utilization to the similar situations by different model actors/performers;
- **Fruitfulness**: Model could pose a relevant research, practical gaps or hypothesis for further validation;

The RMDP in the context of the adopted recommendations and guidelines for RM was analyzed analytically by the model developer and designer (**Section 6.4**), including assessment by various experts and

Date	Venue	Duration	Audience	Participants	Responses
Q2 2023	The Analytics Summit	0.7 hour (presentation)	Practitioners	75	6
Q4 2023	Hokkaido University, Sapporo Privacy Workshop, DP	1 hour (presentation)	Academics, Practitioners	20	12
Q1 2024	Technical University of Munich, Security Research Group I20	1,5 hour (presentation)	Academics	7	5
Q3 2024	University of Tokyo, Momentum Lab	1,5 hour (presentation)	Academics	7	3
Q3 2024	International Workshop on Security, ICC Kyoto	1,5 hour (presentation)	Academics, Practitioners	21	15
Q3 2024	Tokyo Institute of Technology (Tokyo Tech), TDSAI Lab	1 hour (presentation)	Academics	7	6
Q3 2024	Organization	1 hour (presentation)	Practitioners	10	6

Table 6.1 RMDP development evaluation presentations

Note: Table 6.1 provides an overview and summarises the details of the presentation, an iterative process that was performed before running the primary evaluation process of RMDP. The presentation included the development of the reference model, further clarifications, and the concepts utilized. This was followed by discussions, critical observations, and feedback, all of which were crucial for considering diverse aspects and initiating further improvements.

practitioners in the research domain using different approaches (semi-structured interviews and surveys, **Sections 6.2**, **6.3** respectively). In the following sections, we present and discuss the specified steps for assessing the RMDP. We selected eight adapted criteria [94] because we believe that they satisfy the following considerations:

Purpose: The criteria presented can be directly applied to the evaluation of both conceptual and reference models, including those that are relevant to the area of study;

Applications: The proposed criteria represent an interdisciplinary research area for the evaluation of the reference model, as they do not reflect only theoretical considerations and are more widespread in academia than others considered earlier (**Section 6.1.1**);

Completeness: The number of criteria is more multifaceted in comparison with other criteria or perspectives, and thus, we are able to obtain a comprehensive and clearer picture of the study, including more diverse comments from experts (and more thoroughly analyze the disadvantages and advantages);

Further, this section provides details of the additional steps undertaken by the author prior to the primary evaluation cycle, including the interview process as recommended by the authors [133]. Also, **Table 6.1** provides an overview and summarises the details of the presentation, which was performed before running the primary evaluation process of RMDP. The results from the referred presentations significantly enhance the quality of the reference model and substantially strengthen its practical and research contributions. Further, the developed question list (**Appendix C**), presents the final list of questions used for the semi-structured interview and includes sixteen questions. The developed questions were based on the Pilot Test Questionnaire, which was conducted with various academics, PhD students, and lecturers (specializing in Information Security and Privacy, Data Protection at TUM, School of Computation, Information and Technology (CIT)) in early 2024. The group was familiarized with the topic and objectives of the study, including the study and the artefact being developed (the RMDP reference model). Based on the results of the pilot study and the feedback received, the questions were revised to provide greater clarity and precision; including adding explanations and clarifications to the questions and changing their tone and sequence).

6.2 Qualitative evaluation

One of the purposes of qualitative evaluation is flexibility and the ability to respond to the context; sometimes, these goals and qualitative research can be called emergent [59]; in other words, research questions develop depending on various data (knowledge), the analysis conducted, and the form of a respective study. In particular, the goals of the study and research questions should be clearly formulated [187].

6.2.1 Approach

For the qualitative assessment, we conducted a comprehensive series of almost twenty interviews with various experts from different domains in the research domain (data protection, software development, privacy engineering, and privacy legal practices), as demonstrated in **Table 6.2** partly following [86], considering [2] and instructed by the following thematic analysis [23]. The interviews were conducted in 2024 (Q3) via (video) Zoom, Google Meet and Microsoft Teams conference calls. Seven respondents were recorded and transcribed; other respondents were documented with notes due to their unwillingness to be recorded. To ensure the anonymity and confidentiality of the participants, each participant was identified by a unique numerical code (e.g. P3). The interviews were organized as follows: First, we presented the motivation for creating the RMDP reference model, including the approaches and concepts used to develop the RMDP. Next, we presented a visual representation of the RMDP, which was accompanied by an explanation for each dimension or part of the model; then we supplemented the explanation with a step-by-step application of the model and the resulting privacy document; in some cases, we commented and added synthetic examples for a more illustrative example. The interviews concluded with prepared open-ended questions and sections presented below. The process continued till sufficient data was collected, and we also stopped conducting interviews at the time we started receiving similar or identical responses from the participants.

Background: Your industry? Work experience? Privacy or data protection experience?

Correctness: Does the presented model match your experience? Is the presented model complete and covers all relevant aspects of GDPR compliance? Model has a clear structure for enhancing and monitoring GDPR compliance (concepts incorporated and related design arrears)?

Applicability: Is it simple to apply and communicate model and model results? Is the model useful for the organizations? Is it easy to integrate knowledge by utilizing the RMDP model? Could you imagine RMDP utilization in your daily tasks/jobs? Could you imagine RMDP utilization in diverse tasks with DP or other PET?

Simplicity, Understandability: Is the model easy to understand? Does the model have an appropriate level of detail? Does the model have an appropriate level of design?

Completeness: Do you feel that RMDP is feasible and complete for executing and monitoring GDPR compliance? Could you say that RMDP and detail level are enough to address the compliance problem(s)?

Interpretation: Would you say if a model is robust enough to reflect changes in the GDPR compliance domain?

Representation: Are questions regarding GDPR compliance addressed adequately? Is the model representation clear?

Related approaches: Are you familiar with similar models which aim to address the same questions?

ID	Role	Experience	Industry	Duration	Organization size
P1	Senior Software developer*	10 years	Software Development	67 min.	Large
P2	Senior Software developer*	11 years	Software Development	63 min.	Medium
P3	Senior Privacy Researcher and Developer*	8+ years	IT Services and Research	46 min.	Large
P4	Data Protection expert*	4+ years	Consulting	59 min.	Medium
P5	Data Protection officer*	8+ years	Software Development	47 min.	Large
P6	Senior Privacy Researcher and Developer*	7+ years	Telecommunications	55 min.	Large
P7	Software developer	6 years	Software Development	57 min.	Medium
P8	Data Privacy counsel*	7+ years	Software Development	56 min.	Large
P9	Head of Data Protection*	13+ years	Production	45 min.	Large
P10	Senior Privacy engineer*	7 years	IT Services and Consulting	46 min.	Small
P11	Legal Privacy Expert	3 years	IT Services and Research	57 min.	Large
P12	Data Protection officer	5+ years	Production	56 min.	Large
P13	Senior Privacy engineer	4 years	Software Development	58 min.	Large
P14	Data Protection officer*	7 years	Software Development	44 min.	Medium
P15	Privacy engineer*	3+ years	IT Services	55 min.	Large
P16	Legal counsel*	13 years	Law practice	56 min.	Medium
P17	Data Privacy counsel*	4 years	Law practice	61 min.	Medium
P18	Data Protection officer*	2 years	IT Services and Research	59 min.	Small
P19	Data Protection officer*	5 years	IT Services	39 min.	Large

Note: * have additional data protection or privacy certification e.g. International Association of Privacy Professionals (IAPP)[a]; Organization size[b] explained as follows: *Small*: 10-49 employees, *Medium*: 50-249 employees and *Large*: 250 or more employees;

[a]https://iapp.org/

[b]https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Enterprise$_{s}ize$

Table 6.2 Overview of study participants interview

Further remarks and suggestions: Could you share your remarks, suggestions and other comments regarding the RMDP reference model?

6.2.2 Results

The interview was conducted in the English language, and on average, the interview took about 54 minutes; in case of consent from the respondents, the interview was recorded and conducted at no cost (participation was voluntary). The interviewers, who all showed significant interest in the presented artefact, expressed considerable interest in participating in the proposed interview. Taking into account the anonymous nature of the interview, personal data, including affiliation with companies, as well as the age and gender of the interviewer, are excluded.

Table 6.2 details the in-depth interviews conducted with various experts (data protection expert, software developer, privacy engineer, and privacy legal expert), highlighting and illustrating their differences, including related information regarding their experience, industry and organization size.

Table 6.3 presents summarised feedback from interviews regarding the categories selected. Firstly, we provide the main statement and a condensed overview with results considering each group.

P5: "*I have seen so many models, but they are only parts of RMDP... I don't think I have seen something that put them together like this... every part of RMDP feels like you have already dealt with it before .... the thing I have seen here the way you put it together it makes it more compelling ...*"

The results indicated that the shown categories of interview participants and their roles in the organization play a considerable role in achieving the required level of compliance with the GDPR requirements and in the context of PET utilization. Organizations are experiencing an urgent necessity for the offered reference model RMDP and presently face multiple difficulties in adapting existing academic or industrial approaches to evaluate the degree of GDPR compliance (particularly the estimating anonymization level, matching threats and followed countermeasures). Since none of the available methods or approaches fit specific use cases or organizational needs, refereed methods need much effort to be adopted and modi-

	Software developer, P[1,2,3,6,7]	Data protection expert, P[4,5,12,14,18,19]	Privacy legal expert, P[8,9,11,16,17]	Privacy engineer, P[10,13,15]
Correctness	selected abstraction level offers accurate knowledge	demonstrates sufficient correctness	represents specified domain with minor comments	feasible, match practical experience
Applicability	applicable for diverse organizations	usable, but depends on the size/ resources of the organization	template is usable, missing a detailed overview of specific data elements	agree, PDLC and DPG were elaborated
Simplicity, understandability	easy and adequate, practical to the diverse expertise levels	can easily utilize demosntrated model	no blind issues, missing context for specific tasks	holistic and structured, PDLC should be clarified
Completeness	represents all the critical concepts	could be utilized as a systematic method/ advance audit	information taxonomy could be extended, lawfulness goal demanded	clarification for DPG relation to the Countermeasures needed
Interpretation	no blind spots	robust and reflects future changes	no concerns, discussed PDLC DPG should be checked against SDM	agree, Information taxonomy should be clarified
Representation	practical, lack of an example makes it difficult to adopt	likely clearly represented	usable, data transfers could be explained	functional, PDLC shoud be elaborated against security
Additional remarks	overlooking the checklist for utilizing the model/ privacy document	Countermeasures categories need more details against other models	performance depends on the organization's resources	crosswalk to diverse PETs could be demosntated

Note: **P** in the group name belongs and relates to the ID column, *Table 6.2*

Table 6.3 Summary of interview results considering four groups

fied. Simultaneously, the results demonstrate that organizations are experiencing difficulties in translating the GDPR requirements for various stakeholders, and RMDP utilization can enable organizations and diverse roles to combine and optimize the process of communication and translation of results between diverse actors (categories and roles).

Results indicate that each of the demonstrated groups (and referred roles) highlights the importance and consistency of the proposed RMDP model[1].

Considering analyzed results, RMDP enables and delivers the following steps to fit GDPR requirements while utilizing PET: consistently and repeatedly follow and supplement the process of fitting the GDPR requirements, including interpreting the specific requirements utilizing the DPGs. Moreover, RMDP enable and delivers:

- flexibility in adapting the model (the capacity to supplement and modify the presented quadrants, including supplementing or revising the suggested concepts, taking into account the various tasks and sizes of organizations);
- completeness (the ability to interpret the results of various personal data protection processes in a single framework and using a taxonomy that is understandable and accessible for interpretation);
- easy start and monitoring of results (easy of use of the RMDP model due to the not overloaded level of detail and the absence of complications in exploring outcomes from the model);

Additionally, the suggested Privacy document delivers an alternative to record and standardize: a) the implementation results from the model for every phase of PDLC, b) the accessibility and knowledge of each PDLC phase for each role interested, c) the capability to interpret the model with similar approaches or easy of transformation in the context of varying legislative requirements or technical aspects.

Further, we demonstrate detailed results utilized for **Table 6.3** according to the categories and groups presented, namely Software developers, Data protection experts, Privacy legal experts and Privacy engineer, with corresponding criteria given in **Section 6.2.1** and additional remarks. We also want to know that only selective and applicable opinions from interview participants for the current study are presented due to the purpose of excluding similar by content or intentional comments received. Nevertheless, all accumulated feedback from participants *P1-19* is aggregated in **Table 6.3**, as explained earlier.

Software developer

This section presents results from discussion from software developers, P1, P3 and P6 are working in very large enterprises; P2 and P7 as researcher and consultant in medium large enterprises.

[1] P5: "*I have seen so many models, but they are only parts of RMDP... I don't think I have seen something that put them together like this... every part of RMDP feels like you have already dealt with it before .... the thing I have seen here the way you put it together it makes it more compelling ...*"

- **Correctness**:

For the level of abstraction, P1-P3, the presented concepts, structure and elements are reflected acceptably; however, as noted by P1 and P7 the possibility of high-level abstraction can be supplemented with specific examples for various types of organizations, including Countermeasures (for organizational and technical categories). P3 especially emphasized that the basis and base for monitoring and enhancing GDPR in the presented area for the dimension Information Taxonomy and DPG corresponds without any comments (note: for the selected level of abstraction).

- **Applicability**:

The majority affirmatively confirmed that the presented model can be applied in the organization. P1 noted that in this graphical version, this model and the proposed level of abstraction can be used for various organizations and tasks, including expanded to other PETs, including SMPC (P1), but it implies that there will be a need to supplement the concepts for the four dimensions presented. P2 suggested supplementing the privacy document with possible crosswalks to interpret the results in the context of utilizing alternative approaches. P1 and P3 also commented that to improve the model's utilization, it is necessary to add various use cases, including examples that will allow the filling out of the privacy document more organically and confidently. In other words, P1 and P3 proposed to prepare step-by-step instructions for various PETs DP, FL. P3 also noted that the inherent flexibility of the model allows integrating knowledge more effectively, so it can easily change or supplement the concepts presented; at the same time, P3 recognized that indicating the priority of the DPG is extremely important, as well as suitable for use for daily tasks or work routines with various additions.

P3: "I can imagine using it* (RMDP) (in daily/job tasks) and to be honest probably I would adapt it to my needs but I think it could be useful."

- **Simplicity and understandability**:

P1 and P2 agreed that the approach chosen and the understandability of the presented model are adequate; P2 also noted that he imagines that this model can be used to interpret the results for other roles in an optimal form, and thus, this model can be a platform for decision-making and discussions. P3 noted that additional visualization of the relationship between the dimensions (arrows) can be presented in a more simplified way, or additional descriptions of what each of the arrows means or explanations can be added to the visual representation and the representation of the reference model. At the same time, P3 noted that the visual representation of the model is very self-governing and also very intuitively shown. P3, P7 also noted that the importance of the intersection of the DPG is correctly represented and entirely logical.

- **Completeness**:

P1 noted that the abstractness of the model represents and shows all the necessary concepts for tracking the level of compliance. At the same time, P1 and P2 noted that selected level of abstraction is suitable for large and medium-sized organizations. P3 noted that the selected level of abstraction of the concepts presented in the model serves as a starting point or basis for execution and monitoring GDPR compliance for various organizations, which are subsequently not limited and can supplement or change each quadrant's level of detail and content. P3, P7 noted the possibility of complexity in mapping DPG in the context of DP utilization.

- **Interpretation**:

Respondents believe that the interpretation of the model is adequate, as it uses a structured approach and simple and straightforward graphical approaches and elements. P3 highlights that every dimension or concept incorporated can be easily adapted or adjusted.

- **Representation**:

All interviewed participants agreed that the presented model reflects the necessary GDPR requirements at the presented abstract level. Although P1 noted that although the model includes the necessary dimensions, the lack of an example makes it difficult to use or adapt. P3 noted that requirements such as data portability, data transfers, or specific use cases of joint data controllers may arise in the interpretation, concluding that the model has limited application if we consider all current and possible GDPR requirements.

- **Related approaches**:

P1 and P2 noted that they have not encountered similar models. However, they observe similarities in some elements with approaches in the context of certification of data protection experts of the international organization IAPP (aimed at training qualified specialists to assess various GDPR compliance issues). P3 emphasized the similar and closest approach of PRIPARE. However, PRIPARE is intended for other purposes and in the context of the research objectives of the current dissertation, it is not easy to imagine that PRIPARE can be adapted and utilized.

- **Additional remarks**:

All respondents, except P7, noted that the capability to evaluate that something needs to be added (concepts or further categories), considering only the presented model (without modifying and utilizing it in the organization), is quite comprehensive. Experts noted that they would like to get acquainted with a demo of using the given model in organizations of different sizes and products to compare the significance of the model.

Data protection expert

This section presents results from in-depth discussions with data protection experts; P4, P5, P12, P14, P18, P19 are working in medium and large enterprises. During our presentation, we ensure comprehensive coverage of the interconnection to the data protection goals, selected information taxonomy and communication of the knowledge from the resulting privacy document. We also debate the interrelation and interpretation of results obtained between various roles in one organisation and the adaptation results to the DPIA.

- **Correctness**:

P4 and P5 noted that RMDP acknowledges GDPR requirements, including the integrated concepts. P4 also highlighted the importance and feasibility of original privacy by design and by default principles in the second phase of the presented PDLC, which aligns significantly with the GDPR requirements. They also noted the potential of using the presented model as a systematic method, emphasizing that the accuracy of GDPR compliance is not limited by the ability of organizations to complement or modernize the suggested RMDP concepts. P4 further noted the opportunity to expand the context for legal examinations in the second phase of PDLC for the DPG quadrant, enabling more reasonable interpretation between various parties involved (various stakeholders). Thus, experts see a significant potential for RMDP model to handle the increased volume of requests from data subjects to organizations over time, as well as the responsibilities of organizations to fulfil and monitor the data request, instilling confidence in the model's effectiveness.

P4: "*So, I think this presented reference model includes accountable perspective that might very helpful for companies or institutions to use this current model to apply not only for the GDPR but also for other data protection regulations in the same country*"

P5 concluded that to find a gap in the demonstrated model, P5 would need to spend a lot of time applying RMDP multiple times to diverse use cases to find specific gaps. However, P5 admitted that the delivered level of abstractions could possibly decrease the chance of finding a particular gap, and considering P5's experience with other models, RMDP has much promise.

P5: "*I would say the model exceeds my experience...I have a broad understanding of it, and the model seems very intuitive and makes a lot of sense; however, I have a small challenge regarding the data protection goals and mapping them against various international legal frameworks*"

P5 also comments that the chosen definition, taxonomy, and design strongly perform and match the security aspects and deliver synergy when communicating results between various experts from the RMDP implementation. Lastly, P5 comments that the Countermeasures dimensions and categories introduced are depicted as a stand-alone entity. However, experts prefer to represent them in another way since the referred blocks have (possibly) many interconnections with each other.

- **Applicability**:

P4 noted that as an expert sees this model as a possibility for organizations to demonstrate the conduct of DPIA and disclose the results in the context of increased oversight from data controllers. Together with P5, they agreed that this model is a reasonable tool for various data protection activities, including being applicable for PET DP. Still, it relies on the size and resources of the organization. P4 noted that he would use this model in everyday tasks only after specifying each element presented in the reference model in his organization. The expert also noted the possibility of improving the interpretation of results and communications between different roles (for the first phase of PDLC) and suggested defining specific operations or cross-checkbox fields for various processes.

P4: "*In case for companies who has not have enough specific resources for audit their data protection level so that might be not useful and depends on company size and depends on the corporate resources but in total I suppose it's very useful.*"

P5 indicates significant interest in trying the presented model in an ongoing project for his current organization, which is immediately applicable even without adoption for specific use cases in the organization. In addition, RMDP seems to be orientated toward various levels of expertise, considering the diverse stakeholders in the organization and also delivering a clear interpretation for various external data processors to which degree the anonymization level should be adjusted. Additionally, it highlights the importance of the first phase of PDLC, which in some industry models or approaches has less attention. Moreover, P5 underline the applicability of RMDP in terms of geo-location data and DP utilization.

P5: "*Particular and looked in it, and one of the main questions which are most challenging interfaces or the most I deal with (interface between different roles business, privacy and technology, particular software developers)... they kind of speak different languages and you need kind of this model to help people. I think this model is technically enough to be understood by a developer, and it is privacy-focused to be understood by privacy professionals that the way I would see this reference model. The immediate thing I look into this is the model I would use as an interface to more technology and privacy-driven scenarios.*"

P5 raised a question regarding the need for additional loops during the PDLC phases, which could be discussable when no information transformations happen, in particular, changes in the privacy document

or affect the Data Protection Goals' criticality level.

- **Simplicity, understandability**:

P4 and P5 commented that the various actors in RMDP execution who are familiar with the (privacy) audit process can efficiently operate this model. However, additional knowledge or comments for the arrows depicted between the quadrants may be required to interpret the outcomes from the privacy engineer roles. P4 also noted that the components' design or simplicity is optimal for this model's level of abstraction. P5 stated that the repeated cycles in the cycles could be unclear and interpreted differently; however, it would not be an issue for the domain experts.

- **Completeness**:

P4: "*I think the model is very comprehensive and enough for basic compliance requirements and its parts*"

P4 commented that expanding the information taxonomy category attributes would enable more accurate model service for diverse roles or stakeholders. However, the expert specified that this would lead to degradation of the model utility and would problematize the use of the model for the selected level of abstraction. Participant himself suggested adding the "other" category, thereby allowing organizations to be more accessible in the context of the complexity of determining the category of information.

P4: "*I'm a bit concerned or sceptical about the definition of human a bit broad context in countermeasures quadrant I would consider maybe more suitable name*"

P5: "*I have a feeling, and I would say it is too much in some areas, rather than too little, then some blocks could be missing.... It (the RMDP model) a little too detailed if it could be cut a bit further down in some parts, but the completeness is evaluated*"

- **Interpretation**:

P4 claims that the model is fully robust and can anticipate and reflect future changes in the GDPR legislation process. Experts also acknowledge that the model can be adapted to new provisions or laws, such as the EU AI Act. However, given the EU AI Act's interpretation of integrated components and concepts, including visual representation, this is a significant challenge. Experts have suggested a long-term vision for the model, which includes a separate dimension that allows for the addition or combination of other legislation, such as the EU Act, Digital Act, or others.

P5: "*I think this is actually the strongest point of the (RMDP) model; the abstraction layer is relatively high... I would not call this a GDPR model, I would call it a privacy data protection model and with a very little effort, the model could be adapted even not only for GDPR changes*"

- **Representation**:

Both experts noted that the RMDP model is precisely consistent with the current presentation of the GDPR requirements and commented that the model could broadly harmonise to future changes to the GDPR due to its flexibility and adaptability. However, P5 noted that some issues, such as data transfer, could be problematic to intercommunicate in the offered model. In addition, P5 mentioned the DP or other PET implementation could limit the model. However, the model provides enough flexibility for improvement.

-**Related approaches**:

P4 stated that they (the current organization) have yet to come across similar models in the industry with such a comprehensive approach to data lifecycle. However, the expert pointed out from his academic expertise and assessment (various academia venues related to the Data Protection and Privacy tracks) that the presented attributes in Information taxonomy and Countermeasures attributes characterize RMDP as more extensive (from an engineering perspective) when compared with similar approaches. P5 stated the similarity of the lifecycles presented in the model and noted that the established design approach and the execution logic correspond to industry standards. Moreover, the uniqueness of the RMDP enables data protection experts to evaluate the issues of data protection and privacy differently (separately or together). Lastly, experts admit that the model contains or is incorporated partly by definition from various industry standards, thereby simplifying and minimizing the amount of labour in exhaustive use cases.

P5: "*I have seen so many models, but they are only parts of RMDP... I don't think I have seen something that put them together like this... every part of RMDP feels like you have already dealt with it before .... the thing I have seen here the way you put it together it makes it more compelling ...*"

Additional remarks:

P4 noted that data subject requests are presently crucial, as the data protection authorities stressed. Suppose offered concepts could be integrated, or the proposed concepts could perform on reducing the labour regarding data subjects' requests and lawfulness. The offered model will be more applicable and highly utilized in that case. P5 admit that the demonstrated model (RMDP) fits and intuitively delivers a relation to various interconnected expertise, such as security, in a broad sense. Experts also point out that RMDP has to be presented in specific privacy and data protection regular meetings organized by IAPP in the EU (primarily focusing on privacy and data protection). P5 raised the main concerns regarding categorizing the Countermeasures quadrant. Generally speaking, experts would like to see referred categorization, not standalone, while at some point, then having them collapse or having an intersection. In other words, the categories' absoluteness could be adjusted by considering various international regulations in addition to the GDPR or simply including a category such as "other". In addition to other international standards, category others (Information Taxonomy dimension) could leave the options or interface to integrate similar approaches (developed earlier or already utilized by the organizations) in RMDP.

Legal privacy experts

This section presents results from discussions with diverse legal privacy experts; P8, P9, P11, P16 and P17 perform duties in large and medium enterprises (EU), which also outsource the referred expertise for various privacy technology vendors and third parties over the EU. In addition to our general presentation, we also incorporated an additional slide that is particularly important, as it focuses on DPG with mapping referred DPG and GDPR articles and recitals presented in **Table 2.2**. An additional presentation slide was supplied with additional explanation, interconnection, and feasibility with Information Taxonomy and Countermeasures Dimensions.

- **Correctness**:

In addition to other experts, privacy legal counsel and head of data protection experts (P8 and P9) commented that the referred level of abstraction in the presented model is accurate and that it responds to various tasks or use cases for diverse organizations. P9 expert also commented that the overall logic, concepts, and definitions introduced are utilized in the same way in the current organization in the data protection domain imposed by the legislation notation. Additionally, P9 experts highlight the importance of Privacy document (RMDP), which partially follow the process in the current organization, neglecting the specific (use case, PII) details.

P8: "*Ideally, what you presented, I think in an ideal world it should work, and that works in the same way on my previous work experience*".

Additionally, a P8 expert raised concerns regarding the offered privacy document. If the company already has a comparable document, how could the knowledge be transferred to the drafted new RMDP privacy document? Further, P8 show the importance that in some use cases, it is important to take a step back and might be while performing the reference model, the actor could perform one quadrant more than one time to clarify the complexity; however, P8 agree that iterative process could solve such and issue and performing one again the fool loop will provide more clear and precise information for fulfilling the drafted privacy document. P8 and P9 highlight the importance of the DPG and Countermeasures quadrants, which tackle the studied area and cover relevant aspects. Nevertheless, the P9 expert noted that it is not expected to assign DPG goals. Instead, they assign risks in the current workflow (however, the taxonomy is not defined, while it could be named differently, e.g., risks or controls). Referred quadrants could also be adjusted to the specific use cases for diverse organizations. Respondents also mentioned that the two referred below are highly interconnected with the DPIA procedure and assume the most likely where the DPIA happened somehow.

P8: "*I would definitely say that the presented model has a clear structure for enhancing and monitoring GDPR compliance*".

P8 and P9 experts highlight the importance of the evaluation dimensions and PDLC phase for privacy management and monitoring and highlight possible extension descriptions of each of the specified phases. P9 noted that the Countermeasures dimension and Legal category play a significant role in fulfilling current duties, among others offered. Lastly, the experts would like to have a listed Countermeasures catalogue for the specific use cases and also some lawfulness explanation of the purpose related to the privacy assessment (for example, introduced in the offered Privacy Document or before/during the PLDC first phase). Both experts commented that the model and the execution logic need attention. However, it takes a minimal amount of time to grasp the knowledge.

- **Applicability**:

P8 points out that he would be more interested in assessing the specific data element or data asset in detail, which could be represented in the developed privacy documents. Nevertheless, experts acknowledge and agree that being a template and utilizing the presented reference model could give the flexibility to modify the privacy document regarding their needs. P9 expert remarked that the offered model could be extended either for a unit or project from an organizational perspective. However, expert would like to notice the interconnection of the offered model in such a use case. Lastly, experts demonstrate great interest in further methodology, explicit nomenclature, and intention of each concept offered, as well as an extension of a mode in an interactive way for developed model and/or Privacy document.

P8: "*I would say organizations at some point, and refer to my industry experience, are using a similar model or maybe in a different form or stages that are already in place because that will be the way how to confirm that you are in compliance with GDPR*".

Expert P8 mentioned that this happens many time in a daily task they have much bunch of data or related to the project, so picking up the piece of data provide an action plan for mitigating the privacy risk, however is most likely in addition it more feasible to see the big picture and he suggest it could be optional include the examples of the piece of the data or line in the privacy document, which definitely will simplify experts job tasks and optimize the daily routine (in other words decrease workload). Otherwise, it could decrease the countermeasure's actions and feasibility. In conclusion, P8 mentioned that the model could even satisfy his expectations for one data element.

P8: "*We have the same presented approach; however, for our organization's needs or while utilizing the RMDP, we need to break it down into more detailed specifications. . . the similar models from various providers already available on the market*".

P9: "*If the organization spent much time developing and adopting the ontology behind the model, it would not be efficient... This is why other models are more successful on the market.*".

Legal Privacy experts mentioned that depending on the organization structure, they would like to have additional options or interconnections to the security-related departments or interconnection between privacy and security domains; in addition, experts raised questions since the model is purely focused on GDPR, they would like to have a crosswalk with other national or international legislations. P9 expert noted the chosen abstraction level of the model clearly offers a big picture for various stakeholders involved. In addition, considering the logic outlined in the model, assigning roles for specific tasks or responsibilities is easy.

- **Simplicity, understandability**:

Experts agree that the presented model has a straightforward grasp approach and can be utilized by actors with various expertise levels. However, they raised the questions while adding more context and details for specified organizations. Additional information regarding the execution process could be needed. Moreover, excerpts agree on the referred level of abstraction; the presented model design is totally feasible and has a universal characteristic.

P8: "*This model and level abstraction is clear enough, but I can also imagine that you will go in more detail for the specific use cases*".

P9: "*On the conceptual level, this model is not very difficult to understand*".

P9 expert noted that the graphical representation could be mode digital and interactive, giving more enthusiasm for various stakeholders to interact with the model and communicate the results. In addition, experts point out the importance of framing four arrows, which attract more attention from the first catch.

- **Completeness**:

P8: "*In overall the presented model is sufficient to execute and monitor GDPR compliance*".

Experts raised the question of the lawfulness goal is not included in data protection goals, because at this point, the legal person (or legal expert) should perform dedicated tasks (assess if the legal basis for this is solid). In addition, expert would like any confirmation or flag for the PDLC phase monitoring, which highlights the lawfulness needed or basis to perform the privacy review assessment, possibly triggering the privacy legal lawyer's actions. Experts also highlight the importance of the retention deadline and would like to see it somewhere in a more detailed layer. The P9 expert also noted that considering various organisational tasks, the model could be equipped with extended granularity for concepts utilized and followed extended vocabulary.

- **Interpretation**:

Neither expert had specific concerns regarding the presented model and the incorporated concepts of abstraction. In addition, they agree that the chosen abstraction level or template of presented dimensions could serve as an option to reflect adequately any changes that could happen to GDPR. However, P9 expert also noted that providing more information or extended layers or levels of details will describe the

model less robustly, considering possible changes in GDPR legislation and notations.

P8: "*If you get a new requirement or any update (from authorities) you will just place it in one of those stages (depicted in the reference model) or just create a new phase, and it then it will be flexible and reflect GDPR changes . . . and the RMDP is easily updatable in way that you can add or modify the concepts introduced*".

- **Representation**:

P8 and P9 agree that the presented model is most likely clearly represented. Nevertheless, P8 mentioned he would like to have an additional explanation for the model itself, incorporating a bit more background or either a specialized guidance document like a walkthrough in an easy way. Hoverwer practically, if the model is part of a written process or process in an organization, it could have more added details, which helps every actor involved to have the same understanding of what is required.

P8: "*Sometimes, the lifecycle presented does not ram every time in the same way as it was designed, but this is more, I assume, related to the use case or organization needs*".

P8 was missing the lawfulness topic, which was addressed in the Completeness section. In addition, participant mentioned the place where experts can place the DPIA procedure for the presented lifecycle and that they would like to understand the risk level. Nevertheless, experts highlight the light DPIA or privacy assessment could be fulfilled by utilizing the correct two dimensions (Information Taxonomy and DPG Dimensions). However, in case the privacy risk is high, the legal expert should perform the DPIA and then again perform the steps for the same PDLC phase. Lastly, P9 expert noted that the model does not cover the whole scope of GDPR requirements (for example, organisation such as training) but perfectly addresses various data protection (e.g. DPIA) or privacy matters.

- **Related approaches**:

P8 is aware of similar models that organizations adopt for their needs; however, referred models are transformed and accommodated to the organization's needs. Considering RMDP and comparing similar models in use, expert highlighting for this level of abstraction, the current model is motslikey easy to use as a template. Nevertheless, in diverse cases, it is based on the use case. From the participant's personal experience in big organizations, the participant mentioned that each stage in referred models has more detailed sub-stages of up to seven or even more, and presented level abstraction could be utilized as a template for various use cases. However, in general, the ISO/IEC industry standards should be utilized. P9 noted that organizations are familiar with similar approaches introduced by NIST or ISO/IEC 27701 standards and operate depending on diverse organizational or project objectives.

Additional remarks:

All experts commented that adaptation and modification of the presented model depend on the organization's scale and how it operates; however, if the organization does not have enough capacity or resources, model implementation could not be efficient. Also, the P9 expert commented that privacy risk management and further countermeasures are more applicable than offering DPGs and criticality levels to current organizations. Nevertheless, the offered PDLC loop and general RMDP loop (loops in the loop) do not limit the adaptation of RMDP and incorporation results from additional Privacy risk management frameworks already incorporated in the organization. In addition, expert (P9) admit that utilization of PET DP is complex and incorporated mainly as an experiment since the estimation of anonymization level, as well as tuning and selecting proper DP form, is a complex and challenging part for which organizations require more resources. Lastly, the P9 expert noted that the organization does not have an efficient model such

as RMDP, which could act as a logical measure of how to perform privacy analysis between units, projects or platforms, considering the various stakeholders involved and offered DPGs.

Privacy engineers

This section presents results from a discussion with privacy engineers; P10,P13, and P15 are working in large and medium enterprises in the EU, respectively. In addition to our general presentation, we also incorporated an additional slide focusing on DPG and DP. This was supplied with further explanation, interconnection, and feasibility with Countermeasures, DPG and Information Taxonomy Dimensions.

- **Correctness**:

P10 states that the model covers mainly all relevant aspects considering the abstraction level; the proposed Privacy document plays one of the essential roles besides the model presented, which enables the interconnecting of various stakeholders. Additionally, participants pointed out and paid great attention to the DPG, which, generally, is not clearly presented, explained, and addressed in the day-to-day tasks. All experts admit that the presented model provides a better structure against unclear definitions for anonymization levels and how to quantify or design evaluation approaches specific to the use case.

P10: "*I do think you are covering the whole four phases (four phases Evaluation, Information Taxonomy, DPG and Countermeasures) and the loop within the loop. For me, it makes perfect sense*."

In addition, all experts point out that actors for RMDP without any knowledge would probably not be able to utilize the model. However, all admit that, after a short explanation, this model is easy to utilize in interstate. Nevertheless, P15 express that data protection goals have limited application and interconnection with specific NIST interpretation for data protection goals are not defined.

P10: "*It feels very, very complete as it is in presented form. Also, it would not be easy to understand without your presentation and explanation*."

- **Applicability**:

P10 states that the model is superior for building communication with various stakeholders involved in the process of privacy and data protection. Given that, the referred model can decrease complexity in the most common issues related to the lack of knowledge between the privacy counsel and privacy engineer, which reduces the end product value. All experts point out the importance of such an easy-grasp model, considering the clearly presented phases of the PDLC loop. Additionally, experts state that the model definitely serves many organizations; from experts' expertise, they would like to point out the importance of the various threats and interconnection to the DPG and followed Countermeasures. Also, experts admit that filling out presented abstractions for organization-specific needs could take time and additional effort. P10 and P15 agree with the applicability of the privacy document and claim that it could be one of the essential parts of communication results during the implementation of RMDP and also fulfil the GDPR recommendations, which could be translated into terms of DP application for various use cases. Further, the iterative process of the model advances the monitoring phase (including the management and Monitoring PDLC phase), which many organizations highly demand. P10 mentioned that RMDP could be limited in the application regarding the specific use cases in the domain of threat modelling. However, the model has some flexibility to join in the privacy flow to fulfil the task with other models or approaches (particularly for the Countermeasures categories and technical aspects of measuring or estimating various threats and responses to them).

- **Simplicity, understandability**:

Experts mentioned that the model without explanation could lead to misunderstanding; in addition, without justification, many cycles and not easily illustrated start of the reference model could be more apparent to the actor. In addition, experts stated that the RMDP model is carefully balanced in terms of tradeoffs between making the model more complex with various explanations and definitions and just step-by-step instructions. Both experts state that the mode has the right design level and is comprehensible.

P10: "*I think the reference model for the overall process has the right amount of details. Of course, if the company would like to implement the current model, it should adjust or solve many issues (specifically for the use case).*"

- **Completeness**:

The P10 expert noted that the model covers the primary knowledge applicable to the studied domain and GDPR requirements. In addition, as a practitioner (P10 expert) would also like to have more profound knowledge or an explanation of where the DPIA or PIA process and results are related or could be incorporated. Additionally, experts noted the amount of information registered in the presented privacy document related to the completion and progress of a model. Experts highlight that Privacy Document play a more significant role in considering the RMDP visualisation, which results in proper interpretations and fulfilling the data protection goals. However, P10 expert believe that it is not enough for vast organizations to reach the entire level of compliance with GDPR; considering various parameters for DP, an additional role of the data processor or legislative authority could be included in assessing the results of the model performed. In other words, no experts would like an additional perspective or transparency that could address the referred issues or have a clear crosswalk for interpreting the results for data controls.

- **Interpretation**:

All experts agree that the model is generic and modelled in a way that the RMDP reflects the possible changes in the GDPR.

P10: "*The fact that this is a loop and the model has a, let's say, a continuous reassessment... I do think so (the model fully reflects changes in GDPR privacy and legal domain), and I think about continuing this study, especially for the EU AI Act*."

- **Representation**:

Participants agreed at some point that the depicted abstraction-level questions regarding GDPR compliance could be adequately addressed and have an appropriate level of design. However, both experts (P10, P13) were pessimistic, stating that such representation could lead to a holistic view and confirm the question of implementation in various use cases. This could lead to a very high probability that the final product will be GDPR compliant since there is no available threshold (for estimating anonymization level) or particular method to re-validate the current statement. Nevertheless, the model provides a platform to bridge the expertise between various stakeholders involved and enable more transparent discussion of the various data protection issues from the GDPR perspective.

- **Related approaches**:

P10 expert noted that the dedicated entity performing the privacy and data protection task and the whole organization are familiar with various similar approaches and industry standards. However, due to the organization's needs, none of the available approaches or methods fulfil the organization's or end product needs. Thus, the organization adopts and modifies or sometimes joins various approaches to fit the organization's needs and reach the most reasonable result. P10 also notes that the presented RMDP is similar

and tackles similar phases which the organization is currently trying to develop. In addition, P15 expresses that NIST recommendations and frameworks are more likely to be more applicable than RMDP.

- **Additional remarks**:

The P10 expert suggested considering roles in the loop and incorporating them either into the model or into the demonstrated privacy document. In addition, a possible situation could occur to any organization when the PIA is outdated due to the fast pace of technological changes or additional dynamic factors. The model presented could mitigate referred issues due to the continuous loop presented. Nevertheless, P15 expects to have a more transparent structure regarding the different roles, involvement, and activities that are taken during various PDLC phases. Experts also noted that while utilizing the DP for the countermeasures dimension, the process of quantifying the DPG stated before could be complex and challenging to evaluate. In conclusion, experts noted that having such a loop or phases presented in the model is essential to have all stakeholders on the same page; moreover, quantifying the impact of various forms of DP for defined DPG is still a challenge. All experts point out and note that the process of fulfilling GDPR requirements is distributed between multiple stakeholders (DPO, privacy engineers, privacy layers, software engineers) within the company, and it takes considerable time to finalize a small change or update.

6.2.3 Discussion

Simultaneously, considering the results (**Section 6.2.2**), we observe that the proposed RMDP model, Privacy document, and model execution way have the following practical (and theoretical) shortcomings and overlook interpretations. Referred shortcomings are distributed and highlighted further. The RMDP model lack:

- **tracking** and **monitoring** capabilities of various roles activities, including the essential control points in the privacy document;
- various **templates** for various use cases, including extended interpretations for the mentioned use cases;
- **technical methods** for evaluation of the effectiveness of various DP forms in the context of the proposed DPG and methods for estimating or **scoring** criticality levels for the DP;

In addition, previously highlighted complexity enables further shortcomings in the complexity of interpreting RMDP model:

- countermeasures and the dependence between the suggested countermeasures;
- results for experts with an insignificant level of expertise in this domain;
- the starting point and beginning position in RMDP model;
- also concluded that the effectiveness of using the model depends on the degree of effectiveness of the presentation;
- the results and communicating them with the Security experts;

We also note that referred weaknesses are partly recalled by the decided abstraction level for the produced reference level (which is relevant for all reference models and partly applied for conceptual models). Nevertheless, we argue that the organisation could neglect demonstrated shortcomings while adapting or modifying the RMDP model.

Further, we came to the conclusion that interview participants without additional certifications in the domain of privacy and data protection provided less critical and expanded comments, so in quantitative

interviews (**Section 6.3**), we attempted to select candidates with additional certificates confirming improvement or extra knowledge in the data protection and privacy domains. We also want to note that interviews with more comprehensive industry expertise have more critical comments regarding the presented RMDP model; nevertheless, comparing less experienced and trained participants samples, we can conclude that the sample with less expertise level has a more general understanding of the presented study. However, the group with more extensive expertise and practical experience has more in-depth knowledge regarding particular quadrants, or integrated concepts and less broad knowledge of all demonstrated quadrants. Depending on the organization's size, job responsibilities for different roles may differ, affecting the adoption processes and time intervals for communicating results to diverse stakeholders.

Lastly, experts' expertise level, including the taxonomy they operate within an organization, may differ depending on the approach or industry standard used within the organization. Thus, organizations should ensure or approve a single taxonomy before interpreting the results between different stakeholders. This delays or distorts stakeholder knowledge and reduces the suitability of adaptation and monitoring of the GDPR requirements. Also, due to the dynamic concept and definition of privacy for different contexts or use cases, various stakeholders can distort the GDPR requirements. As a result, the suggested RMDP model avoids distortion of the GDPR requirements and allows all stakeholders to use or adopt the taxonomy without distorting the GDPR requirements.

6.3 Quantitative evaluation

6.3.1 Approach

In order to evaluate the alpha version of the developed artefact (**Figure B.3**), the author of this dissertation conducted a standardized survey among various domain experts. We sent a request for an interview to more than 200 potential respondents in the EU directly via private messages to Linkedin[2]. Candidates were determined based on the principle that they had sufficient practical experience and related roles, also endorsed by additional certificates focused on privacy and data protection (e.g. IAPP). We also want to note that some of the participants in the Quantitative evaluation phase were among earlier selected experts via the qualitative evaluation phase. We also highlight the importance of distributing our request to the professional community or inside the organization. Participants showed a significant level of interest in participating, which we greatly appreciate. Participation was based on a voluntary basis. **Appendix C** contains a list of questions, including thirteen five-point Likert-scale questions distributed among various experts and follows the earlier criteria presented by the author [6]. After the presentation of the artefact (RMDP model) and a short discussion, the participants were provided with a link to Google Docs to evaluate the suggested RMDP model according to the adapted criteria of **Section 6.1.1** (*Simplicity; Accuracy: Scope; Systematic Power; Explanatory Power; Validity; Reliability; Fruitfulness*).

We would also like to highlight the careful consideration that went into our survey design. We aimed to keep the survey as short as possible, believing that this would ensure a high level of respondent engagement and improve completion rates. We also took care to formulate the shortest questions possible without losing context, as we understand that long questions may hinder completion or motivate respondents to skip questions. These considerations, along with other optimizations, were made to obtain more significant results from the survey we produced and conducted. In addition, we incorporate the net promoter score (NPS) [156], which is utilized in surveys to gauge the respondents' commitment and fulfilment, providing a reasonable and quick snapshot of the fulfilment of various groups. The NPS score (0-10) is calculated by subtracting the percentage of Detractors from the percentage of Promoters: NPS= %Promoters - %Detractors; (Detractors (score 0-6), Passives (7-8), Promoters (9-10)). An advantage of the referred approach could be highlighted and not limited only to simplicity. Nevertheless, utilizing the referred approach as a standalone could lead to) oversimplification, lack of context and response bias. Thus, we incorporate this approach in our comprehensive multi-phase evaluation strategy, which is enriched by qualitative feedback, to ensure a thorough and robust analysis.

[2] linkedin.com

6.3.2 Results

Following the approach explained in *Section 6.3.1*, we explain our results by first providing a snapshot of the participants' sample, including supplementary information, followed by more detailed aspects of the multi-phase evaluation strategy, complemented by our analyses.

The survey was distributed among 250 recipients per private message over the social business network Linkedin; we also approached respondents to share the invitation among colleagues or among the community; our search cast in, and our screening was scoped by filtering possible applicants in terms of a) current roles (privacy engineer, data protection officer, software engineer, privacy or legal counsel) in the diverse organizations, b) documented further knowledge or confirmation of expertise via additional certification (e.g. IAPP) and c) outlined the scope of practical experience or accomplishments regarding data protection or privacy issues, implementation of DP or other PETs.

Among all of the approached candidates, only 14% participated in the survey, and only 8% participated and contributed to a survey, as demonstrated in **Table 6.4**. Among 8% participants, every participant stated that they have experience in Data Protection and Privacy or a related domain with average professional experience 3.15 years (**Figure 6.2 left**) , and 55% confirmed that professional experience related to the industry and 25% comes from industry and academia, simultaneously (**Figure 6.2 right**). In other words, we could state that 75% of the respondents who took part and completed the survey have practical experience related to the various industries and applicable to diverse organizations. Additionally, the results (**Table 6.5**) demonstrate that the suggested criteria and their evaluation by diverse experts satisfy development and designer goals with minor comments. Further, **Table 6.6** demonstrates a detailed analysis of each criteria (i.e. *Simplicity, Accuracy, Scope, Systematic Power, Explanatory Power, Reliability, Validity and Fruitfulness*), explained and discussed in **Section 6.5**.

Considering the weighted average score[3] (**Table 6.6**), most scores for earlier presented criteria in a range between 3.1 to 3.95, which indicates overall positive sentiment. Additionally, most criteria have weighted average scores above 3.5, indicating participants generally agree with the model's elements and structure. This demonstrates that the developed reference model meets the expected design and developers' goals. However, areas of concern where a weighted average score near 3 or lower need attention or such scores reveal strong disagreement or mixed feedback. Also C3.2 with weighted average score 1.7, could be explained that model doesn't include any irrelevant concepts or elements to the studied domain.

[3]

$$\text{Weighted Average Score} = \frac{(w_{\text{SA}} \cdot n_{\text{SA}}) + (w_{\text{A}} \cdot n_{\text{A}}) + (w_{\text{N}} \cdot n_{\text{N}}) + (w_{\text{D}} \cdot n_{\text{D}}) + (w_{\text{SD}} \cdot n_{\text{SD}})}{n_{\text{Total}}}$$

- $w_{\text{SA}}, w_{\text{A}}, w_{\text{N}}, w_{\text{D}}, w_{\text{SD}}$ are the weights for the responses:

$$\begin{aligned} w_{\text{SA}} &= 5 \\ w_{\text{A}} &= 4 \\ w_{\text{N}} &= 3 \\ w_{\text{D}} &= 2 \\ w_{\text{SD}} &= 1 \end{aligned}$$

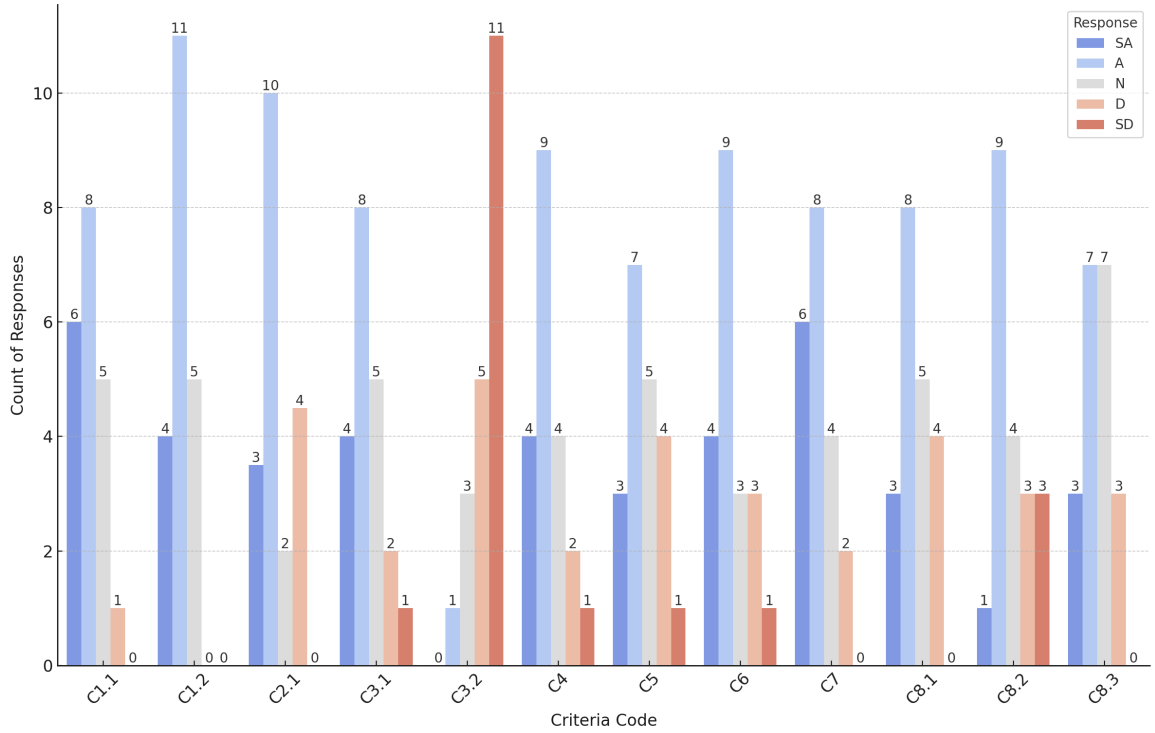
- $n_{\text{SA}}, n_{\text{A}}, n_{\text{N}}, n_{\text{D}}, n_{\text{SD}}$ are the counts for each response type.

- n_{Total} is the total number of responses for that criterion:

$$n_{\text{Total}} = n_{\text{SA}} + n_{\text{A}} + n_{\text{N}} + n_{\text{D}} + n_{\text{SD}}$$

Further, analysis of measures of the variability or dispersion of responses Standard Deviation (SD)[4] for each criteria demonstrates the indicated consensus among respondents regarding the quality of the developed RMDP model and adapted criteria (*Section 6.1.1*). Additionally, C2.2 with 5.04 related to the high SD, which demonstrates areas of disagreement or confusion[5] that require further clarification or refinement. Low standard deviation points to strong consensus and stability, suggesting these aspects of the model are well-understood and accepted. Particularly, C3.1 (SD=2.74) the lowest SD points to strong consensus and stability, suggesting these aspects of the model are well-understood and accepted.

Additionally, analysis (**Table 6.6**) for Positive, Neutral and Negative responses[6]. Indicate that *% Positive* criteria with high % in **bold** are well accepted (e.g., C1.2 suggest these criteria are well-accepted.), low % reveal irrelevance or dissatisfaction (e.g., C3.2 reveal irrelevance). *% Neutral* criteria with high % in **bold** demonstrate uncertainty (e.g., C8.3), and *% Negative* criteria with high % in **bold** illustrated need to improve or redesign (e.g., C3.2 indicates areas needing improvement or redesign).



Note: (SA): Strongly Agree, (A): Agree, (N): Neutral, (D): Disagree, (SD): Strongly Disagree

Figure 6.1 Distribution of survey results

We provide distributions of roles / specified domains as demonstrated in **Table 6.2**. The distribution demonstrates a strong bias towards data protection experts. Well-presented groups included software

[4]Standard Deviation (SD):

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

where:

- σ: Standard deviation
- N: Total number of data points
- x_i: Each individual data point
- μ: Mean of the data points, given by $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$

[5]or uncertainty regarding the accuracy of interrelationships

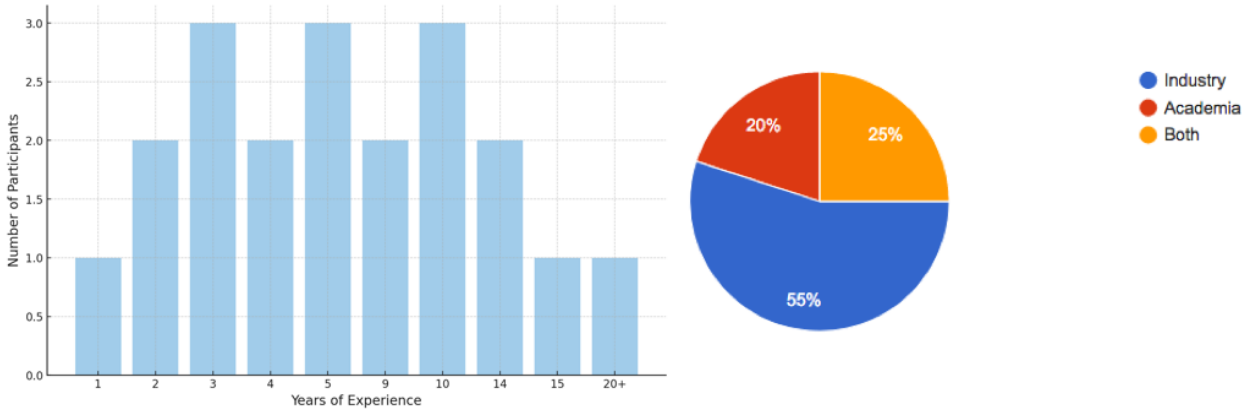
[6]% Positive: Responses (SA + A), % Neutral: Responses (N), % Negative: Responses (D + SD)"

developers, privacy engineers, and privacy legal experts, and referred groups were less represented as data protection experts. We also want to note that every represented group has additional privacy or data protection certification (the dominant certification body is IAPP). As depicted in **Figure 6.1**, the respondent distributes the preference considering criteria defined in **Section 6.1.1**. The results also demonstrate that when considering various roles, such criteria (*Simplicity, Accuracy, Scope, Systematic Power, Explanatory Power, Reliability, Validity and Fruitfulness*) highlighted the relevance of RMDP and feasibility for specified roles.

Group	Completed (%)	Partly Completed (%)	Not Completed (%)
Data Protection Experts	2.80	1.60	20.60
Software Engineer	1.60	0.80	22.60
(Privacy) Legal Expert	1.60	1.60	21.80
Privacy Engineer	2.00	2.00	21.00

Note: Total number of participants: 250 individuals

Table 6.4 Overview of the survey response rate for different groups



Note: **a) Left**: Experience in years, **b) Right**: Distribution of Professional experience

Figure 6.2 Participants sample supplementary information

In addition to already collected data and in line with a multi-phase evaluation strategy, as explained before, we assess the satisfaction of respondents via the introduced NPS score by stating further question: "*On a scale of 0 to 10, how likely is it that you would recommend RMDP to a colleague?*". Results demonstrate that mostly third of respondents related to the Passive satisfied group (**Figure 6.3**), which is not included in the calculation of the NPS score [7]. That partly explains the negative results demonstrated (**Figure 6.4**). Results also could be interpreted that even though the NPS scores for Data Protection Experts groups close and satisfy the overall NPS score, it demonstrated the need for a more significant number of respondents for other roles, which presented a limited sample of participants. In addition, the results obtained were slightly opposite to the results from the Qualitative evaluation (*Section 6.2*), where the referred groups stated that the current model partly fits current daily tasks and could be applied with the adjustment of the dimensions. Nevertheless, considering the drawbacks documented (*Section 6.2.3*)

[7]

$$\text{NPS} = \frac{\text{Number of Promoters} - \text{Number of Detractors}}{\text{Total Number of Respondents}} \times 100$$

where:

- Promoters: Respondents who score 9–10 on the survey
- Detractors: Respondents who score 0–6 on the survey.
- Total Number of Respondents: Total participants including Promoters, Passives, and Detractors

Criteria Code	Survey Question	SA	A	N	D	SD
C1.1	Are the elements of the Model shown simple?	6	**8**	5	1	0
C1.2	Are the relationships between the elements simple?	4	**11**	5	0	0
C2.1	Are the classifications included in the Model accurate?	6	**8**	3	3	0
C2.2	Are the interrelationships between the elements of the Model accurately represented?	1	**12**	1	6	0
C3.1	Does the Model include all elements/concepts essential for the Privacy domain?	4	**8**	5	2	1
C3.2	Does the Model include any elements not applicable/relevant to the Privacy domain?	0	1	3	5	**11**
C4	Does the Model organize elements of the Privacy domain and the interconnection between them in a structured, systematic way?	4	**9**	4	2	1
C5	Could the Model assist with clarifying (tracing back) and forecasting issues related to Privacy?	3	**7**	5	4	1
C6	In your opinion, would the Model be reasonable for the majority of business organizations?	4	**9**	3	3	1
C7	Would the methodology incorporated into the Model lead to reasonable results?	6	**8**	4	2	0
C8.1	Does the Model deliver a convenient structure for framing existing research?	3	**8**	5	4	0
C8.2	Would you utilize the current Model in your research/practice?	1	**9**	4	3	3
C8.3	Could a model be utilized or assist in bridging the gaps in the existing research/practice?	3	7	7	3	0

Note: (SA): Strongly Agree, (A): Agree, (N): Neutral, (D): Disagree, (SD): Strongly Disagree; Values in **bold** indicate the highest value per criteria.

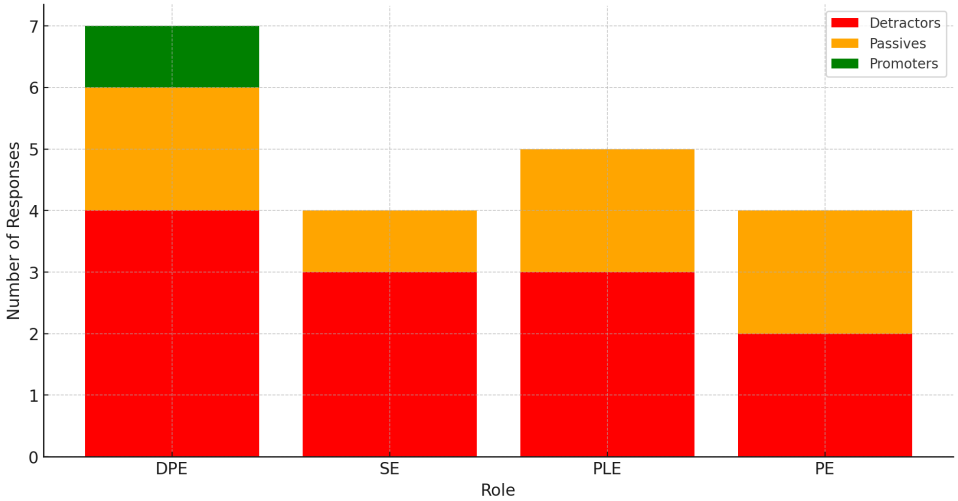
Table 6.5 Survey summary results

related to the abstraction of the model and also the need to tune the model for diverse use cases can partly explain the results received. We also consider that involving more participants in the less represented groups can enhance the overall and each group's NPS score, demonstrated by the most representative group of data protection experts. Lastly, the results highlight that the distribution of the participants is relevant and practical for the evaluation of the developed model, despite the possible implications of not having a representative sample for groups except data protection experts with a potential impact regarding the choice of evaluation method (explained in *Chapter 7*).

6.4 Analytical Discussion

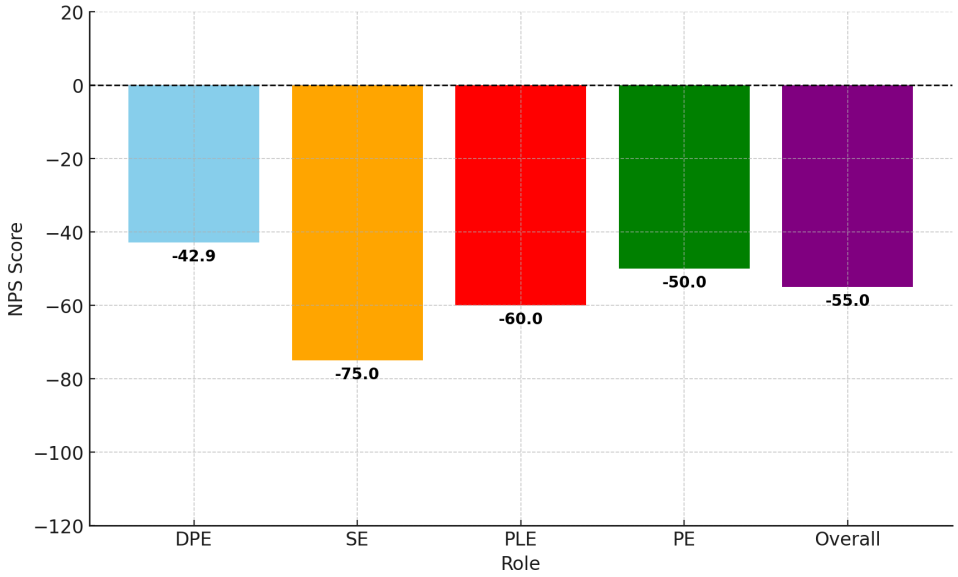
6.4.1 Fulfilment of Guidelines of Modeling

Conceptual models, like reference models, must satisfy not only the main requirements (**Section 6.1.1**) but also secondary ones (relevant to the area under study or the purposes of design and creation of the reference model). In addition, in agreement with the authors [172, 159, 173] the scientific rigour of RMDP can be assessed according to the basic modelling requirements, which can be a normative basis for a qualitative assessment. In this case, we carry out the assessment according to the Guidelines of Modeling (GoM). On the one hand, the concept of GoM is a framework intended for the development of specific recommendations during design to improve the quality of models; on the other hand, it adapts elements



Note: DPE: Data Protection Expert, SE: Software Engineer, PLE: Privacy Legal Expert, PE: Privacy Engineer; *Passives* are not included in NPS computation.

Figure 6.3 Survey distribution of NPS categories by groups



Note: DPE: Data Protection Expert, SE: Software Engineer, PLE: Privacy Legal Expert, PE: Privacy Engineer

Figure 6.4 NPS scores by group and overall

Criteria Code	Weighted Avg.	SD	% Positive	% Neutral	% Negative
C1.1	3.95	3.39	**70**%	25%	5%
C1.2	3.95	4.53	**75**%	25%	0%
C2.1	3.85	3.08	**70**%	15%	15%
C2.2	3.4	5.05	**65**%	5%	30%
C3.1	3.6	2.74	**60**%	25%	15%
C3.2	1.7	4.36	5%	15%	**80**%
C4	3.65	3.08	**65**%	20%	15%
C5	3.35	2.24	**50**%	25%	25%
C6	3.6	3.00	**65**%	15%	20%
C7	3.9	3.16	**70**%	20%	10%
C8.1	3.5	2.92	**55**%	25%	20%
C8.2	3.1	3.00	50%	20%	30%
C8.3	3.5	3.00	50%	**35**%	15%

Note: To calculate the weighted average score for each criterion code based on the Likert scale, we assign the following weights to the responses: Strongly Agree (SA) = 5, Agree (A) = 4, Neutral (N) = 3, Disagree (D) = 2, Strongly Disagree (SD) = 1; Percentage of Positive Responses: (SA + A). Percentage of Neutral Responses: (N). Percentage of Negative Responses: (D + SD); *Positive Responses*: High positive percentages in bold (e.g., 70%, 75% for C1.1 and C1.2) suggest these areas are well-accepted. Low positive percentages (e.g., 5% for C3.2) reveal irrelevance. *Neutral Responses*: High neutral percentages (e.g., 35% for C8.3) suggest insufficient clarity in these criteria. *Negative Responses*: High negative percentages (e.g., 80% for C3.2) highlight areas needing improvement or redesign.

Table 6.6 Survey results analysis

Guideline	Justification
Construction adequacy	Consensus on problem definition and model representation (minimality. intra-model relationship and inter-model relationship);
Language adequacy	Linguistic accuracy (consistency, completeness); language aptitude (semantic power); understability of the language and application (including tool); formalization of language.
Efficiency	Consensus building; language understanding and application; translatability across Perspectives.
Clarify and understandability	Clarity of hierarchy, Comprehensibility of the layout, Filtering
Systematic Structure	Architecture, inter-model view relationship;
Comparability	Semantic compatibility (with similar models);

Table 6.7 Guidelines of Modeling for reference model [174]

of existing approaches intended for the evaluation of models [15]. According to the authors [174] in the context of the main requirements (primary level one) for conceptual models, the following perspectives are presented together with comments in **Table 6.7**. The following comments and discussions consider the main requirements of the GoM for reference models and additional ones (**Section 4.3**).

The author [173] points out the first point (**construction adequacy**) as one of the most relevant and critical of the six presented. Throughout the dissertation, we try to transparently demonstrate the chronology and process of model design and the corresponding decisions. We also inform the readers at each stage of the decision-making process, corresponding to each design and development stage, including providing critical analysis and facts influencing certain decisions.

The second guideline (**language adequacy**) is presented and discussed in **Section 5.8**, including, to some extent, **Section 4.3.3**. Visual presentation and description of each concept (and visual display), as well as each of the four areas (blocks) in mind-map format and in accordance with written language and inclusion of notation-free, is the most appropriate choice. Since the RMDP model is conceived and designed as a reference model allowing to be used by diverse participants with different expertise levels, including from diverse expert domains. Therefore, the appropriate choice of written language, in the context of simple concepts, is one of the conditionally sufficient and acceptable for achieving the stated goal.

Accordingly, the third guideline (**efficiency**) points to the possibility of a reference model and the need to take into account various factors, as well as the ability to adapt to external changes or other factors. The

RMDP model allows for internal changes to be taken into account at different stages of model utilization in the context of various phases of the Evaluation block, including monitoring and access provided for all the different roles involved in the process of compliance and GDPR requirements. We also believe that the design and graphical execution of the reference model can be effectively supplemented by various GDPR amendments or interpretations for the Information Taxonomy, DPG and Countermeasures Dimension. At the same time, the reference model has the ability to complement or adapt the taxonomy for different organizations and their goals, thereby providing a reliable basis for the integration and adaptation of changes of different natures or characters.

The guideline (**clarity**) is presented using various concepts and elements at a certain level of abstraction. We also argue that the selected layout for the RMDP reference model is a) the most optimal due to the vastness and complexity of the GDPR requirements and b) self-sufficient and essential since it is presented by various concepts that can be supplemented either at the expense of each other or be independent in the context of the tasks of assessing the degree of compliance with the GDPR requirement. The basis of the reference model - structure, wording and visual presentation - were optimized and supplemented repeatedly during the development of the model and the concepts incorporated.

The RMDP model is depicted in such a way that it provides a clear understanding of the stages and goals for protecting personal data (in the context of the presented and existing data protection goals). In particular, the RMDP displays four dimensions (blocks) in the context of interpreting the GDPR requirements for personal data and PET, which are interconnected by a continuous cycle of PDLC improvements, resulting in a Privacy document. Also, each of the blocks is defined by success criteria based on recommended industry and academic practices.

In addition, the use of RMDP allows one to keep records, register and fill out a privacy document intended to supplement (**language adequacy**) guideline capabilities and also performs the primary role presented in the guideline (**systematic structure**). It is worth noting that the questionnaire participants confirmed the straightforward structure and relationship inherent in the RMDP reference model.

The final guideline (**comparability**) in the RMDP reference model is formalized using various concepts integrated into a single structure and a graphical representation with explicit subordination of execution supplemented by definitions for various concepts, including using a metamodel. Thus, this allows for the interpretation and integration of the RMDP reference module with other types of conceptual models that are not limited to reference models only.

We also include commentary from designer and developer of RMDP considering the principles (*Abstraction, Originality, Justification and Benefit*) outlined in **Section 1.7** applied to the developed artefact (reference model); further, in **Chapter 7**, we demonstrate the limitations, followed by a discussion. *Abstraction*: developed artefact could be applied to diverse classes of problems such as privacy, data protection, and fulfilling GDPR requirements, neglecting gaps in previously identified methods (**Chapter 3**). *Originality*: RMDP aggregates distributed interdisciplinary knowledge associated with various domains such as privacy legal, security, privacy, data protection and PETs implementation. *Justification*: RMDP is presented and partly evaluated by various practitioners and academics, with limitations partly demonstrated in **Chapter 5** and **Section 7.3**. *Benefit*: Selected participants and stakeholder groups demonstrate immediate advantages and disadvantages in utilizing the RMDP (**Chapter 6**).

6.4.2 Discussion and Evaluation for special requirements

As we noted earlier, DSR artefacts can be characterized in the context of practical relevance and scientific rigour [83]; Next, we evaluate practical rigour in the context of the eight recommendations presented earlier (**Section 4.3**) and their adherence to the RMDP. We present the evaluation of the recommendations and our comments below, which presents the recommendations and accompanying comments. Which can then be interpreted as indicating that the recommendations presented earlier (**Section 4.3**) are satisfied.

- **R1**: *Adequate and clear visualization and (step-by-step and consistent) compliance with GDPR requirements*.

The RMDP contains DPG declared in the GDPR, including *Data Minimization, Availability, Integrity, Confidentiality, Unlinkability, Transparency, and Intervenability* that reflect the path to compliance and management of compliance requirements and transform the DPG for PET, including considering the diversity of many different organizations and the corresponding diversity of goals and requirements presented by organizations. The visual representation of the RMDP links the various essential concepts for compliance with the GDPR, represented by four blocks (dimensions) and the sequence of use of the RMDP (which is directly recorded and results in the corresponding privacy document).

- **R2**: *The ability to trace the relationship and phasing between various stages/dimensions and/or concepts*.

RMDP represents four main blocks: *Evaluation, Information Taxonomy, DPG and Countermeasures* (dimensions) for following and fulfilling the GDPR requirements. Additionally, the Evaluation dimensions define five additional phases for the suggested PDLC. According to the assessments during the survey, the presented RMDP model includes the concept of data lifecycle in the Information taxonomy area. Thus, it allows tracking all processes at all stages (*Creation, Transmission, Processing, Storage and Destruction*), which are also in accordance with the requirements of the GDPR and meet the requirements to track the relationship and phasing between various stages/dimensions and/or concepts.

- **R3**: *Ability to adapt the model depending on the context, needs, objectives and requirements of organizations*.

RMDP delivers the capability to adapt and modify or supplement the presented areas depending on the various requirements or tasks of organizations. Each of the four dimensions is consistent with the presented requirement.

- **R4**: *Possibility of progressive and repetitive execution*.

RMDP is presented and developed as follows as: a sequential execution of each phase for each measurement, where all results are reflected and recorded in a single privacy document. The use of the model occurs clockwise, and for each stage of the Evaluation dimensions, a stage (execution) and analysis of each of the four areas is performed.

- **R5**: *The possibility of intersection with the GDPR requirements the process of integrating them in the context of PET DP utilization*.

RMDP includes the areas of information taxonomy, data protection goals and privacy countermeasures; in addition, we have adopted the appropriate criticality level (high-medium-low), which relates to the interpretation of the GDPR requirements and concerns aspects of the integration and disposal of PET DP.

- **R6**: *Presentation of the possibility and practical information for fulfilling or following individual requirements or elements of individual parts (such as Article, Recital and others) of the GDPR*.

RMDP includes all available data protection goals presented in the GDPR, including the additional data protection goal, Data Minimization; according to the adapted table (**Table 2.2**), each of the Data protection goals can be associated with different parts of the GDPR (**Table 2.2**) represented by the corresponding Article, Recital or their combinations. In addition, Information Taxonomy includes the area that relates to the definition and impact of requirements in the context of anonymization levels.

- **R7**: *Promote and encourage the reuse of existing concepts and elements*.

RMDP (as a reference model) contains a simplified visual representation and notation-free approach to concepts and elements related to assessing the level of compliance in the context of the GDPR, including a privacy document that consistently records all relationships and interrelations.

- **R8**: *Take into account the importance and significance of various stakeholders and interested parties in the context of GDPR compliance, thereby ensuring the ability to interpret the results and support compliance relationships between the various stakeholders directly involved in the continuous data protection and compliance process*.

RMDP is developed and presented in such a way that it is suitable for different actors with different levels of expertise. Simultaneously, incorporated concepts enable interpreting the results of diverse actors and interested parties in a single format, and the model is a platform for combining the examination results into a single format comprehensible to all actors and interested parties.

6.5 Summary

In this section, we summarize the results and the multi-phase approach to assessing the developed RMDP artefact. According to [14, 173, 172], the scientific rigour of RMDP can be assessed according to the basic requirements of modelling (**Section 6.2-6.4**), which can be a normative basis for a qualitative assessment. We also want to note that before the start of the main phase of the model evaluation, we carried out some preparatory steps to improve and enhance the quality of the assessment of the developed model (**Section 6.1.1**). The constructed interview questions, including the selected criteria and evaluation methods, were refined and modified, considering the given comments.

The developed model was rigorously evaluated through in-depth interviews with a variety of experts in the industry and academia. These experts, specializing in data protection, privacy, software development, privacy engineering, and privacy legal counsel, provided valuable insights and comments summarised in **Table 6.3** and **Section 6.2.2**, followed by a discussion of the results (**Section 6.2.3**).

Qualitative evaluation provides rich data for analysis in the context of the four expert groups. In-depth interviews, presentations of the RMDP model, and subsequent discussions, including 19 open-ended questions, led to the following results. The evaluation results demonstrate that RMDP (as a reference model) can represent the domain of personal data protection (GDPR and also not limited to the referred legislation) in the context of the application of PET, namely DP, including advancing the gaps from comparable analyzed approaches (**Chapter 3**). All interviewed participants agreed that the presented model reflects the necessary GDPR requirements at the presented abstract level. However, it's important to note that the model, while comprehensive, may not fully address all current GDPR requirements. For instance, experts noted that the lack of an example makes it difficult to use or adapt the model, and experts pointed out that specific GDPR requirements, such as data portability and data transfers, may not be covered entirely. This balanced view of the model's applicability is crucial for understanding its limitations. Also, the proposed RMDP model is not limited to representing the data protection domain but also contributes to the privacy domain. Moreover, experts agree that the capability of the developed model dominates the privacy domain and could be applicable to diverse PETs, showing promising potential for future applications.

Quantitative evaluation with diverse experts considering the eight criteria (*Simplicity, Accuracy, Scope, Systematic Power, Explanatory Power, Validity, Reliability, Fruitfulness*) were conducted with thirteen questions (Likert Scale) explained in **Section 6.3**. We obtained unambiguous answers, and the results and comments are presented below. Most experts participating in the survey came to a unanimous conclusion and agreed that RMDP could describe and reflect the complexity of the privacy and data protection domain, including corresponding to the qualitative characteristics by which the model was assessed. In the context of use and thematic presentations and research results, we have demonstrated that the development method and refinement process (also improvement and adaptation of various comments) of the developed RMDP model is a) reasonable for various groups of stakeholders involved in evaluating the degree of compliance with the GDPR norms, b) feasible and applicable for various groups with different levels of expertise and c) in demand in the context of the need to communicate the results of the assessment between various stakeholders and roles who interact on an ongoing basis (daily) in the context of keeping and monitoring the implementation of the GDPR norms. More then 65% of respondents imagine and comprehend the requirements introduced by regulators in the context of the GDPR requirements could be

traced with RMDP. Simultaneously, more then 75% state that their activities as being related to the protection of the rights of data subjects and their activities related to the privacy and data protection domain applied in organizations. More then 65% of respondents confirmed that the reference model is practical for their everyday activities in the context of such criteria as Accuracy, Scope, and Systematic power. According to the presented criteria, the most remarkable assurance among respondents is located in descending order respectively *Accuracy, Scope, and Systematic power*. Also, a significant part of respondents agreed that the suggested criteria for the reference model correspond most accurately. This allows us to conclude that the model reflects the reality of assessing the compliance level with the GDPR requirements by considering the DPGs given for PET DP and various DP forms. On the other hand, the most critical assessed criteria are (*Accuracy, Explanatory power, Fruitfulness*). In addition, the criteria (*Fruitfulness*) and the analysis results reflect the need for further refinement of the reference model, which was also partly revealed during the in-depth interviews with various experts. Nevertheless, the results for the NPS score are slightly opposite and inconsistent. Even though the selected groups are highly practical for evaluating the developed model, we record implications regarding the chosen approach as explained further. The passive group, which comprises mostly one-third of the respondents, is not included in computing the NPS score. Additionally, the results underscore the necessity of increasing the number of participants for other groups, except the data protection group, to validate results more comprehensively.

Further, we present an analytical assessment in relation to the requirements of a) the reference model modelling, b) the artefact under development and c) the analytical assessment by the developer and designer of the model. Our analytical discussion (*Section 6.4*) demonstrates that the reference model development meets the requirements outlined in **Chapter 4** and follows the development methodology (**Section 1.7**). Results demonstrate that the developed RMDP reference model communicates knowledge in the domain of personal data protection considering GDPR requirements for personal data and is applicable for diverse groups or roles accepting and utilizing the RMDP model, including those with various experience levels and not limited to the privacy or data protection domain. Also, the results demonstrate that the suggested model partly enhances complexity in such specific and not established industry-specific intricacies: a) dynamic taxonomy and definition of privacy, b) trade-off utility and privacy of PETs as well diverse amount of DP forms, c) suggest privacy development lifecycle applicable and accessible for diverse stakeholders d) outline cooperative effort in fulfilling the GDPR requirements relevant to the data subjects data.

Initially, we want to note the complexity of assessing the developed artefact due to the interdisciplinarity and goals of this work, which will be demonstrated more explicitly in the following **Chapter 7**. Furthermore, we want to remark on the complexity of evaluating the artefact among the various previously presented stakeholder groups (due to various responsibilities performed by the same roles and the difficulty of locating interview participants due to the interdisciplinarity of the research goal and questions, as well as the limited number of respondents with the necessary knowledge in the examined domain and still the reasonable implementation gap between research and industry in adopting PETs). Also, difficulties accompanied the development and evaluation of the proposed RMDP models. Even though the evaluation process and expert groups were scoped for the four groups, we faced multiple challenges in locating an appropriate expert who could contribute to the RMDP development due to issues with expertise levels and also the interdisciplinary nature of research. We also want to note that the adaptation of DP or other PETs in the organization is limited.

7 Conclusion and outlook

This Chapter summarises the current dissertation's results, demonstrates the contribution to the advancements of the body of scientific knowledge, and importantly, expands the results to the wider research community. We also acknowledge the limitations of the research, developed artefact, and propose the next steps for further research. The summary and contribution of research, considering the research questions presented in the first chapter, are demonstrated in **Section 7.1**. The contribution, limitations, and possible threats to our results are described in **Sections 7.2** and **7.3**. We further elaborate on the future research steps and conclude with reflections and outlook in **Section 7.4**.

7.1 Summary

This dissertation delivers particular attention to the role of personal data and further issues with data protection in the context of the European Union (EU) General Data Protection Regulations (GDPR) legislation, including privacy issues and GDPR compliance in the context of utilizing additional Privacy Enhancing Technologies (PET) approaches through Data Protection Goals (DPG). Moreover, the focus is also directed not only to different types and sizes of organizations, including diverse stakeholder roles in the organization that accept and play a significant role in complying with the GDPR requirements for personal data protection and compliance with the privacy principles addressed through Privacy by Design (PbD) but also partly to the related applicable issues of determining the level of anonymization and establishing privacy taxonomy. Nevertheless, the matters recalled are receiving close attention both in the industry and academia and are among the top priorities due to the high demand. Yet, researchers and practitioners have actively investigated and continue to study issues of privacy and data protection, the degree of compliance with the GDPR and develop additional usable approaches or arguments for utilizing various PETs to fulfil the goal-based GDPR requirements. However, the attention of researchers and practitioners is not sufficiently focused on interdisciplinary issues and the synthesis of knowledge between technical and legal aspects of the use of PET. There are still many open questions and gaps related to the interpretation of:

- the GDPR requirements for the technical and legal domain;
- the security and privacy domain;
- dedicated taxonomy related to privacy and data protection, including also difficulties in adapting a single definition and methods for assessing the levels of anonymization (additionally defining the boundaries between the levels of anonymization);
- how to comprehend and assess the degree of compliance of supplementary PET approaches within the GDPR;

Moreover, despite numerous open technical and practical issues, including conceptual ones, many organizations and researchers are primarily focused and pay the most significant attention to the so-called data utility tradeoffs and the effectiveness of using PET in the context of the GDPR requirements for personal data and their improvement due to the high demand from the industry. Further, besides the issues of interpretation and communication, the results of such complex interpretation for technical enhancements, practitioners also experience multiple difficulties and further problems in joining them to the legal interpretations, followed by further complexity in articulating results between diverse stakeholders involved in such a process. However, following the gradual increase in the use of referred PETs, issues of compliance with the GDPR requirements in the context of personal data processing are becoming an integral part of

any organization in various industries since personal data and their multiple arrangements (transformation of personal data) are becoming more critical and essential for any organization. Thus, organizations are faced with the questions (related not only to the tradeoffs between data utility and privacy) of how and by what means or frameworks they can claim and demonstrate that the requirements of data controllers considering GDPR for processing of personal data are fully met and decided technical methods corresponds with the necessary level of anonymization, particular observed for Differential Privacy (DP). Since the adaptation of personal data and its use by organizations is becoming the most critical for many organizations and non-compliance with the GDPR requirement can potentially lead to various critical consequences for both organizations, various stakeholders and data subjects, the assessment of compliance with the GDPR requirement in the context of the use of PET for personal data needs to be rethought.

Chapter 1 outlines the various challenges in the EU for diverse organizations and discusses the difficulties of compliance and fulfilment of the GDPR requirements for varied transformations of personal information and PET implementation, as the accompanying difficulties in interpreting the data protection and privacy requirements, emphasizing the interdisciplinarity of this study. Thus, five research questions are formulated, including the high-level goal of this study, which is proposed and explained. The suggested research questions were developed considering the discussion of the problem specifics in the studied domain, which are accompanied by a dialogue of the most feasible research strategy for the development of the primary artefact of this work, namely, a reference model for evaluating the degree of compliance with the GDPR for PET in the context of adapted DPGs. We also accompany explanations and demonstrate our research activities and contributions for various chapters, including the methods used, which are presented utilizing visualization of this dissertation structure. Simultaneously, we supplement and refer to the already presented academic results chronologically and debate potential limitations.

In order to deliver comprehensive knowledge and specificity of this study, in *Chapter 2*, we discuss and explain the terminology and the most essential GDPR definitions and accompanying concepts for privacy, data protection and related domains and supplement and revise them partly with already published results in academia. We also discuss in more detail the complexity of the interpretation of technical and legal aspects for the studied PET and the suggested Data Protection Goals to deliver a more adequate knowledge and further use of the complexity in the interpretation of the GDPR requirements.

Chapter 3 demonstrates various industrial and academic approaches that are to some degree relevant to the research objective of this dissertation and also to earlier presented research questions, as well as the subsequent results of the critical analysis and shortcomings of the examined approaches. The examined approaches address and support to some extent the issues related to the transformation and protection of personal data in the context of GDPR, privacy and the PET utilized. We critically debate and examine the appropriateness of these approaches for the research questions, based on which we determine further research gaps that we attempt to address and fill with the help of this dissertation.

Chapter 4 is devoted to the beginning of the development of the primary artefact of the reference model, which starts with the representation of the methodology for developing the reference model, the reasoning for including various concepts and data protection goals, followed by the formulation of the main requirements for the reference model. This includes fulfilling and guiding requirements for the development of the reference model directly corresponding to the research goal of this dissertation and the presentation of the metamodel, which demonstrates the predicted and planned goals and capabilities of the future reference model.

We discuss and provide explanations for incorporating the various necessary concepts into the core reference model artefact in *Chapter 5*. This chapter starts with the presentation of the developed RMDP reference model artefact. Next, we demonstrate and discuss each block in more detail and explain the reasons for each block of the reference model, including the motivations and goals pursued for integrating the various concepts into the reference model. After that, each block is discussed in more detail and separately with a more in-depth examination and justifications for incorporating the suggested concepts to satisfy the GDPR requirements based on earlier published results in academia, supplemented by further academic contributions in the research domain. We complement the developed reference model with various design decisions and an explanation of the decisions taken. The chapter concludes with various aspects related

to the utilization of the reference model, which includes explanations and relationships between the presented blocks and concepts. Moreover, we demonstrate and discuss the Privacy Document, which enable to register the implementation of the reference model and reflects the results of operating the reference model, which can be broadcast to various stakeholders with diverse levels of expertise. In conclusion, we conclude the chapter with a discussion of the developed reference model, including a critical dialogue of abstraction and limitations, including implications for the designed artefact.

Chapter 6 demonstrates in detail the process and methods for assessing the developed reference model, which is accompanied by explicit results of the reference model examination utilizing qualitative/quantitative methods with various stakeholders involved (data protection office, privacy legal experts, privacy engineers and software developer) in assessing the degree of compliance with the GDPR provisions for personal data. Further, we deliver a supplementary analytical examination of the developed model and observation with various requirements presented both to the model and to the developed reference model itself, considering earlier decided and suggested requirements. This Chapter concludes with a critical discussion and review of the multifaceted results obtained in the context of evaluating the developed reference model.

In the context of the dissertation, we delivered five research questions. In the following, we debate the results of our research according to the research questions and objectives.

RQ1 *How can the different levels of GDPR anonymization for PET (DP) be interpreted?*

Although the GDPR provides a description of the available anonymization levels (Article 4 (1) and Recital 26 GDPR as complete anonymization and pseudo-anonymization), the proposed phrasing leaves considerable gaps for organizations (or data processors, third parties) and freedom to perform (including assessment) the mandated anonymization level. Simultaneously, the GDPR notations do not provide specific requirements (or clear instructions and methods) for executing the offered levels of anonymization or criteria for which it can be argued that the conditional level of anonymization has been achieved. In addition, the GDPR, referring to various TOMs conditions (Art. 24 GDPR), leaves the right for organizations to tete-a-tete, choose and demonstrate (including selecting technologies to accomplish the required level of anonymization). Thus, organizations are left without precise, transparent technical methods or criteria that can be applied in this case w.r.t GDPR. Our research [147] first ensures the complexity outlined before and secondly demonstrates the complexity of interpreting anonymization levels for PET, including examining the complexity in analyzing if DP[1] satisfies the GDPR regulation and the complexity in assessing such an issue. The results of our work [149] demonstrate that despite different levels of anonymization and issues in estimating them, many different stakeholders are involved and prefer to achieve complete anonymization instead of pseudonymization, including with the help of PET; nevertheless, it raises questions about communication of the results between referred stakeholders and further implications for the examination of anonymization level. Further, research and the results of our work [148], despite the confirmed additional complexity of the assessment of anonymization levels and whether DP complies with the GDPR requirements (which could be addressed and mitigated from various perspectives) indicated the possibility of interpreting the presented anonymization levels using an alternative approach (operating the so-called Data Protection Goals (DPG)). Additionally, our further effort and analysis in mitigating the lack of a single privacy taxonomy for DP [150] partly demonstrated the possibility of interpreting the presented anonymization levels using DPGs as one of the alternative ways of assessing the level of anonymization, which can also be applied to studied PET.

RQ2 *Which stakeholders and related objectives influence compliance, monitoring and compliance with GDPR?*

Although the GDPR introduces various entities (such as data processors, data controllers, third party and supervisory authorities, including DPOs) in the context of compliance with and implementation of the GDPR requirements (**Figure 2.2**), we recognized supplementary roles that are continuously or with varying

[1] in the current thesis, DP is related to the Differential Privacy

intensity involved in the process of compliance with the GDPR requirements relevant to personal data in organizations. Using the previously obtained research results [149], we supplemented the groups of roles in addition to those discussed in the GDPR, such as DPO, Privacy Legal Expert, Privacy Engineer and Software Developer. We also argue that despite the variety of specializations and different expert levels, the previously mentioned roles are involved in different combinations depending on various factors (organizational tasks, level of expertise). In other words, in the context of compliance with the GDPR requirements when working with personal data, the previously mentioned roles are directly involved in compliance with and implementation of the GDPR requirements, including the results of their assessment or other actions must be interpreted between the previously presented roles or groups of stakeholders in order to achieve the most transparent and effective implementation of the GDPR requirements.

RQ3 *What methods exist in the scientific literature/industry to achieve and comply with GDPR compliance standards (for reference models)?*

We conducted a broad search and analysis of similar or suitable approaches/methods to assess the degree of compliance with the GDPR both in academia (demonstrated in *Section 3.1*) and in industry (shown in *Section 3.2*). In addition, we conducted a more detailed and thorough investigation of similar applicable approaches and theoretical approaches in academia, taking into account alternative methodologies such as privacy patterns or privacy strategies. At the time of writing this thesis, we found that no approach or method was entirely satisfactory or appropriate for the PET under study. Despite this, the most suitable holistic approaches with exceptions were identified as the most similar and partly suitable for the research purposes in this thesis (SDM [163], PRIPARE [39] and Capability Model [109]). However, the referred approaches have several drawbacks presented and discussed in *Chapter 3*.

Despite the fact that examined approaches in the industry claim or are developed to address full compliance with the GDPR requirements, we concluded that only SDM (*Section 3.2.1*) mostly satisfies the declared statements. Furthermore, SDM developers continue the process of improving the referred framework[2] and optimizing various drawbacks declared by practitioners. Simultaneously, among other approaches investigated, only the SDM framework most widely and reasonably integrated the Data Protection Goals (DPG), which are clearly incorporated considering the GDPR, and SDM includes DPG such as DPG *Data Minimization*. Additionally, declared DPGs, as shown by SDM, are partly capable of bridging the gap between the legal and technical domains by addressing TOMs[3]. To this extent, DPGs were incorporated into the reference model being developed (*Section 4.1*). Further, SDM as a framework and declared DPG as introduced can be partly interpreted with country[4] security concepts[5] from the specific data protection risk focus. Also, we came to the conclusion that other considered industrial standards[6] only partially satisfy the DPG, including the fact that, according to the interview results, we found difficulties in integrating various industrial standards. Simultaneously, we came to the conclusion that among the analyzed approaches, only SDM can be considered up-to-date, but still, we did not find a confirmation[7] of the widespread use of the referred framework among organizations, as well as practical results among organizations in the context of adapting various PETs.

RQ4 *How can the GDPR compliance reference model (for DPG adaptation) be defined and presented?*

We have followed systematic and established principles devoted to developing and designing the reference model explained in *Section 4.1*. We have prepared and suggested eight diverse requirements (**R1**-**R4** declared in *Section 4.3.1* and **R5**-**R8**, shown in *Section 4.3.2*) for the developed reference model, which should enable the research objectives of this dissertation. The first four requirements (**R1, R2, R3, R4**)

[2] at the time of finalizing the dissertation SDM framework offered in v.3
[3] Technical and Organisational Measures, Art. 32 GDPR
[4] Germany
[5] BSI's IT-Grundschutz, https://www.bsi.de
[6] ISO/IEC
[7] or it is not available publicly due to the sensitivity of the topic

are general, and we characterize them as the main requirements that can be applied to the developed reference models. Also, we introduced four additional requirements (**R5, R6, R7, R8**) that should satisfy the requirements of compliance with the GDPR standards and also fill in the research gaps for the previously analyzed methods and approaches. We consider that the four additional requirements (**R5-R8**) highlight the peculiarity of the proposed reference model, which allows various stakeholders not only to step by step and effectively follow the requirements of the GDPR in the context of PET disposal and personal data processing but also allows them to effectively communicate the results of these processes and use a single taxonomy and vocabulary, thereby developing synergies in complying with the requirements of the GDPR and thereby providing an organization with a more rapid response to various challenges related to the protection of personal information, issues of assessing the level of anonymization and compliance with the requirements of data controllers. We also proposed a metamodel and blocks of the reference model (*Section 4.3.3*), as well as an additional privacy document (*Section 5.7*), which enables the utilization of the reference model by different organizations for diverse tasks, thereby using the developed template of the proposed model. Moreover, the visual representation (*Section 5.8*) of the reference model (notation-free and mind map approach) satisfies and allows diverse experts (with varying levels of expertise) to communicate the results and perform daily tasks between various stakeholders.

RQ5 *How do experts and practitioners evaluate aspects of the proposed compliance reference model for PET from various perspectives (feasibility, necessity, importance, economic and technical components)?*

To answer this question, we should first note that the scope of perspectives **RQ5** was slightly changed. However, the intentions of **RQ5** remain the same and further explanations and results are demonstrated. From the beginning of writing the current dissertation till finishing the final draft and analyzing the results, the perspectives initially addressed in **RQ5** have slightly changed and adopted due to the spin-offs obtained in various phases of the current research. Considering the designed choices of the developed artefact, initially mentioned perspectives were adopted to the more feasible criteria (demonstrated in *Section 6.1.1*), which enhanced current research and enabled more flexibility in assessing the developed artefact quality. To answer **RQ5**, we follow a multi-phase approach (explained in *Section 6.1*) by performing various steps and obtaining qualitative and quantitative data from diverse stakeholder groups defined earlier. The pre-development phase (**Table 6.1**) was supported by a critical analysis of the modelling requirements for reference models and an assessment of the model development according to the previously presented requirements for the reference model, resulting in more specific interview questions and further demonstrated results in *Section 6.3.2*. Further, we utilized the interview results to integrate the feedback[8] received from various experts and practitioners to refine the developed reference model. Refinement of the model according to the obtained results (*Chapter 6*) and the conducted analysis enable us to enhance the model's significance from the practitioner focus considering the various organization sizes and lead to the fulfilment of research goals of this dissertation. Remarkably, the selected stakeholder groups involved in the assessment of the quality of a developed artefact demonstrate satisfaction (**Table 6.6**); additionally, results indicate that there is no disagreement between groups and results obtained (**Figure 6.1**). We also want to provide interpretations of the results, considering the initial perspectives stated in **RQ5**. *Feasibility*: the developed model and incorporated concepts, as well as the dimension interconnection, enable transparent processes of fulfilment and recording results of the model execution and enable a more practical assessment of fulfilling GDPR requirements. Nevertheless, all groups indicated the need for more concise model templates[9]. *Necessity*: practitioners highlight that the developed model fits daily tasks and could be adopted considering organizations' needs. Moreover, some experts mentioned that such model design choices and logical interpretation not only enable easy communication of results between diverse stakeholders but also enhance it, fostering a more collaborative environment. Additionally, model could be utilized in education to explain complex GDPR notations more simply. Lastly, the model demonstrated

[8] also includes feedback from the RMPD Poster presented (Best Poster Award at https://www.iwsec.org/2024/index.html)
[9] additional templates for diverse use cases

benefit and traceability across all phases addressed with PDLC[10] or DPGs. *Importance*: in addition to the various recommendations from data protection authorities or industry standards, the model illustrates how such recommendations imposed by regulatory authorities could be demonstrated. *Economic*: the model could potentially reduce expenses for large organizations through unified templates applicable to diverse use cases or products, which also decreases the time for results communications between the stakeholders or business units, in addition considering small and medium organizations having less or no budget for such sophisticated assessment, the model provides added value and such organizations can benefit from the developed model. *Technical components*: practitioners note that organizations could benefit from the offered reference model, which introduces and adopts such concepts as PDLC together with DPGs, which supplement diverse tasks to validate or estimate the anonymization level applicable to the supplementary PET utilized and complex issues with privacy and data protection tasks. Nevertheless, they note that extending the model with a specific scoring function would increase the model's value.

7.2 Contribution to the State of the Art

The proposed and developed reference model established on best evidence synthesis, on the one hand, conceptualizes various knowledge and allows achieving the required level of compliance with the GDPR in the context of personal data processing in the context of PET use, focused on multiple stakeholders with different levels of expertise in the studied area. Unlike alternative and competing artefacts or approaches, the proposed reference model is developed taking into account the ontological basis, accepting the interdisciplinarity of the studied area, and levelling the shortcomings of existing and examined approaches. In addition, the proposed reference model accumulates various perspectives and results of various academics and practitioners, providing a consistent and relevant body of knowledge, and at the same time, expands the available and existing knowledge base in this and subsequent time intervals by integrating various privacy and data protection knowledge, aimed at meeting the requirements of the GDPR in the context of PET disposal. In other words, the reference model is based on various academic literature (accepting and critically evaluating different points of view for the studied area), including a large body of knowledge on privacy (about 30+ years), data protection (15+ years), security (40+ years) and conceptual and reference models (35+ years). The proposed reference model embodies and presents practical and theoretical knowledge from various experts in the studied area (data protection, privacy legal experts, software engineers and privacy engineers), which characterize diverse industries and organizations (of mixed sizes). The development process and the designed artefact are developed taking into account the DSR principles that are relevant and not limited to the following domains: researchers (DSR[11], IS[12], PbD[13], DPbD[14], DP[15], compliance, and data protection) and practitioners from the previously mentioned communities including PET[16] and DP.

For the *DP* research and practitioner *community*, accepting that the main body of research focused on enhancing the technical capabilities of DP applications. Our results assist and align with the identified interdisciplinary gaps in robust data protection introduced by Kobbi Nissim [136], with our specific intention for the PET DP, which partly contributes also to the "*need to co-design law and computer science*" and pointed legal challenges for PETs [66]. Firstly, the studies [148, 147, 149] debate and propose an additional perspective in assessing the anonymization level for PET considering GDPR goal-based regulation with DPGs. Furthermore, being interdisciplinary research, the study [150] studied existing privacy risks for various DP considering current at this time types of attacks and secondly suggested a way to estimate such privacy levels[17] for DP, which could further be applied for evaluating anonymization level.

[10] Privacy Development Life Cycle
[11] Design Science Research
[12] Information Systems
[13] Privacy by Design
[14] Data Protection by Default
[15] Differential Privacy
[16] Privacy Enhancing Technology
[17] with privacy risk-utility measures

For the *reference model* research community, the reference model represents the best evidence synthesis for managing the complexity of various approaches and integrating interdisciplinary knowledge, explaining the possibilities of assessing the degree of compliance for multiple organizations. This also authorizes us to assert that approaches to evaluate the degree of compliance in the past and in the future will increasingly differ for organizations utilizing personal data and also data controllers who are required to transparently estimate the fulfilment of various GDPR requirements for processing personal data.

For the *Privacy* community, the reference model represents the best evidence synthesis for governing the level of compliance with the GDPR, demonstrating the possibility of assessing the level of compliance with the GDPR using DPGs as one of the feasible supplementary and alternative approaches. In particular, the RMDP shows that the issues of assessing the level of compliance with the GDPR requirements imposed on personal data and the use of PET should be considered in a broader sense, and not only in terms of the effectiveness of technical approaches and assessing the data utility for PET. We acknowledge that the presented reference model is one of the first attempts to combine traditional and developed methods and concepts to evaluate the level of compliance with the GDPR for the use of PET, taking into account the different needs and requirements of various stakeholders with varying levels of expertise. Simultaneously, the demonstrated reference model offers an alternative view and perspective on the importance of DPGs in assessing anonymization levels and compliance with the GDPR.

For the *DSR* community, the RMDP reference model as a result of research activities according to the DSR principles demonstrates how actually the DSR principles can be practical in the context of interdisciplinary research. This includes aggregating and utilizing long-term accumulated data and experience (ontologies, reorientation, and capability formation) of the DSR performance for both the maturity model and the alternative models, which are debated in *Chapter 3*. The knowledge and observations obtained enable us to obtain various multifaceted knowledge about the mechanisms of accumulation and application of knowledge for forthcoming challenges.

For various *practitioners*, the reference model represents both descriptive and prescriptive knowledge about meeting the requirements of the GDPR in the context of PET utilization for personal data. Considering the developed reference model as a descriptive artefact, RMDP synthesizes interdisciplinary knowledge for knowledge communication, training and assessment of the degree of compliance and personal data transformation activities according to the DPG for PET (i.e. the use and application of the artefact). As a prospective artefact, the reference model provides the necessary interdisciplinary knowledge about PDLC for creating, maintaining or improving the multidisciplinary knowledge and the capability of organizations to complete various requirements of the GDPR for multiple tasks and goals of organizations that play an active role at present. At the same time, the developed and proposed reference model enables configuration, supplementation, and customization of the model to specific needs. Including depending on dynamic factors, thereby, after adaptation and specification of the model, the reference model allows you to meet the specific requirements and goals of the organization (in some cases, the use of different technologies). The various use cases presented in the thesis reflect and illustrate the changing role of personal data, the process of assessing compliance with the GDPR requirements and what consequences organizations may experience given the various use cases and the current GDPR requirements for personal data.

7.3 Study Limitations

The results of this study presented in this dissertation have various limitations. Referred limitations are primarily related to but not limited to the following remarks with the research context, research design, and interdisciplinary context, as well as feasible ambiguities that could arise during the study or the results of other studies utilized as results for this study. Simultaneously, the validity of this work depends on the validity of the materials considered and used in the development, design and composition of the reference model while accepting the additional opinions and feedback of the various expert groups that participated in the evaluation of the developed artefact. Further, we will provide more detailed comments in the context of the previously mentioned limitations.

Initially, it is worth noting the interdisciplinary nature of the study since the data protection domain represents and belongs to the intersection of simultaneously following research domains (such as legal sciences, computer sciences, IS research, management and others that to one degree or another impact the GDPR) or, in other words, at the junction of various domains. Saying differently, interdisciplinarity and complexity lie in the fact that the issues under study affect from technical implementations and implementations to organizational ones, thereby complicating the use and adaptation of various sources of data and knowledge included in this research project and dissertation. In other words, it can be argued that the list of literature and knowledge used can be considered limited or partial. Despite this, we utilized various alternatives for searching and validating components for research questions. In other words, we can declare that reverse and direct searches were used to identify relevant academic works in the area under study. The selection and search for relevant publications were carried out by analyzing academic literature, as well as by analyzing the provided links and comments; in addition, we analyzed various grey sources or aggregated press releases of various privacy agencies[18] or comments or other sources published by various ICOs in the EU[19] and UK[20]. It is worth noting that in some cases, during extended interviews and the exchange of expert opinions with various participants, the source of recommendations or knowledge was interview participants or industry representatives during the presentation of intermediate results during the project implementation. We also want to note that the developed artefact is based on the current recommendations of the GDPR, which was launched and applied in 2018 and still remains one of the central laws aimed at protecting personal data in the EU. In addition, it is worth noting that although the interdisciplinary research work considered related disciplines that affect the GDPR, we do not exclude that the inclusion of materials or knowledge, for example, from the security domain and not limited to this area (including experts from this area) could shift the focus of the study or represent a bias with the most significant influence of the data protection and privacy domains. Therefore, the inclusion of additional areas related to the GDPR may lead to a change in the perspective of the developed reference model.

Secondly, qualitative research is limited by the selection of participants and organizations. Despite the fact that the primary preference (for selecting candidates) was based on competence, as well as the level of expert qualification verified by both professional experience and respective certifications. We understand that the data could be recalled incorrectly or with shortcomings. In particular, we believe that the knowledge reflected in the developed reference model can be determined by both the choice of organizations and the choice of candidates at various phases of the current research project. However, according to the principles of candidate selection, we attempted to include organizations of various sizes in qualitative research, including those from diverse industries. Simultaneously, it can be considered that the results are objective and accurate or appropriate for the examined industries; it is likely that they can be interpreted for other industries not included in the selected sample. Considering that the research design is largely based on qualitative data, it can be argued that the offered research results are reliably and analytically accurate, but the quantitative or statistical component can be disputed or, in other words, cannot be demonstrated. In other words, we would like to point out that the results should be treated with caution, and in order to obtain more accurate results, we recommend or require supplementary experiments with a more significant number of participants and a greater variety of organizations in the future. In addition, we would like to point out that the most significant part of the respondents, including the distribution of invitations to participate in the study and the analysis of the reference model, occurred from organizations in Germany (although it is worth noting that large organizations are representatives of international organizations in Germany). We suppose that the respondents could have a different perception and interpretation of the generally recognized European legislation GDPR, which could be influenced by national legislation (e.g. Bundesdatenschutzgesetz or other national ones). On the other hand, the results indicated that we did not record any bias relative to other EU countries. Thus, we would like to point out that despite the preponderance of candidates from Germany, we consider that the prospects for data protection of personal data in

[18] European Data Protection Board (EDPB), Federal Commissioner for Data Protection and Freedom of Information (BfDI), Commission Nationale de l'Informatique et des Libertés (CNIL), partly EU agency for cybersecurity (ENISA) and National Institute of Standards and Technology (NIST)

[19] Bayerisches Landesamt für Datenschutzaufsicht (BayLDA)

[20] Information Commissioner's Office (ICO)

the proposed reference model are not violated and are applicable to international standards, and the proposed reference model could be utilized in other EU countries (or where GDPR needs to be applied). We also want to note that some of the participants held diverse functions and performed diverse roles. Thus, we want to note that there is a possibility that other candidates (who did not participate, refused to take part in the evaluation or did not finalize the evaluation) could have assessed the reference model differently. In addition, it is worth noting that the results of the interviews could also be influenced by the following factors such as English language proficiency, including lack of necessary familiarity with the terminology, concepts presented to the interview participant and limited time for presenting the results.

Further, *thirdly*, we want to remark on the limitations associated with the reference model such statements presented earlier as the reference model represents a) application domain (universality) and on the other hand, suggests a blueprint reference model for research goal outlined earlier (recommendation character), Despite the fact that both of these statements have been critically discussed in the academic literature, the author of this dissertation, including other users of this reference model may not agree that the proposed reference model is a reference model or in other words have a different opinion about the name of this artefact. The author and designer of the developed reference model reserve the right to call the developed artefact a reference model and leave freedom in choosing names or defining the user by this artefact. Additionally, we accept that the presented reference model, like other reference models, may have a limited validity period (associated not only with various progress in various domains and technologies but also with the adaptation of legal requirements in the EU territory), but we accept that the presentation of the following and revised version of the reference model can mitigate this limitation.

In *conclusion*, it is worth noting the possible limitations in the evaluating phase of the reference model presented in *Chapter 6*. The evaluation of the reference model is complex due to methodological, philosophical and practical aspects. There are various frameworks for assessing the quality of the reference model and its declared practicality. We used diverse approaches to evaluate the quality of the reference model, but it is worth noting that the possibility of assessing the quality of the reference model is not limited and can be supplemented by other methods. A separate threat to the validity of the reference model may be the lack or, in other words, the possibility of practical assessment of the developed reference model. However, the results of the model assessment in the academic community provide an opportunity to declare that the debate about the validity of the model is limited; in addition, these limitations can also apply to the model due to the use of various concepts integrated into the model.

7.4 Reflection, future research and enhancements

We want to note that among the presented goals (*Section 1.5*), the purpose of the current study or any other research projects or dissertations is not limited only to enhancing and developing the notably recognized gaps but also to elaborate on how further researchers or contributors could enhance the current study or received results and introduced uncertainties, providing guidance and awareness for more feasible generalization. We consider that the outcomes and ideas presented and developed in the current study can be a beginning for further research dedicated to various interdisciplinary topics of privacy and data protection, considering diverse PETs and further implications in estimating privacy risk in sophisticated systems or other practical applications. A current section is devoted to the issues related to observed limitations and enhancements applications and also incorporates future work dedicated to matters referred.

This dissertation demonstrates the development process of the reference model and also illustrates the design choices for each concept chosen, and explains the interconnections between the studied domains as stated in the beginning research questions. The primary artefact of the current dissertation is a developed and offered reference model (*Chapter 5*) evaluated with various stakeholders (*Chapter 6*), supplemented with quality analysis of the provided artefact. Looking back at the time of sketching and drawing the whole pipeline of current dissertations (**Figure 1.1**), it was not expected that the demonstrated interdisciplinary topic would demand more effort and could be the subject of documented research spin-offs due to the complexity of choosing appropriate methods and methodologies as well as the approaches to validate the results developed. Also, the results indicate that developed artefacts could be applied in var-

ious organizations or industries and for data protection and compliance tasks; it is necessary to note that the developed model should be adjusted precisely to meet the needs of the organizations to be executed with intended efficiency. Nevertheless, by writing down the final sections and reflecting on the delivered results, the chosen way of conducting performed research could be highlighted with faced limitations (*Section 7.3*) that could be translated and outlined as future work. Further, we provide more detailed comments regarding the implications and future work.

Firstly, we primarily address the constraints and limitations. As we remarked in *Section 5.2* and *Section 5.9.1*, the proposed reference model (**Figure 5.1**), like other conceptual or reference models, has an expiration date and requires constant revisions or updates. However, such refinement depends and is not limited to the development phase of various factors in technology and policy development; thus, it is difficult to specify a dedicated period for the next improvement. Nevertheless, time is a key component to mitigate the possible privacy implications, which also partly aligns with Article 32 (1) GDPR[21] and demonstrated in *Sections 5.1*. Accordingly, in the future, the author of this dissertation, with the valuable input and feedback from academic peers, researchers, and experts in data protection and privacy, intends to update the offered model, taking into account the dynamics and current state of developing concepts (definitions, impacts) and the landscape in data protection, privacy and other closely related to the GDPR domains, due to the high dependency of the concepts incorporated for each dimension (*Section 5.2-5.5*). Moreover, considering the ongoing reexamination of academic literature, including the results and reports of various industrial standards and practices integrations, the developed model and integrated concepts are subject to be adjusted, including summarised inferences (*Section 5.9.1-5.9.2*). Similarly, the supplementary observations received (*Chapter 6*) and the analysis of the developed reference model's evaluation results will be considered and incorporated in the next version of the reference model.

Secondly, future research enhancements and steps, considering the evaluation results presented earlier (in the current dissertation), could be explained further mainly in six directions[22], supplemented with our comments:

- developing proposed PDLC[23] and further mapping results of each stage with generic SDLC;
- extending Information Taxonomy categories;
- developing a cross-walk between established DPGs dimension and other legal regulations (i.e. US[24], EU[25]);
- extending privacy countermeasures, supplementing them with possible causes or time intervals, including the development of specifications and definition of a set of catalogs for each group of privacy countermeasures (e.g. NIST Privacy framework);
- expanding the functionality of the proposed Privacy Document due to specialization for different roles and monitoring the implementation of countermeasures and the adoption of amendments and their coordination between different stakeholders[26];
- developing a local and global scoring function for automatization of the developed reference model[27];

Thirdly, we want to stress the application aspect of the developed reference model. On the one hand, the reference model's concept and objectives encourage stakeholders with different levels of expertise to utilize and interpret various facts and take necessary steps in operating knowledge in the data protection domain,

[21] *a procedure for **regularly** testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing*

[22] it is also important to note that demonstrated further enhancements could be depended or related to each other

[23] Privacy Development Life Cycle

[24] CCPA, HIPPA

[25] DSA, AI act

[26] RACI Matrix: Responsible, Accountable, Consulted, Informed

[27] assessment of anonymization level considering[150]

PET and GDPR legal provisions. Which enables referred stakeholders to encounter the most feasible way to accomplish data protection goals considering the GDPR necessities for various organizations and purposes. On the other hand, the developed reference model does not deliver precise recommendations or, as some interviewees stated, there are no different reference model templates for diverse use cases or detailed step-by-step instructions for solving various issues with estimating compliance with the GDPR requirements or data protection issues, which create additional effort on the users of the reference model. This points out additional steps aimed to improve the execution of the offered reference model by utilizing an alternative approach which optimal DPGs to choose or use to determine DPG goals. In addition, the practical implementation of the assessment of referred goals (in the DPG dimension) can be carried out using various prompts in the context of the adaptation of specialized LLMs[28] concentrated on data protection or privacy audits [11, 208]; however, and the challenges regarding the cost of training, fine-tuning and developing such models remains challenging.

Further, during the development and evaluation of the reference model presented in this dissertation, the author and designer of the primary artefact received multiple requests from diverse organizations to update the offered reference model for the needs of advanced digital assistant[29] that would provide recommendations considering the implementation PDLC and offered DPG for demonstrated groups of stakeholders tailored to the specific company units and needs. Where such integrations are aimed to enhance compliance with the GDPR and reduce the amount and costs for dedicated privacy and data protection training. We consider that the development of such advanced assistant based on contextual integrity [68] in the case of aggregation and analysis of various use cases, including enriched information on possible risks from expert assessments, w.r.t possible historical data or accumulation of such data using a separate recommendation service, will enable obtaining considerable different facts for subsequent testing and validation in relation to the adopted recommendations (i.e. choice of DPGs and further privacy countermeasures). However, many other experiments will be required first to adopt the presented model and secondly to normalize and evaluate the quality of such assistant and offered service.

Lastly, we also want to note that the visual representation of the delivered reference model was also one of the aspects of the attention of various experts in the data protection domain. Further investigation is required to enhance the visual representation of the reference model and the components of the model, including the display of the process of using the reference model for different stakeholders. We believe that different approaches to cognitive perception or visual comprehension [168] can allow the reference model to be more effectively understood by various stakeholders. Overall, we hope our suggestions, additional proposals, endeavours and results obtained related to the technical and legal domain will encourage researchers to continue developing and adopting the offered reference model for diverse tasks.

In conclusion, it is worth noting that the results received in this dissertation laid the foundation for enhancing the approach for evaluating the degree of compliance assessment for various types and sizes of organizations in the context of the GDPR requirements, based on the DPG when adapting PET (not limited to DP only). In other words, we can declare that the basis was laid for the development of methods for assessing compliance with the GDPR both for PET and for other possible interpretations and adapted technologies. It is worth noting that various degrees of improvement of methods are possible, which may relate to and are not limited to various methods of assessing the DPG, defining the taxonomy and concept of privacy, developing or adapting various privacy countermeasures and their relationships with each other, supplementing the PDLC and interpreting and supplementing the Privacy Document, taking into account the specifics and progress of adapting legal requirements not only for the GPDR. Further, we argue that the results of this dissertation will allow the production of various automation and optimizations or other products using the primary artefact. Additionally, using the presented results to adapt industrial approaches or develop new industrial standards, taking into account the delivered results, can be achieved by studying the contributions, mechanisms and communication of integrated concepts in the primary artefact.

[28] Large Language Models

[29] AI assistant as an artificial agent with a natural language interface, the function of which is to plan and execute sequences of actions on the user's behalf across one or more domains and in line with the user's expectations[63]

Bibliography

[1] Abdullah, N.S., Indulska, M., Shazia, S.: A study of compliance management in information systems research (2009)

[2] Adams, W.C.: Conducting semi-structured interviews. Handbook of practical program evaluation pp. 492–505 (2015)

[3] Al-Mayahi, I., Sa'ad, P.M.: Iso 27001 gap analysis-case study. In: Proceedings of the International Conference on Security and Management (SAM). p. 1. The Steering Committee of The World Congress in Computer Science, Computer . . . (2012)

[4] del Alamo, J.M., Martín, Y.S., Caiza, J.C.: Towards organizing the growing knowledge on privacy engineering. In: IFIP International Summer School on Privacy and Identity Management, pp. 15–24. Springer (2017)

[5] Alshammari, M., Simpson, A.: Personal data management: an abstract personal data lifecycle model. In: Business Process Management Workshops: BPM 2017 International Workshops, Barcelona, Spain, September 10-11, 2017, Revised Papers 15. pp. 685–697. Springer (2018)

[6] Alter, S.: Defining information systems as work systems: implications for the is field. European journal of information systems 17(5), 448–469 (2008)

[7] Altman, M., Cohen, A., Falzon, F., Markatou, E.A., Nissim, K., Reymond, M.J., Saraogi, S., Wood, A.: A principled approach to defining anonymization as applied to eu data protection law. Kobbi and Reymond, Michel Jose and Saraogi, Sidhant and Wood, Alexandra, A Principled Approach to Defining Anonymization As Applied to EU Data Protection Law (May 9, 2022) (2022)

[8] Andenmatten, M.: It-qualität durch service management. HMD Praxis der Wirtschaftsinformatik 53(2), 185–199 (2016)

[9] Andenmatten, M.: It4ittm-das agile betriebskonzept der it der zukunft. HMD Prax. Wirtsch. 54(2), 261–274 (2017)

[10] Aridor, G., Che, Y.K., Salz, T.: The economic consequences of data privacy regulation: Empirical evidence from gdpr. National Bureau of Economic Research Cambridge, MA, USA (2020)

[11] Aydin, I., Diebel-Fischer, H., Freiberger, V., Möller-Klapperich, J., Buchmann, E., Färber, M., Lauber-Rönsberg, A., Platow, B.: Assessing privacy policies with ai: Ethical, legal, and technical challenges. arXiv preprint arXiv:2410.08381 (2024)

[12] Baldassarre, M.T., Barletta, V.S., Caivano, D., Scalera, M.: Integrating security and privacy in software development. Software Quality Journal 28, 987–1018 (2020)

[13] Becker, J., Delfmann, P., Dreiling, A., Knackstedt, R., Kuropka, D.: Configurative process modeling–outlining an approach to increased business process model usability. In: Proceedings of the 15th IRMA International Conference. pp. 1–12. Gabler New Orleans (2004)

[14] Becker, J., Probandt, W., Vering, O.: Grundsätze ordnungsmäßiger Modellierung: Konzeption und Praxisbeispiel für ein effizientes Prozessmanagement. Springer-Verlag (2012)

[15] Becker, J., Rosemann, M., Von Uthmann, C.: Guidelines of business process modeling. In: Business Process Management: Models, Techniques, and Empirical Studies, pp. 30–49. Springer (2002)

[16] Becker, R., Thorogood, A., Bovenberg, J., Mitchell, C., Hall, A.: Applying gdpr roles and responsibilities to scientific data sharing. International Data Privacy Law 12(3), 207–219 (2022)

[17] Bélanger, F., Crossler, R.E.: Privacy in the digital age: a review of information privacy research in information systems. MIS quarterly pp. 1017–1041 (2011)

[18] Berghel, H.: Malice domestic: The cambridge analytica dystopia. Computer 51(05), 84–89 (2018)

[19] Bernau, D., Eibl, G., Grassal, P.W., Keller, H., Kerschbaum, F.: Quantifying identifiability to choose and audit ϱ in differentially private deep learning. In: Proceedings of the Conference on Very Large Databases (2021)

[20] Bichler, M., Frank, U., Avison, D., Malaurent, J., Fettke, P., Hovorka, D., Krämer, J., Schnurr, D., Müller, B., Suhl, L., et al.: Theories in business and information systems engineering. Business & Information Systems Engineering 58, 291–319 (2016)

[21] Blumer, A., Ehrenfeucht, A., Haussler, D., Warmuth, M.K.: Occam's razor. Information processing letters 24(6), 377–380 (1987)

[22] Bowen, P., Hash, J., Wilson, M.: Information security handbook: a guide for managers. Tech. rep., National Institute of Standards and Technology (2006)

[23] Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative research in psychology 3(2), 77–101 (2006)

[24] vom Brocke, J.: Design principles for reference modeling: reusing information models by means of aggregation, specialisation, instantiation, and analogy. In: Reference modeling for business systems analysis, pp. 47–76. IGI Global (2007)

[25] Bygrave, L.A.: Data protection by design and by default: deciphering the eu's legislative requirements. Oslo Law Review 4(2), 105–120 (2017)

[26] Caiza, J.C., Martín, Y.S., Del Alamo, J.M., Guamán, D.S.: Organizing design patterns for privacy: a taxonomy of types of relationships. In: Proceedings of the 22nd European Conference on Pattern Languages of Programs. pp. 1–11 (2017)

[27] Calder, A.: IT governance: A manager's guide to data security and ISO 27001/ISO 27002. Kogan Page Publishers (2008)

[28] Cantiello, P., Mastroianni, M., Rak, M.: A conceptual model for the general data protection regulation. In: International Conference on Computational Science and Its Applications. pp. 60–77. Springer (2021)

[29] Cavoukian, A.: Privacy by design in law. Policy and Practice (2011)

[30] Cavoukian, A., Dixon, M.: Privacy and security by design (2013)

[31] Cavoukian, A., Taylor, S., Abrams, M.E.: Privacy by design: essential for organizational accountability and strong business practices. Identity in the Information Society 3, 405–413 (2010)

[32] Cavoukian, A., et al.: Privacy by design: The 7 foundational principles. Information and privacy commissioner of Ontario, Canada 5, 12 (2009)

[33] Cha, S.C., Hsu, T.Y., Xiang, Y., Yeh, K.H.: Privacy enhancing technologies in the internet of things: Perspectives and challenges. IEEE Internet of Things Journal 6(2), 2159–2187 (2018)

[34] Cherdantseva, Y., Hilton, J.: A reference model of information assurance & security. In: 2013 international conference on availability, reliability and security. pp. 546–555. IEEE (2013)

[35] Clements, T., Milton, S.: Maintaining data protection and privacy beyond gdpr implementation. Rolling Meadows, IL: ISACA (2018)

[36] Cleven, A., Winter, R.: Regulatory compliance in information systems research–literature analysis and research agenda. In: International Workshop on Business Process Modeling, Development and Support. pp. 174–186. Springer (2009)

[37] Corbin, J.M., Strauss, A.: Grounded theory research: Procedures, canons, and evaluative criteria. Qualitative sociology 13(1), 3–21 (1990)

[38] Cormode, G.: Personal privacy vs population privacy: learning to attack anonymization. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 1253–1261 (2011)

[39] Crespo, A., Notario, N., Troncoso, C., Le Métayer, D., Kroener, I., Wright, D., Del Alamo, J.M., Martin, Y.S.: Pripare privacy-and security-by-design methodology handbook. Tech. rep., Technical report, 2015. URL http://pripareproject. eu (2015)

[40] Crnkovic, G.D.: Constructive research and info-computational knowledge generation. In: Model-Based Reasoning in Science and Technology: Abduction, Logic, and Computational Discovery, pp. 359–380. Springer (2010)

[41] Cummings, R., Desai, D.: The role of differential privacy in gdpr compliance. In: FAT'18: Proceedings of the Conference on Fairness, Accountability, and Transparency. p. 20 (2018)

[42] Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.H., Metayer, D.L., Tirtea, R., Schiffner, S.: Privacy and data protection by design-from policy to engineering. arXiv preprint arXiv:1501.03726 (2015)

[43] De, S.J., Le Métayer, D.: Priam: a privacy risk analysis methodology. In: Data Privacy Management and Security Assurance: 11th International Workshop, DPM 2016 and 5th International Workshop, QASA 2016, Heraklion, Crete, Greece, September 26-27, 2016, Proceedings 11. pp. 221–229. Springer (2016)

[44] De Hert, P., Papakonstantinou, V., Kamara, I.: The cloud computing standard iso/iec 27018 through the lens of the eu legislation on data protection. Computer Law & Security Review 32(1), 16–30 (2016)

[45] DeMarco, T.: Structured analysis and system specification. In: Software pioneers: contributions to software engineering, pp. 529–560. Springer (2011)

[46] Deng, M., Wuyts, K., Scandariato, R., Preneel, B., Joosen, W.: A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requirements Engineering 16(1), 3–32 (2011)

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

[48] Directive, E.: 95/46. EC of the European Parliament and of the Council of 24, 31 (1995)

[49] Dwork, C.: Differential privacy. In: International colloquium on automata, languages, and programming. pp. 1–12. Springer (2006)

[50] Dwork, C.: Differential privacy: A survey of results. In: International conference on theory and applications of models of computation. pp. 1–19. Springer (2008)

[51] Dwork, C., Roth, A., et al.: The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science 9(3–4), 211–407 (2014)

[52] El Kharbili, M.: Business process regulatory compliance management solution frameworks: A comparative evaluation. In: Proceedings of the Eighth Asia-Pacific Conference on Conceptual Modelling-Volume 130. pp. 23–32 (2012)

[53] Fedorowicz, J., Gelinas, U., Ulric, J.: Adoption and usage patterns of cobit: results from a survey of cobit purchasers. IS Audit and Control Journal pp. 45–51 (1998)

[54] Fettke, P., Loos, P.: Classification of reference models: a methodology and its application. Information systems and e-business management 1, 35–53 (2003)

[55] Fettke, P., Loos, P.: Multiperspective evaluation of reference models–towards a framework. In: International Conference on Conceptual Modeling. pp. 80–91. Springer (2003)

[56] Fettke, P., Loos, P.: Perspectives on reference modeling. In: Reference modeling for business systems analysis, pp. 1–21. IGI Global (2007)

[57] Finck, M., Pallas, F.: They who must not be identified—distinguishing personal from non-personal data under the gdpr. International Data Privacy Law 10(1), 11–36 (2020)

[58] Force, J.T.: Risk management framework for information systems and organizations: A system life cycle approach for security and privacy (discussion draft). Tech. rep., National Institute of Standards and Technology (2017)

[59] Fossey, E., Harvey, C., McDermott, F., Davidson, L.: Understanding and evaluating qualitative research. Australian & New Zealand journal of psychiatry 36(6), 717–732 (2002)

[60] Frank, U.: Evaluation of reference models. In: Reference modeling for business systems analysis, pp. 118–140. IGI Global (2007)

[61] Frank, U., Strecker, S., Koch, S.: " open model"-ein vorschlag für ein forschungsprogramm der wirtschaftsinformatik: Langfassung. Tech. rep., ICB-Research Report (2007)

[62] Friedewald, M., Schiering, I., Martin, N., Hallinan, D.: Data protection impact assessments in practice: Experiences from case studies. In: European Symposium on Research in Computer Security. pp. 424–443. Springer (2021)

[63] Gabriel, I., Manzini, A., Keeling, G., Hendricks, L.A., Rieser, V., Iqbal, H., Tomašev, N., Ktena, I., Kenton, Z., Rodriguez, M., et al.: The ethics of advanced ai assistants. arXiv preprint arXiv:2404.16244 (2024)

[64] Garfinkel, S., et al.: De-identification of Personal Information:. US Department of Commerce, National Institute of Standards and Technology (2015)

[65] Garrido, G.M., Liu, X., Matthes, F., Song, D.: Lessons learned: Surveying the practicality of differential privacy in the industry. arXiv preprint arXiv:2211.03898 (2022)

[66] Garrido, G.M., Sedlmeir, J., Uludağ, Ö., Alaoui, I.S., Luckow, A., Matthes, F.: Revealing the landscape of privacy-enhancing technologies in the context of data markets for the iot: A systematic literature review. Journal of Network and Computer Applications 207, 103465 (2022)

[67] Gaulke, M.: Praxiswissen COBIT: Grundlagen und praktische Anwendung in der Unternehmens-IT. Geeignet als Vorbereitung auf die ISACA-Prüfungen: COBIT Foundation, IT-Governance & IT-Compliance Practitioner, IT-Governance-Manager, IT-Compliance-Manager, CGEIT. dpunkt. verlag (2019)

[68] Ghalebikesabi, S., Bagdasaryan, E., Yi, R., Yona, I., Shumailov, I., Pappu, A., Shi, C., Weidinger, L., Stanforth, R., Berrada, L., et al.: Operationalizing contextual integrity in privacy-conscious assistants. arXiv preprint arXiv:2408.02373 (2024)

[69] Goldberg, I., Wagner, D., Brewer, E.: Privacy-enhancing technologies for the internet. In: Proceedings IEEE COMPCON 97. Digest of Papers. pp. 103–109. IEEE (1997)

[70] Gonzalez-Perez, C., Henderson-Sellers, B.: A powertype-based metamodelling framework. Software & Systems Modeling 5, 72–90 (2006)

[71] Grance, T., Hash, J., Stevens, M.: Security considerations in the information system development life cycle. US Department of Commerce, Technology Administration, National Institute of . . . (2004)

[72] Green, R.E.: The persuasive properties of color. Marketing Communications (1989)

[73] Gruschka, N., Mavroeidis, V., Vishi, K., Jensen, M.: Privacy issues and data protection in big data: a case study analysis under gdpr. In: 2018 IEEE International Conference on Big Data (Big Data). pp. 5027–5033. IEEE (2018)

[74] Guldentops, E., Van Grembergen, W., De Haes, S.: Control and governance maturity survey: establishing a reference benchmark and a self assessment tool. Information systems control journal.-Rolling Meadows, Ill. 6, 32–35 (2002)

[75] Hadar, I., Hasson, T., Ayalon, O., Toch, E., Birnhack, M., Sherman, S., Balissa, A.: Privacy by designers: software developers' privacy mindset. Empirical Software Engineering 23, 259–289 (2018)

[76] Hallinan, D.: Data protection impact assessments in practice. In: Computer Security. ESORICS 2021 International Workshops: CyberICPS, SECPRE, ADIoT, SPOSE, CPS4CIP, and CDT&SECOMANE, Darmstadt, Germany, October 4–8, 2021, Revised Selected Papers. vol. 13106, p. 424. Springer Nature (2022)

[77] Hamri, M.M., Benslimane, S.M.: Building an ontology for the metamodel iso/iec24744 using mda process. International Journal of Modern Education and Computer Science 7(8), 48 (2015)

[78] Hansen, M., Jensen, M., Rost, M.: Protection goals for privacy engineering. In: 2015 IEEE Security and Privacy Workshops. pp. 159–166. IEEE (2015)

[79] Henderson-Sellers, B., Gonzalez-Perez, C.: Granularity in conceptual modelling: application to metamodels. In: Conceptual Modeling–ER 2010: 29th International Conference on Conceptual Modeling, Vancouver, BC, Canada, November 1-4, 2010. Proceedings 29. pp. 219–232. Springer (2010)

[80] Herath, S., Gelman, H., McKee, L.: Privacy harm and non-compliance from a legal perspective. Journal of Cybersecurity Education, Research and Practice 2023(2), 3 (2023)

[81] Hernon, P.: Information life cycle: Its place in the management of us government information resources. Government Information Quarterly 11(2), 143–170 (1994)

[82] Hevner, A.R.: A three cycle view of design science research. Scandinavian journal of information systems 19(2), 4 (2007)

[83] Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS quarterly pp. 75–105 (2004)

[84] Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. Management Information Systems Quarterly 28(1), 6 (2008)

[85] Hintze, M.: Viewing the gdpr through a de-identification lens: a tool for compliance, clarification, and consistency. International Data Privacy Law 8(1), 86–101 (2018)

[86] Hoda, R., Noble, J., Marshall, S.: Grounded theory for geeks. In: Proceedings of the 18th conference on pattern languages of programs. pp. 1–17 (2011)

[87] Hoepman, J.H.: Privacy design strategies. In: IFIP International Information Security Conference. pp. 446–459. Springer (2014)

[88] Holzel, J.: Differential privacy and the gdpr. Eur. Data Prot. L. Rev. 5, 184 (2019)

[89] Humphreys, E.: Implementing the ISO/IEC 27001: 2013 ISMS Standard. Artech house (2016)

[90] Huth, D., Faber, A., Matthes, F.: Towards an understanding of stakeholders and dependencies in the eu gdpr. Proceedings of the MKWI 2018 pp. 338–344 (2018)

[91] Ikhsan, M., Widodo, A.P., Adi, K.: Systematic literature review on corporate information technology governance in indonesia using cobit 2019. Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram 9(2), 354–364 (2021)

[92] Initiative, J.T.F.T., et al.: SP 800-37 Rev. 1. Guide for Applying the Risk Management Framework to Federal Information Systems: a Security Life Cycle Approach. National Institute of Standards & Technology (2010)

[93] Irti, C.: Personal data, non-personal data, anonymised data, pseudonymised data, de-identified data. Privacy and Data Protection in Software Services pp. 49–57 (2022)

[94] Järvelin, K., Wilson, T.D.: On conceptual models for information seeking and retrieval research. Information research 9(1), 9–1 (2003)

[95] Jin, H., Shen, H., Jain, M., Kumar, S., Hong, J.I.: Lean privacy review: Collecting users' privacy concerns of data practices at a low cost. ACM Transactions on Computer-Human Interaction (TOCHI) 28(5), 1–55 (2021)

[96] Josey, A.: The IT4ITTM Reference Architecture, Version 2.1–A Pocket Guide. Van Haren (2017)

[97] Josey, A., et al.: The it4it™ reference architecture, version 2.0– (2015)

[98] Kabanov, I.: Effective frameworks for delivering compliance with personal data privacy regulatory requirements. In: 2016 14th Annual Conference on Privacy, Security and Trust (PST). pp. 551–554. IEEE (2016)

[99] Karampela, M., Ouhbi, S., Isomursu, M.: Exploring users' willingness to share their health and personal data under the prism of the new gdpr: implications in healthcare. In: 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC). pp. 6509–6512. IEEE (2019)

[100] Klymenko, O., Meisenbacher, S., Matthes, F.: Identifying practical challenges in the implementation of technical measures for data privacy compliance. arXiv preprint arXiv:2306.15497 (2023)

[101] Knoke, F., Nwankwo, I.: Managing data protection compliance through maturity models: A primer. Eur. Data Prot. L. Rev. 8, 536 (2022)

[102] Koç, H., Eckert, K., Flaig, D.: Datenschutzgrundverordnung (dsgvo): Bewältigung der herausforderungen mit unternehmensarchitekturmanagement (eam). HMD Praxis der Wirtschaftsinformatik 55(5), 942–963 (2018)

[103] Kohlegger, M., Maier, R., Thalmann, S.: Understanding maturity models. Results of a structured content analysis. Citeseer (2009)

[104] Kolain, M., Grafenauer, C., Ebers, M.: Anonymity assessment-a universal tool for measuring anonymity of data sets under the gdpr with a special focus on smart robotics. Rutgers Computer & Tech. LJ 48, 174 (2021)

[105] Koops, B.J., Leenes, R.: Privacy regulation cannot be hardcoded. a critical comment on the 'privacy by design'provision in data-protection law. International Review of Law, Computers & Technology 28(2), 159–171 (2014)

[106] Kostova, B., Gürses, S., Troncoso, C.: Privacy engineering meets software engineering. on the challenges of engineering privacy bydesign. arXiv preprint arXiv:2007.08613 (2020)

[107] Kozyreva, A., Lorenz-Spreen, P., Hertwig, R., Lewandowsky, S., Herzog, S.M.: Public attitudes towards algorithmic personalization and use of personal data online: Evidence from germany, great britain, and the united states. Humanities and Social Sciences Communications 8(1), 1–11 (2021)

[108] Kurtz, C., Semmann, M., Böhmann, T.: Privacy by design to comply with gdpr: a review on third-party data processors (2018)

[109] Labadie, C., Legner, C.: Understanding data protection regulations from a data management perspective: a capability-based approach to eu-gdpr. In: Proceedings of the 14th International Conference on Wirtschaftsinformatik (2019) (2019)

[110] Labadie, C., Legner, C.: Personal data protection inside and out: Integrating data protection requirements in the data lifecycle. Enterprise Modelling and Information Systems Architectures (EMISAJ) 15, 9–1 (2020)

[111] Labadie, C., Legner, C.: Building data management capabilities to address data protection regulations: Learnings from eu-gdpr. Journal of Information Technology 38(1), 16–44 (2023)

[112] Lachaud, E.: Iso/iec 27701 standard: Threats and opportunities for gdpr certification. Eur. Data Prot. L. Rev. 6, 194 (2020)

[113] Lenhard, J., Fritsch, L., Herold, S.: A literature study on privacy patterns research. In: 2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA). pp. 194–201. IEEE (2017)

[114] Lester, P.M.: Syntactic theory of visual communication. Retrieved December 3, 2010 (2006)

[115] Libert, T.: An automated approach to auditing disclosure of third-party data collection in website privacy policies. In: Proceedings of the 2018 World Wide Web Conference. pp. 207–216 (2018)

[116] Ligett, K., Nissim, K.: We need to focus on how our data is used, not just how it is shared. Communications of the ACM 66(9), 32–34 (2023)

[117] Linden, T., Khandelwal, R., Harkous, H., Fawaz, K.: The privacy policy landscape after the gdpr. arXiv preprint arXiv:1809.08396 (2018)

[118] Lopes, I.M., Guarda, T., Oliveira, P.: How iso 27001 can help achieve gdpr compliance. In: 2019 14th Iberian Conference on Information Systems and Technologies (CISTI). pp. 1–6. IEEE (2019)

[119] MacKenzie, C.M., Laskey, K., McCabe, F., Brown, P.F., Metz, R., Hamilton, B.A.: Reference model for service oriented architecture 1.0. OASIS standard 12(S18), 1–31 (2006)

[120] Malgieri, G.: The concept of fairness in the gdpr: a linguistic and contextual interpretation. In: Proceedings of the 2020 Conference on fairness, accountability, and transparency. pp. 154–166 (2020)

[121] Marjanov, T., Konstantinou, M., Jóźwiak, M., Spagnuelo, D.: Data security on the ground: Investigating technical and legal requirements under the gdpr. Proceedings on Privacy Enhancing Technologies 3, 405–417 (2023)

[122] Martin, N., Friedewald, M., Schiering, I., Mester, B.A., Hallinan, D., Jensen, M.: The data protection impact assessment according to article 35 gdpr (2020)

[123] Martín García, Y.S., Álamo Ramiro, J.M.d.: A metamodel for privacy engineering methods. CEUR Workshop Proceedings (2017)

[124] McCallister, E.: Guide to protecting the confidentiality of personally identifiable information, vol. 800. Diane Publishing (2010)

[125] McCallister, E., Grance, T., Scarfone, K.: Identifiable information (pii). NIST Special Publication 800, 122 (2010)

[126] McDonald, A.M., Cranor, L.F.: The cost of reading privacy policies. Isjlp 4, 543 (2008)

[127] Mhaidli, A., Fidan, S., Doan, A., Herakovic, G., Srinath, M., Matheson, L., Wilson, S., Schaub, F.: Researchers' experiences in analyzing privacy policies: Challenges and opportunities. Proceedings on Privacy Enhancing Technologies 4, 287–305 (2023)

[128] Mike, N.: European Privacy by Design. Ph.D. thesis, Budapesti Corvinus Egyetem (2023)

[129] Mollaeefar, M., Ranise, S.: Identifying and quantifying trade-offs in multi-stakeholder risk evaluation with applications to the data protection impact assessment of the gdpr. Computers & Security 129, 103206 (2023)

[130] Moody, D.L.: Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. Data & Knowledge Engineering 55(3), 243–276 (2005)

[131] Moody, D.L., Shanks, G.G.: What makes a good data model? evaluating the quality of entity relationship models. In: International Conference on Conceptual Modeling. pp. 94–111. Springer (1994)

[132] Mourby, M., Mackey, E., Elliot, M., Gowans, H., Wallace, S.E., Bell, J., Smith, H., Aidinlis, S., Kaye, J.: Are 'pseudonymised'data always personal data? implications of the gdpr for administrative data research in the uk. Computer Law & Security Review 34(2), 222–233 (2018)

[133] Myers, M.D., Newman, M.: The qualitative interview in is research: Examining the craft. Information and organization 17(1), 2–26 (2007)

[134] Nissim, K.: Privacy: From database reconstruction to legal theorems. In: Proceedings of the 40th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems. pp. 33–41 (2021)

[135] Nissim, K., Bembenek, A., Wood, A., Bun, M., Gaboardi, M., Gasser, U., O'Brien, D.R., Steinke, T., Vadhan, S.: Bridging the gap between computer science and legal approaches to privacy. Harv. JL & Tech. 31, 687 (2017)

[136] Nissim, K., Wood, A.: Foundations for robust data protection: Co-designing law and computer science. In: 2021 Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA). pp. 235–242. IEEE (2021)

[137] Notario, N., Crespo, A., Martín, Y.S., Del Alamo, J.M., Le Métayer, D., Antignac, T., Kung, A., Kroener, I., Wright, D.: Pripare: integrating privacy best practices into a privacy engineering methodology. In: 2015 IEEE Security and Privacy Workshops. pp. 151–158. IEEE (2015)

[138] Ohm, P.: Broken promises of privacy: Responding to the surprising failure of anonymization'(2010). UCLA Law Review 57, 1701

[139] Ohm, P.: Broken promises of privacy: Responding to the surprising failure of anonymization. UCLA l. Rev. 57, 1701 (2009)

[140] OMG, S., Notation, O.: Software & systems process engineering meta-model specification. OMG Std., Rev 2, 18–71 (2008)

[141] Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., Sinz, E.J.: Memorandum on design-oriented information systems research. European journal of information systems 20(1), 7–10 (2011)

[142] Pallas, F., Koerner, K., Barberá, I., Hoepman, J.H., Jensen, M., Narla, N.R., Samarin, N., Ulbricht, M.R., Wagner, I., Wuyts, K., et al.: Privacy engineering from principles to practice: A roadmap. IEEE Security & Privacy 22(2), 86–92 (2024)

[143] Perera, C., McCormick, C., Bandara, A.K., Price, B.A., Nuseibeh, B.: Privacy-by-design framework for assessing internet of things applications and platforms. In: Proceedings of the 6th International Conference on the Internet of Things. pp. 83–92 (2016)

[144] Powar, J., Beresford, A.R.: Sok: Managing risks of linkage attacks on data privacy. Proceedings on Privacy Enhancing Technologies 2, 97–116 (2023)

[145] Presthus, W., Sørum, H.: Consumer perspectives on information privacy following the implementation of the gdpr (2019)

[146] Price, T.: It4it™: the new enterprise architecture framework. Hewlett-Packard Enterprise (2016)

[147] Prokhorenkov, D.: Alternative methodology and framework for assessing differential privacy constraints and consequences from a gdpr perspective. In: 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC). pp. 0359–0364. IEEE (2022)

[148] Prokhorenkov, D.: Anonymization level and compliance for differential privacy: A systematic literature review. 2022 International Wireless Communications and Mobile Computing (IWCMC) pp. 1119–1124 (2022)

[149] Prokhorenkov, D.: Toward compliance implications and security objectives: A qualitative study. In: 2023 IEEE 39th International Conference on Data Engineering Workshops (ICDEW). pp. 138–145. IEEE (2023)

[150] Prokhorenkov, D., Cao, Y.: Towards benchmarking privacy risk for differential privacy: A survey. In: Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. pp. 322–327 (2023)

[151] Pub, F.: Standards for security categorization of federal information and information systems. NIST FIPS 199, 122 (2004)

[152] Purtova, N.: From knowing by name to targeting: The meaning of identification under the gdpr'(2022). International Data Privacy Law 12, 163

[153] Radack, S.: The system development life cycle (sdlc). Tech. rep., National Institute of Standards and Technology (2009)

[154] Rastogi, V., et al.: Software development life cycle models-comparison, consequences. International Journal of Computer Science and Information Technologies 6(1), 168–172 (2015)

[155] Regulation, G.D.P.: General data protection regulation (gdpr). Intersoft Consulting, Accessed in October 24(1) (2018)

[156] Reichheld, F.F.: The one number you need to grow. Harvard business review 81(12), 46–55 (2003)

[157] Ridley, G., Young, J., Carroll, P.: Cobit and its utilization: A framework from the literature. In: 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the. pp. 8–pp. IEEE (2004)

[158] Rösch, D., Schuster, T., Waidelich, L., Alpers, S.: Privacy control patterns for compliant application of gdpr (2019)

[159] Rosemann, M., Rosemann, M.: Grundsätze ordnungsmäßiger modellierung. Komplexitätsmanagement in Prozeßmodellen: Methodenspezifische Gestaltungsempfehlungen für die Informationsmodellierung pp. 85–152 (1996)

[160] Rosemann, M., Schütte, R.: Grundsätze ordnungsmäßiger referenzmodellierung. Entwicklungsstand und Entwicklungsperspektiven der Referenzmodellierung 32, 16–33 (1997)

[161] Rosemann, M., Schütte, R.: Multiperspektivische referenzmodellierung. In: Referenzmodellierung: State-of-the-Art und Entwicklungsperspektiven. pp. 22–44. Springer (1999)

[162] Rost, M.: Dsk: Standard-datenschutzmodell v2. DuD-Datenschutz und Datensicherheit 43(1), 6–7 (2020)

[163] Rost, M., Pfitzmann, A.: Datenschutz-schutzziele—revisited. Datenschutz und Datensicherheit-DuD 33, 353–358 (2009)

[164] Rubinstein, I.S., Good, N.: The trouble with article 25 (and how to fix it): the future of data protection by design and default. International Data Privacy Law 10(1), 37–56 (2020)

[165] Rubinstein, I.S., Hartzog, W.: Anonymization and risk. Wash. L. Rev. 91, 703 (2016)

[166] Sabo, J., Willett, M., Brown, P., Janssen, G., Jutla, D.: Privacy management reference model and methodology (pmrm) version 1.0. OASIS Open (2012)

[167] Saltarella, M., Desolda, G., Lanzilotti, R.: Privacy design strategies and the gdpr: A systematic literature review. In: International Conference on Human-Computer Interaction. pp. 241–257. Springer (2021)

[168] Sample, K.L., Hagtvedt, H., Brasel, S.A.: Components of visual perception in marketing contexts: A conceptual framework and review. Journal of the Academy of Marketing Science 48, 405–421 (2020)

[169] Sangaroonsilp, P., Dam, H.K., Choetkiertikul, M., Ragkhitwetsagul, C., Ghose, A.: A taxonomy for mining and classifying privacy requirements in issue reports. Information and Software Technology 157, 107162 (2023)

[170] Schermann, M., Böhmann, T., Krcmar, H.: Fostering the evaluation of reference models: Application and extension of the concept of is design theories. 8. Internationale Tagung Wirtschaftsinformatik 2007-Band 2 p. 181 (2007)

[171] Schiffner, S., Serna, J., Ikonomou, D., Rannenberg, K.: Privacy technologies and policy. Edited by S. Schiffner et al 11498 (2019)

[172] Schuette, R., Rotthowe, T.: The guidelines of modeling–an approach to enhance the quality in information models. In: International Conference on Conceptual Modeling. pp. 240–254. Springer (1998)

[173] Schütte, R.: Grundsätze ordnungsmäßiger Referenzmodellierung: Konstruktion konfigurations-und anpassungsorientierter Modelle, vol. 233. Springer-Verlag (2013)

[174] Schütte, R., Schütte, R.: Die neuen grundsätze ordnungsmäßiger modellierung. Grundsätze ordnungsmäßiger Referenzmodellierung: Konstruktion konfigurations-und anpassungsorientierter Modelle pp. 111–175 (1998)

[175] Selzer, A., Woods, D., Bohme, R.: An economic analysis of appropriateness under article 32 gdpr. Eur. Data Prot. L. Rev. 7, 456 (2021)

[176] Shapiro, S.S.: Privacy by design: moving from art to practice. Communications of the ACM 53(6), 27–29 (2010)

[177] Shivashankarappa, A.N., Smalov, L., Dharmalingam, R., Anbazhagan, N.: Implementing it governance using cobit: A case study focusing on critical success factors. In: World Congress on Internet Security (WorldCIS-2012). pp. 144–149. IEEE (2012)

[178] Shostack, A.: Threat modeling: Designing for security. John Wiley & Sons (2014)

[179] Sion, L., Dewitte, P., Van Landuyt, D., Wuyts, K., Emanuilov, I., Valcke, P., Joosen, W.: An architectural view for data protection by design. In: 2019 IEEE International Conference on Software Architecture (ICSA). pp. 11–20. IEEE (2019)

[180] Sirur, S., Nurse, J.R., Webb, H.: Are we there yet? understanding the challenges faced in complying with the general data protection regulation (gdpr). In: Proceedings of the 2nd International Workshop on Multimedia Privacy and Security. pp. 88–95 (2018)

[181] Slavin, R.E.: Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational researcher 15(9), 5–11 (1986)

[182] Sokolovska, A., Kocarev, L.: Integrating technical and legal concepts of privacy. Ieee Access 6, 26543–26557 (2018)

[183] Spiekermann, S., Cranor, L.F.: Engineering privacy. IEEE Transactions on software engineering 35(1), 67–82 (2008)

[184] Spindler, G., Schmechel, P.: Personal data and encryption in the european general data protection regulation. J. Intell. Prop. Info. Tech. & Elec. Com. L. 7, 163 (2016)

[185] Stanciu, A.M., Ciocârlie, H.: Integrating security into the software development life cycle: A systematic approach. In: 2023 International Conference on Electrical, Computer and Energy Technologies (ICECET). pp. 1–6. IEEE (2023)

[186] Steuperaert, D.: Cobit 2019: A significant update. EDPACS 59(1), 14–18 (2019)

[187] Stiles, W.B.: Evaluating qualitative research. BMJ Ment Health 2(4), 99–101 (1999)

[188] Tambo, T., Filtenborg, J.: It4it™ as a management of technology framework: Perspectives, implications and contributions. In: International Association for Management of Technology. pp. 1–14. International Association for Management of Technology (IAMOT) (2017)

[189] Tapia, R.S., Daneva, M., van Eck, P., Wieringa, R.: Towards a business-it aligned maturity model for collaborative networked organizations. In: 2008 12th Enterprise Distributed Object Computing Conference Workshops. pp. 276–287. IEEE (2008)

[190] Thomas, O.: Understanding the term reference model in information systems research: history, literature analysis and explanation. In: International Conference on Business Process Management. pp. 484–496. Springer (2005)

[191] Timm, F., Sandkuhl, K.: A reference enterprise architecture for holistic compliance management in the financial sector (2018)

[192] Tolsdorf, J., Iacono, L.L.: Data cart: A privacy pattern for personal data management in organizations. Human Factors in Privacy Research p. 353 (2023)

[193] Tongren, J., Warigon, S.: A preliminary survey of cobit use. EDPACS: The EDP Audit, Control, and Security Newsletter 25(3), 17–19 (1997)

[194] Torre, D., Abualhaija, S., Sabetzadeh, M., Briand, L., Baetens, K., Goes, P., Forastier, S.: An ai-assisted approach for checking the completeness of privacy policies against gdpr. In: 2020 IEEE 28th International Requirements Engineering Conference (RE). pp. 136–146. IEEE (2020)

[195] Truong, N., Sun, K., Wang, S., Guitton, F., Guo, Y.: Privacy preservation in federated learning: An insightful survey from the gdpr perspective. Computers & Security 110, 102402 (2021)

[196] Ulrich, F., Stefan, S., Koch, S.: Open model, ein vorschlag für ein forschungsprogramm der wirtschaftsinformatik wirtschaftinformatik (2007)

[197] Van Blarkom, G., Borking, J.J., Olk, J.E.: Handbook of privacy and privacy-enhancing technologies. Privacy Incorporated Software Agent (PISA) Consortium, The Hague 198, 14 (2003)

[198] Veale, M., Binns, R., Ausloos, J.: When data protection by design and data subject rights clash. International Data Privacy Law 8(2), 105–123 (2018)

[199] Venable, J., Pries-Heje, J., Baskerville, R.: Feds: a framework for evaluation in design science research. European journal of information systems 25(1), 77–89 (2016)

[200] Voigt, P., Von dem Bussche, A.: EU-Datenschutz-Grundverordnung (DSGVO): Praktikerhandbuch. Springer-Verlag (2018)

[201] Vom Brocke, J.: Referenzmodellierung: Gestaltung und Verteilung von Konstruktionsprozessen, vol. 4. Jan vom Brocke (2003)

[202] Voss, W.G., Houser, K.A.: Personal data and the gdpr: providing a competitive advantage for us companies. American Business Law Journal 56(2), 287–344 (2019)

[203] Wagner, I.: Privacy policies across the ages: content of privacy policies 1996–2021. ACM Transactions on Privacy and Security 26(3), 1–32 (2023)

[204] Wang, Y.: Privacy-enhancing technologies. In: Handbook of research on social and organizational liabilities in information security, pp. 203–227. IGI Global (2009)

[205] Willis, C.L., Miertschin, S.L.: Mind maps as active learning tools. Journal of computing sciences in colleges 21(4), 266–272 (2006)

[206] Wuyts, K.: Privacy threats in software architectures (2015)

[207] Wuyts, K., Joosen, W.: Linddun privacy threat modeling: a tutorial. CW Reports (2015)

[208] Yan, B., Li, K., Xu, M., Dong, Y., Zhang, Y., Ren, Z., Cheng, X.: On protecting the data privacy of large language models (llms): A survey. arXiv preprint arXiv:2403.05156 (2024)

[209] Zhu, W., Kairouz, P., McMahan, B., Sun, H., Li, W.: Federated heavy hitters discovery with differential privacy. In: International Conference on Artificial Intelligence and Statistics. pp. 3837–3847. PMLR (2020)

A Appendix

We offer DPGs that were briefly introduced in **Section 2.5**. The list of DPGs presented in this appendix is by no means exhaustive or complete and is intended only to explain the nature of the DPGs (we also point out that a) the DPGs may change due to various factors not taken into account in the dissertation data and b) the relationships and connections between the DPGs may change over time, thus we point out the need to evaluate and update the DPGs according to the latest version of the SDM). The description and explanation of DPGs could be utilized to systematize GDPR requirements. Supplementary materials provided by SDM frameworks could also extend the description of DPGs and numerated accordingly (e.g. B1.3 [1]). Seven DPGs, as introduced in SDM [163] and integrated in RMDP, are presented below:

- **Data Minimization** - covers the fundamental requirement under data protection law to limit the processing of personal data to what is appropriate, substantial and necessary for the purpose. The implementation of this minimisation requirement has a far-reaching influence on the scope and intensity of the protection concept determined by the other protection goals. Data minimisation specifies and operationalises the principle of necessity in the processing process, which requires of this process as a whole as well as each of its steps not to process more personal data than is needed to achieve the purpose of processing (B1.3). The minimisation requirement applies not only to the quantity of data processed, but also to the scope of its processing, its storage period and its accessibility. In particular, it is necessary to ensure that personal data are kept in a form which only permits identification of data subjects for as long as is necessary for the purposes of the processing (B1.5). Data minimisation starts with the design of the information technology by the manufacturer through its configuration and adaptation to the operating conditions (B1.17) to its use in the core processes of processing as well as in the supporting processes, for example in the maintenance of the systems used.

- **Availability** - refers to the requirement that access to personal data and their processing is possible without delay and that the data can be used properly in the intended process. For this purpose, the data must be accessible by authorised parties and the intended methods for processing must be applied to them. Availability includes the concrete retrievability of data, e. g. through data management systems, structured databases and search functions, and the ability of the technical systems used to present data appropriately for humans (B1.18). Furthermore, measures must be taken to implement availability to ensure that personal data and access to them can be rapidly restored in the event of a physical or technical incident (B1.20). Measures must also be implemented to guarantee the availability of personal data and the systems and services that process them when they are under a reasonable expected load and to ensure that the protection of personal data is not compromised in the event of an unexpectedly high load (B1.19). If, in exceptional cases, the protection of personal data with regard to availability is nevertheless violated, it must be ensured that measures are taken to rectify and mitigate the violation (B1.22).

- **Integrity** - refers, on the one hand, to the requirement that information technology processes and systems continuously comply with the specifications that were defined for them to perform their intended functions (B1.6). On the other hand, integrity refers to the property that the data to be processed remain intact (B1.6), complete, correct and up-to-date (B1.4). Deviations from these characteristics must be ruled out or at least detectable (B1.23) so that they can be addressed and corrected (B1.22). This also applies if the underlying systems and services are subject to unexpectedly high loads (B1.19). In addition to the aspect of freedom from errors, the aspect of freedom from discrimination must be maintained, especially

[1] https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methodology$_V2.0b.pdf$

in automated evaluation and decision-making processes (B1.16). The factors and characteristics of an assessment or decision-making process that may have potentially discriminatory effects shall be identified a priori in the legal review, and taken into account in implementation and monitored in operation This aspect is reflected, for example, by measures to clean up training data and validate results when applying AI procedures.

- **Confidentiality** - refers to the requirement that no unauthorised person can access or use personal data (B1.7). Unauthorised persons are not only third parties outside the responsible body, but also employees of technical service providers who do not require access to personal data in order to provide the service, or persons in organisational units who have no connection whatsoever with the content of a processing activity or with the data subject. The confidentiality of personal data must also be ensured when the underlying systems and services are subject to unexpectedly high loads (81.19). Should confidentiality nevertheless be violated in exceptional cases, it must be ensured that measures are taken to remedy and mitigate the accompanying violation of the protection of personal data (B1.22).

- **Unlinkability** - refers to the requirement that personal data shall not be merged, i. e. linked. It must be implemented in practice especially if the data to be merged were collected for different purposes (B1.2). The larger and more meaningful the data base, the greater the potential greed may be to use the data beyond the original legal basis. Such further processing is only legally permissible under strictly defined circumstances. The unlinkability is to be ensured by means of technical and organisational measures. In addition to measures for pseudonymisation, other measures that allow further processing separately from the original processing are also suitable, both on the organisation side and on the system side. The data base can be adapted, for example, by authorisation systems and reduction to the extent necessary for the new purpose.

- **Transparency** - refers to the requirement that both data subjects (B1.1) and system operators (B1.23 Adequate monitoring of processing) and competent supervisory bodies (B1.8) shall be able to identify to varying degrees which data are collected and processed when and for what purpose in a processing activity, which systems and processes are used to determine where the data are used and for what purpose, and who has legal responsibility for the data and systems in the various phases of data processing. Transparency is necessary for the monitoring and control of data, processes and systems from their creation to their erasure and a prerequisite for legally compliant data processing to which, where necessary, data subjects can give an informed consent (B2). Transparency of the whole data processing and of the instances involved can help to ensure that, in particular, data subjects and supervisory bodies can identify deficiencies and, if necessary, demand appropriate changes to the processing.

- **Intervenability** - refers to the requirement that the data subjects' rights to notification, information, rectification (B1.11), erasure (B1.12), restriction (B1.13), data portability (B1.14), objection and obtaining the intervention in automated individual decisions (B1.15) are granted without undue delay and effectively if the legal requirements exist (B1.10) and data controller is obliged to implement the corresponding measures. Where the data controller has information enabling him to identify the data subjects, he must also take measures to identify and authenticate the data subjects who wish to exercise their rights (B1.9). In order to implement the rights of data subjects and supervisory orders (B3 Implementation of supervisory orders) and to remedy and mitigate data protection breaches (B1.22), the controllers must at all times be in a position to take action in data processing, from collection to erasure of the data. Where the processing of personal data is based on the consent of the data subject, measures must be taken to ensure that the personal data are processed only where the data subject has given his or her consent and where that consent has not been withdrawn (B2). For information technology processing to which the data subjects themselves have access (e. g. applications on the smartphone) and for which different data protection settings are intended, data-protection friendly default settings must be defined by the controller and further measures must be taken. These further measures must enable data subjects to make their own config-

urations, differentiated according to the respective processing purposes, and to decide which processing operations they wish to allow that go beyond the minimum required (B1.17).

We acknowledge that it is not a specified period of time after the submission of the current dissertation; the earlier utilized definitions or interpretations (Chapter 2) could be modified and adopted considering changes in regulations related to the studied domain or further technical advancements. Therefore, we outline the sources utilized (demonstrated in the Bibliography and footnotes) and also stress that further studies or investigations should first revalidate the definition or interpretation to keep the model current. Further, a list of sources for the definitions and interpretations used was explained, suggesting that the organization or user of the model should review any changes or modifications that may occur. We also acknowledge and appreciate the right of organizations or practitioners to utilize different interpretations or definitions instead of presented, as the model is designed with their adaptability in mind.

B Appendix

The appendix **B** provides an overview (chronological sprints) of how the reference model (RMDP) design and visualization evolved over time. The various versions of the RMBD presented below are arranged in chronological (time) order.

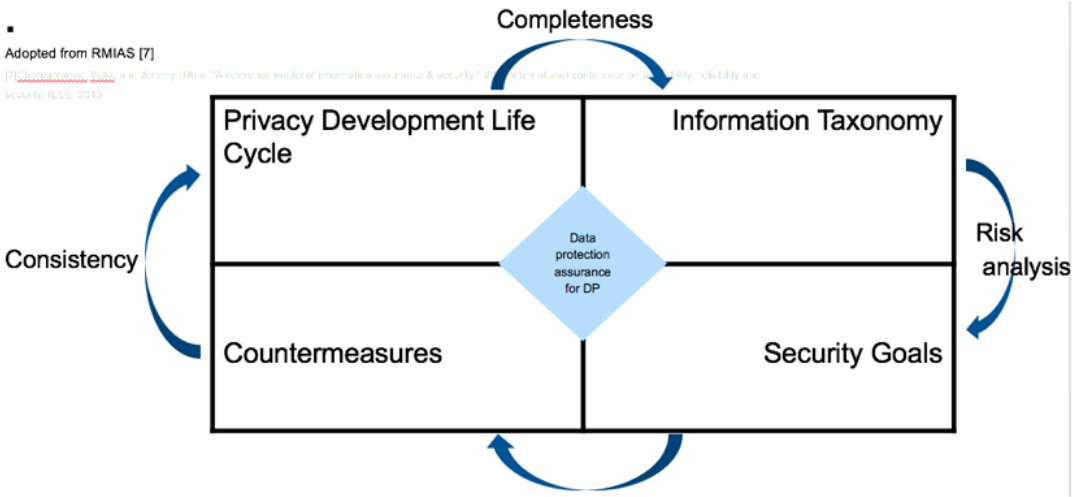


Figure B.1 RMDP ver.A, Blocks visualization, concept
Note: Presented and discussed during Sapporo Privacy Workshop (2023)

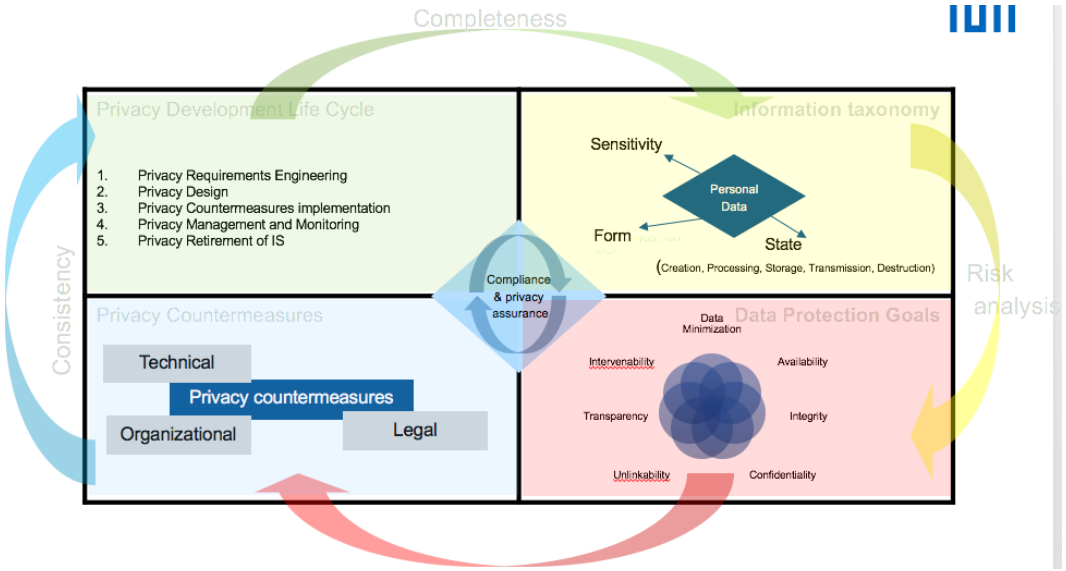


Figure B.2 RMDP ver.B, Content and Blocks visualization
Note: Presented and discussed during TUM I20 chair feedback session (2024)

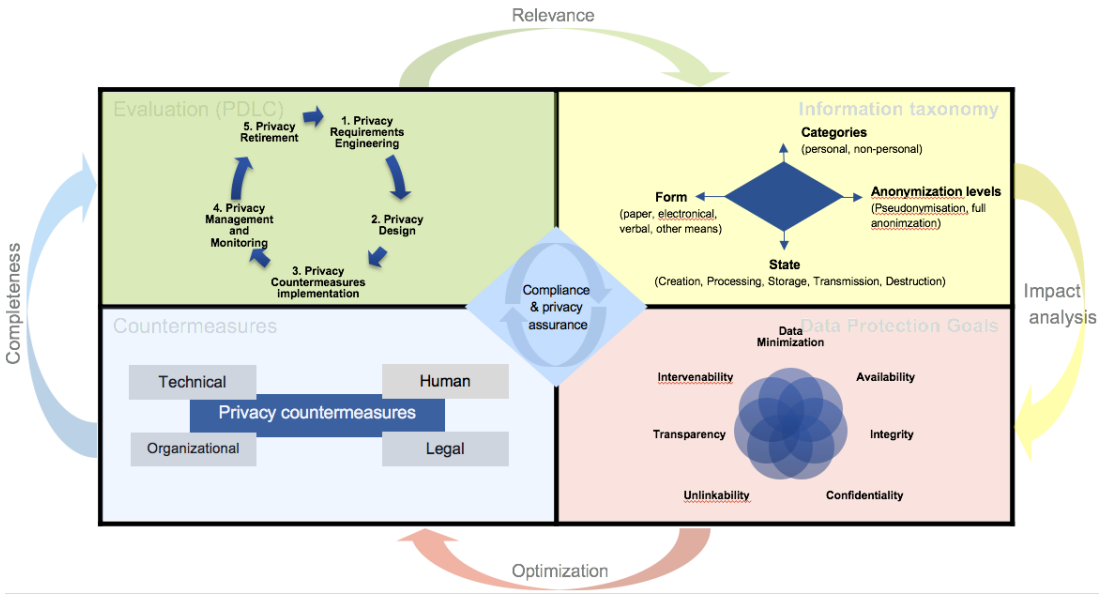


Figure B.3 RMDP ver.C
Note: Presented and discussed during the 19th International Workshop on Security (2024)

C Appendix

Appendix C demonstrates the template of a survey structure organized with Google Survey services. This includes two main blocks of questions, followed by additional question for collecting net promoter score data.

Evaluation Questions:

Do you have any experience in Data Protection, Privacy or related domains?

How many years of experience do you have?

Does your experience come from industry or academia?

RMDP is evaluated with eight criteria: Simplicity, Accuracy, Scope, Systematic Power, Explanatory Power, Reliability, Validity and Fruitfulness.

Criteria 1: Simplicity

Are the elements of the Model shown simple?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Are the relationships between the elements simple? (Relationships illustrated by arrows)
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Criteria 2: Accuracy

Are the classifications included in the Model accurate? (information taxonomy, data protection goals, privacy countermeasures, privacy development life cycle)
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Are the interrelationships between the elements of the Model accurately represented?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Criteria 3: Scope (completeness) (Model should cover the broad scope of the domains represented and should not overlook the essential concepts/elements)

Does the Model include all elements/concepts essential for the Privacy domain?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

If you think some essential elements/concepts are missing, please name them.

Does the Model include any elements not applicable/relevant to the Privacy domain?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Criteria 4: Systematic power (Model should help to organize concepts/elements and relationships between them in a meaningful, systematic way)

Does the Model organize elements of the Privacy domain and the interconnection between them in a structured, systematic way?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Criteria 5: Explanatory power (the Model should assist with explaining and forecasting issues)

Could the Model assist with clarifying (tracing back) and forecasting issues related to Privacy?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Criteria 6: Reliability (the Model should be valid for the situations it was designed and lead to a similar understanding when utilized by different users/stakeholders)

In your opinion, would the Model be reasonable for the majority of business organizations?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

If the current Model could not be applicable to some industry or organization, please name them (or explain your opinion)

Criteria 7: Validity (the Model should provide reasonable representations and findings)

Would the methodology incorporated into the Model lead to reasonable results?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Criteria 8: Fruitfulness (preferably, the Model should suggest research gaps)

Does the Model deliver a convenient structure for framing existing research?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Would you utilize the current Model in your research/practice?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Could a model be utilized or assist in bridging the gaps in the existing research/practice?
1 Strongly Disagree – 2 Disagree – 3 Undecided / Neutral - 4 Agree - 5 Strongly Agree

Additional question

On a scale of 0 to 10, how likely is it that you would recommend RMDP to a colleague?
0 Very Unlikely - 10 Very Likely

D Appendix

Appendix D demonstrates a more detailed overview of the four types of privacy countermeasures categories. It's important to note that the current sections are examples of demonstrated privacy countermeasures. It's also important to note that the categories of these demonstrated privacy countermeasures are not fully representative of the current appendix, setting the reader's expectations accordingly. We also want to stress that declared measures in Article 32 of GDPR are applicable to technical and organizational categories. Nevertheless, the organizations have a choice to adjust referred declared recommendations considering their goals. Additionally, earlier declared privacy countermeasures, namely Technical, Organizational, Legal, and Human, represent a high level of abstraction, which could be adopted by the organizations and further assigned to privacy countermeasures categories accordingly.

Generic Countermeasures include[1]:

- Privacy Impact Assessment as part of change control processes;
- Privacy threat assessment;
- Undertake a privacy risk assessment of all current application portfolio;
- Define privacy requirements early in projects;
- Review privacy throughout process development;
- Privacy awareness training (managers, staff, volunteers, contractors);
- Organizational Privacy Policy;
- Instrument processes to automatically detect that privacy requirements are in place;
- Track changes to legislation, guidance and any privacy-related clauses in other mandates such as contracts;
- Ask local DPA to audit you;
- Map (or create a process flow for) the data in order to understand what is collected, why it is needed, how it is used, and what protections are in place;
- Use anonymization, pseudonymization and data minimisation/avoidance where possible;
- Implement metrics to measure privacy performance in applications/data;

GDPR Countermeasures recommendations:

- the pseudonymization and encryption of personal data;
- the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
- the 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 organizational measures for ensuring the security of the processing;

[1] https://owasp.org/www-project-top-10-privacy-risks/OWASP$_{Top10Privacy Risks Countermeasures v2.0.pdf}$