TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Informatik XIX

Development of a reference process model for GDPR compliance management based on enterprise architecture

Dominik Martin Huth

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Jens Großklags

Prüfer der Dissertation:
1. Prof. Dr. Florian Matthes
2. Prof. Dr. Helmut Krcmar

Die Dissertation wurde am 18.02.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 02.07.2021 angenommen.

Zusammenfassung

Heutzutage bauen viele wirtschaftliche Aktivitäten auf der Nutzung von personenbezogenen Daten auf. Im Jahr 2016 hat die Europäische Union die Datenschutzgrundverordnung (DSGVO) verabschiedet, die die vorhergehende Richtlinie von 1995 ersetzt. Diese wurde zu einer Zeit erlassen, als weniger als ein Prozent der Weltbevölkerung das Internet nutzte. Die DSGVO führte aktualisierte Definitionen für personenbezogene Daten und deren Verarbeitung ein, sie erweiterte die Rechte von Betroffenen, und sie macht Unternehmen strikte Vorgaben für die Dokumentation und Meldepflicht von datenschutzrechtlich relevanten Vorgängen. Als womöglich wichtigste Merkmale etabliert sie zudem die Rechenschaftspflicht und einen Strafrahmen von bis zu vier Prozent des Jahresumsatzes, mit denen Verstöße gegen die Verordnung geahndet werden können.

Über zwei Jahre nachdem die Verordnung im Jahr 2018 in Kraft getreten ist, bemühen sich Unternehmen immer noch um umfassende Konformität mit der DSGVO. Da zum Zeitpunkt der Einführung keine Standard-Rahmenwerke verfügbar waren, haben die Organisationen eigene Ansätze entwickelt, um die neuen Anforderungen umzusetzen und im Geschäftsbetrieb beizubehalten. In vielen Organisationen hat das Unternehmensarchitekturmanagement (EAM) zur Bewältigung dieser Aufgaben beigetragen.

Ziel dieser Arbeit ist es, ein Referenzprozessmodell zu entwickeln, welches Ansätze zum Datenschutzmanagement strukturiert. Das Modell wird in vier Phasen entwickelt: Definition des Problems, Konstruktion des Modellrahmens, sowie Konstruktion und Validierung des Modells. Eine Umfrage mit 38 Teilnehmern erarbeitet die relevanten Aufgaben im Datenschutzmanagement und untersucht den möglichen Nutzen einer Zusammenarbeit mit dem Unternehmensarchitekturmanagement. Der Modellrahmen entsteht aus Literaturquellen aus der Wirtschaftsinformatik und dem Datenschutz. Auf Basis von 24 Interviews mit Unternehmensarchitekten wird das Modell konstruiert. Das Modell wird durch elf qualitative Interviews und eine Umfrage mit 29 Teilnehmern evaluiert.

Der Hauptbeitrag dieser Arbeit, das Referenzprozessmodell *ProPerData*, besteht aus acht Stakeholder Rollen, sieben Informationsquellen, neun Zeiteinheiten, elf Aufgaben im Datenschutzmanagement, 16 Arbeitsschritten und zwölf Arbeitsergebnissen. Es kann Unternehmen, die EAM betreiben, als Vorlage für DSGVO Ansätze dienen. Unter den befragten Stakeholder-Gruppen zeigen Unternehmensarchitekten die größte Zustimmung gegenüber ProPerData. Software-Entwickler würden nur einzelne Aspekte des Modells betrachten, aber schätzen den Beitrag von ProPerData für das Gesamtverständnis der Verordnung. Obwohl Datenschutzexperten den Wert des Modells für die Kommunikation unter Stakeholdern begrüßen, sehen sie keine direkten positiven Auswirkungen auf ihre eigene Arbeit.

Zukünftige Forschungsthemen sind die Entwicklung einer Methode für die Implementierung von ProPerData, weitere Forschung zur Zusammenarbeit zwischen EAM, dem Datenschutz und weiteren Abteilungen, und die Entwicklung von leichtgewichtigen Ansätzen, um die Einhaltung von Datenschutzanforderungen im agilen Kontext zu gewährleisten.

Abstract

Today's economy relies heavily on personal data. In 2016, the European Union passed the General Data Protection Regulation (GDPR) to replace the previous data protection legislation from 1995 – a time when less than 1% of the world population used the internet. The GDPR introduced updated definitions for personal data and processing personal data, enhanced data subject rights, and strict documentation and reporting obligations. Most importantly, it establishes the principle of accountability and allows fines of up to 4% annual revenue for violations of the regulation.

More than two years after the regulation entered into force, companies still struggle to achieve full GDPR compliance. Since there were no standard frameworks available, each organization developed its own approach for implementing the new provisions and maintaining compliance in changing business environments. Many organizations relied on support from enterprise architecture management (EAM) to address the various challenges that the GDPR poses.

The goal of this thesis is to develop a reference process model that structures data protection management (DPM) approaches. It is constructed in a research process with four phases: problem description, identification of the model frame, model construction, and validation. A survey with 38 data protection officers (DPOs) establishes the relevant tasks and investigates possible benefits of collaboration with enterprise architecture (EA). The model frame is constructed from literature sources in the information systems and privacy domain. 24 qualitative interviews with enterprise architects serve as the basis for text coding and model construction. ProPerData is evaluated through eleven qualitative interviews and a survey with 29 participants.

ProPerData, which represents the main contribution of this work, is a reference process model that is comprised of eight stakeholder roles, seven resources, nine temporal entities, eleven DPM tasks, 16 work units, and twelve work products. It can serve as a blueprint for GDPR implementation approaches in companies that engage in EAM. Among the consulted stakeholders, enterprise architects show the highest approval of ProPerData. Software developers would mostly focus on their own tasks, but appreciate the value of ProPerData for overall understanding of the regulation. While DPM experts acknowledge the value of the process model for communication among other stakeholders, they do not see immediate positive effects on their own work.

Implications for future research include designing a method for implementing ProPerData, extending work on the collaboration between EA, DPM and other departments, and creating lightweight methods for ensuring compliance in agile organizations.

Danksagung

Mein aufrichtiger Dank gilt Prof. Dr. Florian Matthes für die Möglichkeit, diese Arbeit unter seiner Betreuung am Lehrstuhl für Software Engineering betrieblicher Informationssysteme durchzuführen. Das anregende Umfeld und sein Reichtum an Ideen haben meine Auseinandersetzung mit IT als wichtigem Teil des gesellschaftlichen Lebens gefördert und meine Haltung gegenüber Lernen, Wissen und Austausch bedeutend gefestigt. Weiterhin bedanke ich mich herzlich bei Herrn Prof. Dr. Helmut Krcmar für die Zweitbegutachtung meiner Arbeit.

Diese Arbeit hat sehr profitiert von den zahlreichen Gesprächen mit Experten aus der Praxis, die mir teilweise mehrmals großzügig ihre Zeit und ihr Wissen in ausführlichen Interviews zur Verfügung gestellt haben. Ohne diese fundierten Einblicke und die fachliche Rückmeldung hätte ich diese Arbeit nicht schreiben können. Ich hoffe, dass ich durch meine Gedanken und Ergebnisse einen Teil davon zurückgeben konnte.

Danke an all meine Kollegen am Lehrstuhl und in der Forschung für den täglichen Austausch zu unseren geteilten Erfolgen und Herausforderungen. Mein besonderer Dank gilt meinen Kollegen Fabian Burmeister von der Universität Hamburg für die hervorragende Zusammenarbeit, Gloria Bondel für das rege Feedback, sowie Klym Shumaiev und Christof Tinnes für kluge Gedanken aus nächster Nähe. Vielen Dank an Aline Schmidt und Jian Kong für die verlässliche Arbeit am Lehrstuhl. Weiterhin danke ich meinen Studenten Laura, Ahmet, Nora und Michael für ihre wertvollen Beiträge und unsere gemeinsame Forschung.

Ich danke meinen Eltern für das Interesse und die Energie, die sie in mich und meine Ausbildung investiert haben, meinen Freunden für die Welt abseits der Dissertation, und meiner Frau Maike für die Liebe, Unterstützung und Verschnaufpausen auf unseren gemeinsamen Wegen zwischen Träumen und Zielen.

München, 14.02.2021

Dominik Martin Huth

Table of Contents

1. Introduction **1**
- 1.1. Privacy and technological progress . . . 2
- 1.2. The General Data Protection Regulation . . . 3
- 1.3. Problem statement and research questions . . . 4
- 1.4. Design-oriented information system (IS) research . . . 7
- 1.5. Reference models . . . 9
- 1.6. Research approach and thesis outline . . . 10

2. Foundations **15**
- 2.1. GDPR definitions . . . 15
- 2.2. GDPR stakeholders . . . 17
  - 2.2.1. Roles in the GDPR . . . 18
  - 2.2.2. Roles in GDPR implementation projects . . . 20
- 2.3. DPM Tasks . . . 21
  - 2.3.1. DPM tasks in other publications . . . 21
  - 2.3.2. Survey results on complexity and challenges of DPM tasks . . . 24
- 2.4. EAM for supporting DPM . . . 26
- 2.5. Summary . . . 32

3. Related Work **35**
- 3.1. Academia . . . 35
  - 3.1.1. PRIPARE . . . 35
  - 3.1.2. Capability-based approach by Labadie and Legner . . . 37
  - 3.1.3. Method by Koç . . . 39
  - 3.1.4. Other contributions . . . 40
- 3.2. Industry . . . 41
  - 3.2.1. Standard data protection model (SDM) of the German data protection authorities . . . 41

3.2.2. ISO 27001 . . . 43
3.2.3. COBIT . . . 45
3.2.4. IT4IT . . . 47
3.2.5. Other approaches . . . 50
3.3. Summary and Research Gap . . . 50

4. Construction of the Reference Model Frame 53
4.1. Requirements for a reference process model that supports GDPR implementation 53
4.1.1. General requirements for a reference process model . . . 54
4.1.2. Requirements to support compliance with the GDPR . . . 54
4.1.3. Requirements that originate from empirical opportunities and barriers in GDPR implementation projects . . . 55
4.2. Reference model frame . . . 56
4.2.1. Metamodel for the reference process model . . . 56
4.2.2. Modeling language . . . 59
4.2.3. Summary . . . 59

5. *ProPerData* - a reference process model for GDPR compliance management based on EA 61
5.1. Construction approach . . . 61
5.2. Data collection . . . 63
5.2.1. Interview series with enterprise architects . . . 63
5.2.2. Other empirical inquiries . . . 65
5.2.3. Academic publications and industry guidelines . . . 65
5.3. Construction of *ProPerData* . . . 65
5.3.1. Interrelations of model elements . . . 70
5.3.2. Roles and collaboration . . . 70
5.4. Design decisions . . . 73
5.4.1. Level of abstraction . . . 73
5.4.2. Relation to existing approaches . . . 75
5.4.3. Design principles to enhance information governance . . . 75
5.5. Interdependencies with internal processes . . . 76
5.5.1. Interdependencies with EAM . . . 76
5.5.2. Interdependencies with software development . . . 78
5.5.3. Interdependencies with other internal processes . . . 80
5.6. Usage of ProPerData . . . 80

6. Evaluation 83
6.1. Evaluation approach . . . 83
6.2. Qualitative evaluation . . . 85
6.2.1. Approach . . . 85
6.2.2. Results . . . 86
6.3. Quantitative Survey . . . 96
6.3.1. Approach . . . 96
6.3.2. Results . . . 98

6.4. Analytical discussion . . . 102
6.4.1. Fulfillment of the guidelines of modeling (GoM) . . . 102
6.4.2. Discussion of specific requirements . . . 103
6.5. Summary of evaluation . . . 105

7. Conclusion **107**
7.1. Summary . . . 107
7.2. Limitations . . . 111
7.2.1. Limitations of the considered material . . . 111
7.2.2. Limitations in the perspective of ProPerData . . . 112
7.2.3. Limitations in the method . . . 112
7.2.4. Limitations of the evaluation . . . 114
7.3. Reflection and outlook . . . 114
7.3.1. Design a method based on ProPerData . . . 115
7.3.2. DPM as EA stakeholder . . . 115
7.3.3. Ensuring data protection in agile organizations . . . 116

Bibliography **117**

Abbreviations **127**

A. Interview partners **129**

B. ProPerData - A process model for GDPR compliance **132**
B.1. Roles . . . 132
B.2. Stages . . . 133
B.3. Resources . . . 134
B.4. Work units . . . 137
B.4.1. Inform & educate . . . 137
B.4.2. Verify existing processing activities . . . 138
B.4.3. Create new processing activities . . . 139
B.4.4. Conduct Data Protection Impact Assessments (DPIA) . . . 142
B.4.5. Cooperate with supervisory authority . . . 144
B.4.6. Maintain records of processing activities . . . 145
B.4.7. Conduct Audits . . . 146
B.4.8. Interact with data subjects . . . 147
B.4.9. Report to management . . . 149
B.4.10. Execute organizational tasks . . . 150
B.4.11. Leverage data protection efforts for business impact . . . 151
B.5. Work products . . . 152

List of Figures

1.1. IS research framework (Hevner et al., 2004) adapted to this research project . . . 8
1.2. Reference modeling approaches in literature and their relationship to our reference modeling approach. . . . 11
1.3. Research approach and overview of this thesis . . . 13

2.1. Stakeholders and relationships between them (Huth et al., 2018) . . . 19
2.2. Mean time consumption of data protection tasks (n=37) . . . 25
2.3. Complexity distribution of data protection tasks (n=38) . . . 25
2.4. Most severe problems in data protection tasks . . . 27
2.5. Enterprise Architecture as Cross-layer view (Winter and Fischer, 2007) . . . 28
2.6. Relation of DPM tasks to EAM elements, based on (Huth et al., 2020c) and (Buckl, 2011, p.3) . . . 29
2.7. Proportion of DPM experts who were supported by EAM and reasons for not collaborating with EAM, by organization size (Huth et al., 2020c) . . . 31
2.8. Overall perceived usefulness of EAM support for DPM (Vilser, 2019) . . . 32
2.9. Usefulness of EAM for each DPM task (n=12) (Huth et al., 2020c) . . . 32

3.1. The PRIPARE method (Crespo et al., 2015) . . . 36
3.2. Capability model for data management (Labadie and Legner, 2019) . . . 38
3.3. The DPM process (Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder, 2019, p.54) . . . 43
3.4. ISO27001 implementation process (Brenner et al., 2011) . . . 45
3.5. COBIT process reference model (Gaulke, 2019) . . . 46
3.6. COBIT domains with management goals (Gaulke, 2019) . . . 47
3.7. The *IT value chain* in IT4IT (The Open Group, 2017) . . . 48
3.8. Levels of abstraction in IT4IT (The Open Group, 2017) . . . 49

4.1. Metamodel, model and instance levels, as used for ProPerData . . . 57
4.2. ProPerData metamodel . . . 58

4.3. Symbols in the reference model 59

5.1. Elements in single company instances as partial sets of the elements in the reference process model 62
5.2. Construction approach for ProPerData 63
5.3. Construction step of the initial version of ProPerData 66
5.4. The ProPerData overview canvas 69
5.5. ProPerData work units with responsibilities (RACI-Matrix) 70
5.6. Dependencies among ProPerData work units 71
5.7. The two abstraction levels of ProPerData 74
5.8. The data protection process in an exemplary process map (representation based on (Gadatsch, 2017, p.85)) 77
5.9. An exemplary EA model of an order placement process, its supporting applications, the processed data and the responsible business units (Huth et al., 2019b) . 77
5.10. Four levels of EAM support for DPM (Huth et al., 2020b) 79
5.11. Usage scenario of ProPerData 81

6.1. Evaluation framework for reference models (Frank, 2006), configured for this evaluation 84
6.2. Parameters for the qualitative evaluation 85
6.3. Parameters for the quantitative survey 97
6.4. Approach for the quantitative survey 97
6.5. Participant roles in quantitative evaluation (non-exclusive) 99
6.6. Distribution of assessments of ProPerData 99
6.7. Average assessments of ProPerData by participant role 100
6.8. Distribution of detractors, passively satisfied and promoters 101
6.9. The net promoter score (NPS) by selected roles 101

7.1. Venn diagram of the perception of reference models. Adapted from Thomas (2005) 113

B.1. Conceptual representation of privacy engineering methods 137
B.2. The PRIPARE lifecycle for privacy-friendly system design (adapted from Crespo et al. (2015)) 140
B.3. DPIA decision diagram (Article 29 Data Protection Working Party, 2017a) . . . 143
B.4. The DPIA process (adapted from Bieker et al. (2016)) 143
B.5. A simple process of creating the record of processing activities (RoPA) (Huth et al., 2019b) 145
B.6. Generic process for answering data subject requests 148
B.7. A reporting process, adapted from Taschner (2015) 149
B.8. Design space for effective privacy notices, cf. Schaub et al. (2017) 150

List of Tables

1.1. The 10 highest fines until January 2021 (CMS, 2021) 5

2.1. Examples of personal data and non-personal data (European Commission, 2020) 17
2.2. Interview partners for developing the list of DPM tasks (Vilser, 2019; Huth et al., 2020c) 22
2.3. List of DPM tasks (Vilser, 2019; Huth et al., 2020c) 23

5.1. Exemplary initial documentation of practices, following a pattern documentation template and the Metamodel for Privacy Engineering Methods (MPEM) structure 67
5.2. Example of ProPerData work unit candidate in work unit consolidation 68
5.3. Design principles for EAM to enhance information governance (partial table from (Burmeister et al., 2020)) 76

6.1. Participants in the qualitative evaluation interviews 87
6.2. Summary of qualitative interviews with three stakeholder groups 96
6.3. The GoM (Schütte, 1998) 102

A.1. Interview partners from interview series in (Huth et al., 2020b) and (Burmeister et al., 2020) 130
A.2. Expert sources with focus on data portability (Huth et al., 2019a) 130
A.3. Expert interviews with focus on the record of processing activities (Huth et al., 2019b) 131

B.1. Privacy properties that must be ensured with technical and organizational measures. Adapted from Huth and Matthes (2019) 153
B.2. Examples for guidelines to process owners and developers 155
B.3. Classification criteria for personal data 156

CHAPTER 1

Introduction

Technology spreads faster today than ever before. While it took eleven to thirteen years for 60% of the U.S. population to adopt personal computers, cellular phones and the internet, this barrier was reached after only seven years for smartphones and social media (Ritchie and Roser, 2019). In many instances, this technological progress relies on personal data. The accuracy and usefulness of personalized apps and services depend on knowledge about user behavior and user interests.

These new services outpaced data protection legislation and demanded clarification of important issues regarding personal data: What is personal data? Who owns personal data? Who is accountable if personal data is compromised? How do national borders affect the global transfer of personal data? To which information are data subjects entitled to and how can they influence how their data is processed?

With the GDPR, the European Union (EU) attempts to catch up with technological and societal developments of the past two decades and intends to establish a legal framework that can handle future developments. The GDPR, which was passed in 2016 and is in force since 2018, clarifies definitions concerning personal data and the territorial scope of data processing. It extends data subject rights and enforces accountability of data processors with fines of up to 4% annual revenue. An important aspect of the principle of accountability is the obligation to document processing of personal data.

However, implementing these provisions is not a trivial task. Data protection is a highly interdisciplinary effort (Kabanov, 2016; CIPL, 2018). It requires a holistic overview of the tasks and careful planning of the integration of data protection processes. In this chapter, we motivate the construction of a reference process model to support organizations in the establishment and maintenance of GDPR compliance.

1.1. Privacy and technological progress

In the past 20 years, technological progress enabled a variety and quality of services that would not have been possible before: Social Media profiles allow connecting friends and interest groups all over the world; search engines provide personalized search results based on location, interests or past search queries; online shops make personalized suggestions based on previous purchases and purchases of other customers; personal health devices track exercise plans based on biometric measurements, and mobility services organize the most convenient transport at a given time and place.

In the course of adopting these services, individuals have handed over massive amounts of personal data to the companies that offer them - in many instances at no monetary charge. Instead, the companies[1] leverage the information that can be extracted from personal data to increase service quality or prediction accuracy. Personal data, which used to be a byproduct of other activities, has become the central resource in an economy that Schneier (2015) calls a *ubiquitous surveillance system*. Zuboff (2019) coins the term *behavioral surplus* for inherent information that does not relate to the primary cause of an interaction - e.g. conducting an online search or purchasing a product in an online store. Consequently, Zuboff calls this new data economy *surveillance capitalism*. Data, as a frequently cited article states, is the *new oil* (Economist, 2017).

Personal data and personal information are inseparably connected to the term privacy, which has been a topic of discussion for at least the past century. Famous definitions for privacy include *"the right to be let alone"* (Warren and Brandeis, 1890) or *"the claim of individuals, groups or institutions to determine for themselves when, how and to what extent information about them is communicated to others"* (Westin, 1967). (Solove, 2006, p.486) relies on Wittgenstein's concept of relatedness to characterize privacy as *"an umbrella term, referring to a wide and disparate group of related things"* and puts forward a taxonomy for these problems.

As Colesky et al. (2016) show from a computer science perspective, the Solove taxonomy covers the elementary IT-based operations that can be performed on data. We adopt this notion of privacy for this thesis, because the GDPR similarly addresses a wide range of problems that are associated with the protection of personal data.

Schneier (2015) suggests that even the notion of being observed causes an individual to act in a manner that the individual assumes will satisfy the observer's expectation, and thus poses a severe threat for freedom of speech. This phenomenon is also called the *chilling effect* (Solove, 2006, p.487).

A famous example for the chilling effect is the panopticon by Jeremy Bentham, an architectural design that introduced an invisible surveillance system. A central authority could choose to watch anyone at any time, while people who were being watched were unable to tell whether they were being watched. The panopticon is often described as a prison, but it was aimed at a broader use that was not limited to correctional facilities: the constant threat of surveillance was intended to promote efficiency and create *"a world without waste, a world in which anything left over is immediately reused"* (Miller, 1987, p.8).

[1] We will use the terms company, enterprise and organization interchangeably throughout this thesis.

The value of privacy, according to (Solove, 2007, p.15), lies not only in the fact that it frees individuals from social control, but that *"it is in itself a form of social control that emerges from the norms and values of society"*. The protection of individual privacy rights guarantees a balance of power between society and the individual, and enables individuals to flourish. Thus, it is rather a protection of the individual according to a society's own norms, such as freedom of speech (Solove, 2007).

Zuboff (2019) argues that the combination of forecasted behavior and highly effective nudging, i.e. predicting and influencing human behavior, takes away the *right to the future*, because these practices tend to create a self-fulfilling prophecy. Protecting privacy therefore means enabling individuals to lead a self-determined life.

1.2. The General Data Protection Regulation

Up until 2018, the legal provisions for processing of personal data in Europe were defined in the EU directive from 1995 (European Parliament, 1995). At the time of its publication, less than 1% of the world population used the internet (Deutscher Bundestag, 2007), compared to more than 58% in 2019 (Miniwatts Marketing, 2020). Mobile internet, social media, e-commerce, streaming services and wearable devices were still negligible, if they existed at all (Ritchie and Roser, 2019).

While stating that the general principles of data processing of the 1995 directive remain sound, the EU acknowledged that rapid technological developments and cross-border data flows have introduced new challenges to the protection of personal data (European Union, 2016, Recital 6). Further, the 1995 directive has not prevented fragmented privacy legislation across the Union (European Union, 2016, Recital 9).

With the GDPR, which was enacted in 2016 and came into force in 2018, the EU aims to address these inconsistencies in both terminology and national legislation. A regulation, as opposed to a directive, becomes effective immediately without further legislative action by a member state of the Union.

Animated by the threat of penalties of up to 4% annual revenue or €20 million, companies have engaged in large efforts to establish GDPR compliance. An industry study reports that budgets of more than $50 million have been allocated to GDPR projects (CIPL, 2018). According to another industry study from 2018, 68% of companies with 500 or more employees have spent over $100,000 on GDPR implementation before May 25, 2018 and 87% expect privacy to become even more important after the passing of the GDPR deadline (TrustArc, 2018).

Despite the general efforts to be compliant with the GDPR, not all companies have engaged or succeeded in GDPR projects. Mikkelsen and Strandell-jansson (2018) state that just before the May 2018 deadline, none of 35 interviewed companies were fully compliant, and 50% still had major gaps to address. IAPP (2020), released in December 2020, confirms the previous years' findings that less than half of the respondents in an industry survey consider their organization to be *compliant* or *fully compliant*.

These observations are supported by publicly disclosed lists of violations: At the time of writing,

CMS (2021) lists more than 500 recorded fines across European countries. The largest fine of €50 million was issued to Google in France for an insufficient legal basis for data processing (CMS, 2021). An intended fine of more than €200 million for British Airways for insecure data processing was later reduced to €22 million (CMS, 2021).

1.3. Problem statement and research questions

As the above industry surveys and penalties show, implementing and maintaining the GDPR provisions is a major challenge and a severe financial risk for organizations. The GDPR defines the objectives and obligations of data protection, but there is no standard approach for transforming an organization to a GDPR-compliant state: GDPR compliance and implementation require the integration into each company's unique processes (Kabanov, 2016), and since no two systems or organizations are the same, any effort necessarily has to be non-trivial (Sirur et al., 2018).

Organizations have to develop and deploy risk and compliance measures that incorporate stakeholders from several internal departments (Kabanov, 2016). The main concerns include the translation to a technical context, readability and ease of understanding, awareness about data privacy within an organization, and clarity about judicial interpretation. Among the issues faced during the implementation were data flow mapping, automated monitoring of data practices, protocol updates and training (Sirur et al., 2018).

Kabanov (2016) describes four major challenges in a GDPR compliance project: the complexity of the external and internal regulatory environments, the initial maturity of policies, the complexity of the IT landscape, and the large size of an organization.

According to Ayala-Rivera and Pasquale (2018), companies struggle with the extraction of requirements from legal texts, mapping legal obligations into software functionality, and understanding how to operationalize the requirements.

In a systematic literature review, Almeida Teixeira et al. (2019) find eight particular challenges in the existing academic work. Challenges that originate from the GDPR itself are its complexity, the subjectivity within the articles, and possible future changes to the regulation. Organization-specific challenges include lack of privacy-related knowledge or technology, lack of budget or human resources, or the lack of practical guidelines.

As earlier studies have concluded, fulfilling compliance requirements draws time and resources from core business activities that could be invested to gain a competitive advantage (Van Roosmalen and Hoppenbrouwers, 2008). In other words, efficient implementation and maintenance of compliance is highly desirable, because it frees up resources for core business activities.

The GDPR has been in force for more than two years, which could imply that the importance of initial implementation efforts has decreased. However, recent industry reports (IAPP, 2019a, 2020), as well as the high number of fines (as of January 2021 at least 500 fines, cf. Table 1.1) suggest that companies still struggle to establish or maintain GDPR compliance.

Further, in rapidly changing business environments, business processes and the associated pro-

cessing activities of personal data change frequently. This creates a need to monitor and maintain GDPR compliance in an efficient way in future years. Thus, sustainable concepts that foster the understanding and operationalization of the regulation are needed.

Country	Fine	Data Controller	Type
France	50 €M	Google	Insufficient legal basis for data processing
Germany	35.3 €M	H&M	Insufficient legal basis for data processing
Italy	27.8 €M	TIM	Insufficient legal basis for data processing
UK	22 €M	British Airways	Insufficient technical and organizational measures
UK	20.5 €M	Marriott	Insufficient technical and organizational measures
Italy	16.7 €M	Wind Tre	Insufficient legal basis for data processing
Germany	14.5 €M	Deutsche Wohnen	Non-compliance with general data processing principles
Italy	12.3 €M	Vodafone Italia	Non-compliance with general data processing principles
Italy	8.5 €M	Eni	Insufficient legal basis for data processing
Spain	5 €M	Banco Bilbao	Insufficient fulfillment of information obligations

Table 1.1.: The 10 highest fines until January 2021 (CMS, 2021)

Compliance in general has been addressed in IS research before, e.g. by Abdullah et al. (2009) and Cleven and Winter (2009). Cleven and Winter (2009) note that *"compliance [...] demands for a unified system of concepts and a pool of methods and models that can be combined for a holistic compliance implementation"*. Timm and Sandkuhl (2018) confirm that regulatory texts do not provide sufficient guidance on how to align an organization with the regulatory requirements.

Before the GDPR was initiated, privacy research in IS mostly focused on describing the state of information privacy and explaining what is occurring, with or without testable hypotheses (Bélanger and Crossler, 2011, p.1023). Bélanger and Crossler (2011) identify a general lack of design and action research in this field. Freitas and Mira da Silva (2018) interview representatives from ten companies and identify a lack of understanding and awareness of the obligations. As a result, they suggest the development of methods that support GDPR compliance.

Since the GDPR has been discussed, there have been notable works in the IS research domain. Based on a systematic literature review on the implementation of Privacy by Design (PbD), Kurtz et al. (2018) discover that existing publications only elaborate on the problem and objectives of a solution, but do not propose artifacts to address the issues. Therefore, we state the following research goal:

> ***Research Goal:*** To support the implementation and execution of GDPR compliance management by developing a reference process model that can serve as a blueprint for GDPR compliance approaches.

This research goal is addressed in multiple, additive steps and research contributions. We address these additive steps in five research questions, which we present and discuss in this section.

> ***RQ1:*** What are the tasks and stakeholders that have to be considered for GDPR compliance?

Firstly, we investigate the necessary tasks that have to be performed for GDPR compliance. This is achieved through analysis of the primary and secondary literature and a survey among DPOs that assesses the findings from literature and puts them into a practical perspective. Additionally, we gain an overview of the relevant stakeholders in the GDPR, both globally and within the organization. These insights will provide an understanding of the core obligations and roles that are discussed in this thesis and add to the structure of the reference process model that is developed later.

> ***RQ2:*** Which methods exist in literature to address GDPR compliance?

Findings from previous academic publications form the basis for our research. We present and analyze the existing publications that address the GDPR as a whole, in contrast to single aspects. Moreover, we analyze existing industry approaches. When the GDPR entered into force, organizations had to develop their own approaches to implement the updated obligations. In addition to previous experience, they had to rely on existing industry methods that either address the topic of GDPR compliance directly or provide a more general reference approach that can contribute to DPM. We can reasonably assume that well-known industry approaches influenced how organizations addressed the implementation of privacy legislation.

> ***RQ3:*** What are requirements and concepts of a reference process model to address GDPR compliance?

Our third research question addresses the specific requirements that a reference process model for supporting the implementation and maintenance of GDPR compliance should address. We derive these from foundational work on reference modeling, as well as a combination of expert interviews, existing research work and observations from the identified industry methods. Additionally, we incorporate requirements that were stated in the interview series that the main development of the reference process model is based on.

> ***RQ4:*** How can a reference process model for GDPR compliance be defined?

We develop the main contribution of this thesis with our fourth research question. Based on the findings from the first three research questions, we conduct a detailed ex-post investigation of the GDPR compliance efforts in 24 organizations from the perspective of enterprise architects. We claim that these approaches follow an implicit reference process model that we define explicitly in this thesis.

RQ5: How do practitioners assess the economic, deployment and engineering aspects of a reference process model for GDPR compliance?

The goal of design-oriented research is to create a relevant and sound artifact. It is therefore essential to evaluate the proposed artifact and asses its fit to the problem at hand. We achieve this by qualitative interviews with main stakeholders of the reference process model, which also yielded suggestions for improvement. After iterative adaptation of the reference process model, a short quantitative survey among stakeholders adds a further perspective of the artifact of this thesis.

1.4. Design-oriented IS research

At its core, IS research has two fundamental research approaches: the behaviorist approach and the design-oriented approach. The behaviorist approach focuses on observing and explaining IS characteristics and user behavior, while the design-oriented approach aims at developing and assessing innovative IS (Hevner et al., 2004; Österle et al., 2011).

A design-oriented IS research process has four generic steps: *1) analysis*, *2) design*, *3) evaluation* and *4) diffusion* (Österle et al., 2011, p.9). To qualify as research, as opposed to solution development, design-oriented IS research needs to comply with four basic principles:

Abstraction The artifact must be applicable to a class of problems.

Originality The artifact must advance the body of knowledge.

Justification The design decisions for development of the artifact must be transparent and traceable.

Benefit The artifact must yield benefit for the stakeholder groups.

Valid results of a design-oriented IS research process can be artifacts, such as constructs, models, methods or instantiations (Österle et al., 2011, p.9).

(Hevner et al., 2004, p.76) identify the design-oriented or *design science* paradigm as *"fundamentally [...] problem-solving"*, which is nonetheless inseparable from the behaviorist approach. To provide assistance in creating synergetic efforts between the two paradigms, the authors present a framework for IS research. We configured the IS research framework to our research project (cf. Figure 1.1).

Relevance of the research project is ensured by the real-world problem environment that the research addresses: Organizations face the challenge of implementing and maintaining the GDPR provisions under severe pressure of large fines. At the same time, the data protection efforts should not disrupt business operations excessively. Employees of the organizations are often unaware which tasks have to be fulfilled and who is responsible for which activities. Further, since the regulation is an interpretable legal document, some uncertainty in implementing the legal provisions remains. We motivate the work in detail in this chapter.

Rigor of the project is supported by the knowledge base for supporting the implementation of privacy legislation. It includes methods for designing privacy-aware systems, industry standards

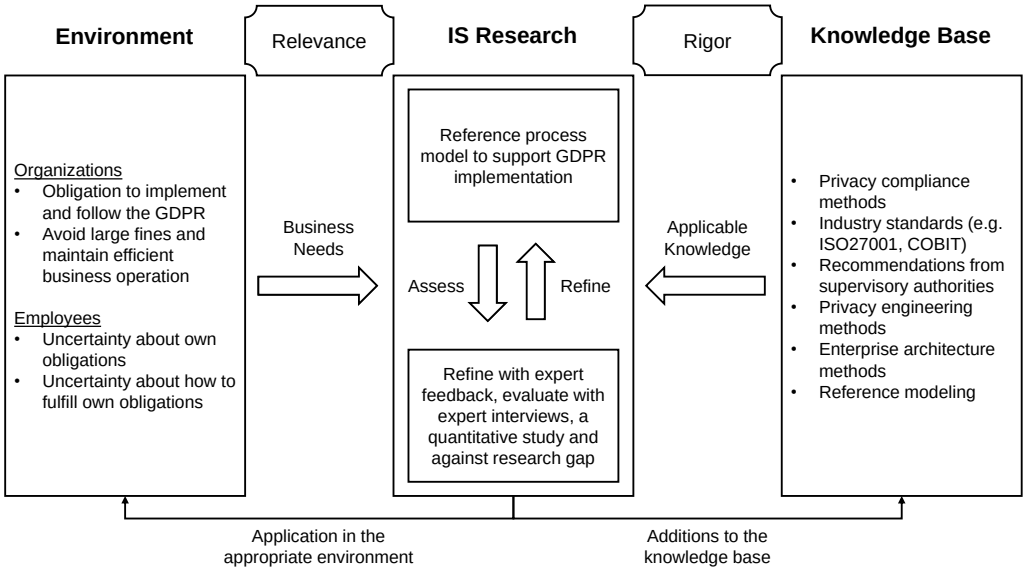

Figure 1.1.: IS research framework (Hevner et al., 2004) adapted to this research project

for managing IT landscapes and ensuring IT security, guidelines by data protections authorities on single aspects of the GDPR, as well as interdisciplinary and overarching enterprise architecture methods. Literature on reference modeling supports the construction and evaluation of the artifact.

At the center of the research framework is the construction of the artifact - a reference process model to support GDPR implementation from an EA perspective - and the refinement and evaluation of the research result.

According to Hevner et al. (2004), seven principles support effective design science research. We took these principles into account as follows:

Design as an artifact The result of our search process is a reference process model, which represents a core subject matter of the IS field and is therefore a viable artifact of an IS research process. According to Frank (2006), reference models represent ideal subjects for design-oriented research.

Problem relevance As discussed in this chapter, the implementation and maintenance of GDPR provisions presents a substantial challenge for organizations. Organizations face the risks of large fines and competitive disadvantages because of ineffective implementation of the provisions.

Design evaluation We demonstrated the quality and utility of the design artifact through qualitative evaluations with potential users of our reference process model. The evaluation

results were integrated iteratively in the design artifact. We then evaluated the revised artifact in a quantitative study with a broad selection of potential users.

Research contributions In the course of designing the final artifact, we contributed to the body of knowledge in research with results on single provisions of the regulation (data portability, records of processing activities, privacy engineering methods), as well as empirical insights into GDPR implementation projects.

Research rigor Rigor was ensured through a structured research approach, the inclusion of extensive literature sources and frequent discussions with other researchers.

Design as a search process Our research process aligned to state-of-the-art research processes for the construction of reference models, which included a refinement of the artifact. We describe our search process in detail in Chapter 5.

Communication of research Partial research results from the search process itself are published at scientific conferences in the IS community. The final artifact *ProPerData*, which mainly addresses a practitioner audience, is published as a technical report and can be downloaded via the chair homepage.

1.5. Reference models

Bichler et al. (2016) describe models as one of the most important elements of computer science and business informatics, with applications in system construction, verification, optimization, explanation and documentation (Bichler et al., 2016, p.313). They abstract their application domain and focus on the core concepts of their subject, while neglecting the technical implementation details (Frank, 2006, p.119).

By reconstructing reality, conceptual models aim to contribute to a better understanding of the problem domain. Since the goal of conceptual models is not only understanding, but improvement of information systems, they also include prescriptive elements for specific implementations (Frank, 2006, p.120). We adopt the following definition:

Definition: Model
A model is the abstraction of observations regarding the contents of a subject that serve a specific purpose. (Vom Brocke, 2003, p.16)

A process model is a special type of conceptual model that is characterized by processes being the subject of the model. It abstracts from the actual process implementation and describes the essential elements and steps of the process they describe. This description does not necessarily have to be linear.

According to (Fettke and Loos, 2003, p.36), reference models are designed as reusable artifacts to address a specified problem. Their main characteristics should be generality and recommendation character (Vom Brocke, 2003, p.31f), (Thomas, 2005). Vom Brocke defines a reference model as

Definition: Reference model
An information model that is developed or used for the construction of applied models, where the relationship between reference model and application model is characterized by reuse of the subject or content of the reference model in the subject or content of the application model. (Vom Brocke, 2003, p.34)

For practical application, reference models promise effective support by reducing the complexity of real world problems (Frank et al., 2007). Possible scenarios for reference models are, among others, the description of organizations, business process (re-)engineering, or knowledge management (Fettke and Loos, 2003, p.38). In these scenarios, the main benefits for using reference models are the increased reuse and a higher level of integration for information systems (Frank et al., 2007, p.3). (Schütte, 1998, p.209) considers the analysis of the current situation and the guidance in constructing specific models as the two main goals of reference modeling.

Fettke and Loos (2004) distinguish between reference models as *encountered artifacts*, where research has the task of reasoning about these reference models, and reference models as *theoretical constructs* that result from a scientific process. Within the realm of theoretical constructs, we adopt the notion of reference models as a *set of normative statements*, and, to a lesser extent, as a statement for a *class of organizations* (Fettke and Loos, 2004, p.333).

As with conceptual models, reference process models are a particular type of reference models. Reference process models are focused on processes or process organization. We adopt the definition by Fettke et al. (2005) for a reference process model:

Definition: Reference process model
A reference process model represents dynamic aspects of an enterprise, e.g. activity sequences, organizational activities required to satisfy customer needs, control-flow between activities, particular dependency constraints etc. (Fettke et al., 2005, p.469)

By modeling an action context, reference process models support the mutual adaptation of business organization and information system (Frank et al., 2007, p.4).

1.6. Research approach and thesis outline

Every scientific process involves a method, which requires a methodological foundation (Schütte, 1998, p.177). An explicit method allows comparability and thus serves as a quality criterion for the obtained results. In this section, we will define the method that we follow in this thesis. To this end, we adopt the following definition:

Definition: Method
A method is a prescription for attaining a certain goal. (Vom Brocke, 2003, p.58)

Reference modeling methods can be subdivided into two categories: *construction* and *application* of the reference model (Ahlemann and Gastl, 2007, p.79). Since our goal is the construction of

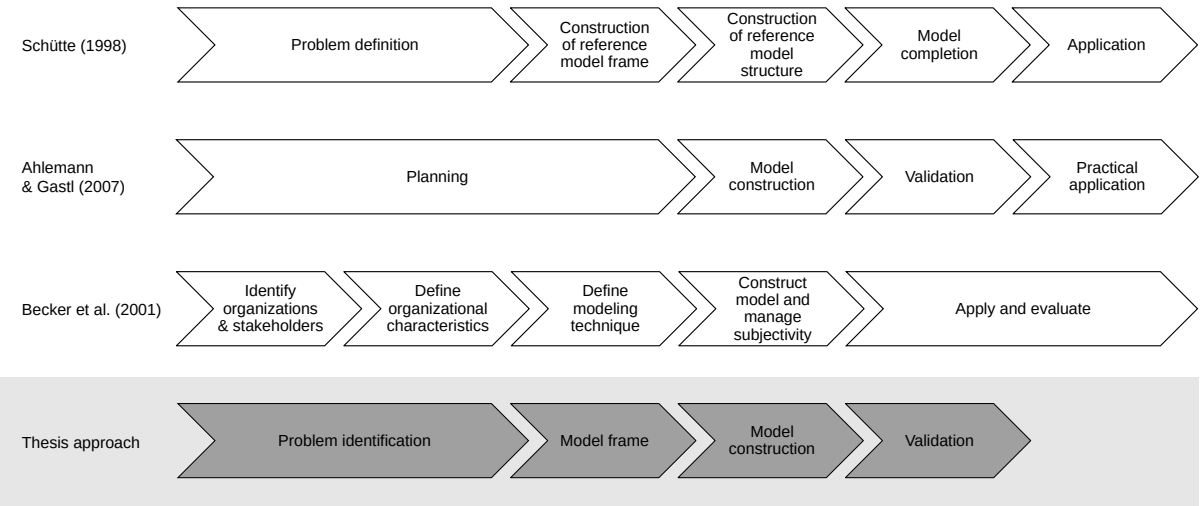

Figure 1.2.: Reference modeling approaches in literature and their relationship to our reference modeling approach.

a reference process model, we only refer to methods that support this first category of reference modeling. The approach followed in this thesis draws from Schütte (1998), Ahlemann and Gastl (2007) and Becker et al. (2001). Figure 1.2 visualizes the three approaches from literature and relates our approach to these approaches.

1. We start the research process with the **problem identification** phase. Becker et al. propose to identify the relevant organizations and their organizational characteristics first. According to Schütte, this enables the researcher to identify the problems that seem particularly important for this class of organizations. Schütte suggests that the modeler should define the problem in a top-down approach to ensure that the reference model fulfills the set goals. We define organizations that engage in EAM as the organizations that our reference process model addresses. The relevant problem is GDPR compliance management, i.e. initiating and maintaining an approach to comply with the regulations that are set out by the GDPR. This implies that the model will be used by different stakeholders in different roles. Becker et al. note that for a multi-perspective reference model, the modeler should identify the model users and consider their perspectives when creating the reference model frame.

   We first investigate the GDPR stakeholders in Chapter 2. We present results on stakeholders that are stated explicitly in the regulation, i.e. external stakeholders, based on Huth et al. (2018). Our discussion of stakeholders in internal compliance projects represents the basis for the roles in the reference process model. Further, we summarize which tasks are involved in the DPM process. We develop a list of DPM tasks which was validated by 38 DPOs (Huth et al., 2020c). The same survey also provides empirical insights into the perceived usefulness of EAM support for DPM tasks.

   In Chapter 3, we present relevant contributions from the research community and industry

methods. We discuss the suitability of the approaches for fulfilling the research goal of this thesis and derive the research gap that we address with the construction of a reference process model for the implementation and maintenance of GDPR compliance approaches.

2. The next phase in our research approach is the **construction of the reference model frame**, which should take into account the relevant organizations and characteristics from the previous phase. Ahlemann and Gastl suggest planning the inter-model relationships and using existing domain knowledge to create an initial frame of reference. According to Schütte, this can be supported by a meta model. Schütte also notes that a suitable modeling language for the reference model must be selected.

   In Chapter 4, we first discuss requirements that the reference process model must fulfill, based on general modeling requirements, the presented literature for supporting privacy compliance and our expert interviews from organizations that are relevant for our reference model. We then present and discuss the metamodel that structures ProPerData, as well as the visual language that is used to model ProPerData.

3. Third, the **construction** of the reference process model takes place. Ahlemann and Gastl propose conducting interviews, which should be recorded and transcribed. Subsequently, the reference model can be constructed from the interview results, as well as standards, norms, existing research results, own domain knowledge and possibly other sources. According to Becker et al., this construction should take into account the perspectives of the prospective model users which were identified previously. To complete the construction, Schütte recommends pointing out intra-model dependencies and enhancing the model with qualitative statements by potential model users.

   We discuss the construction of the reference process model ProPerData in detail in Chapter 5. It covers the initial construction and an iteration after the first qualitative interviews. We describe the data collection process, which incorporated empirical inquiries and literature sources. The chapter discusses in detail the elements of ProPerData and the design decisions we took during the construction of ProPerData. We emphasize the reasoning of these design decisions with quotes from the interviews. As the DPM process is highly connected with many internal processes in the organization, we also reflect on these interrelations. For the full reprint of ProPerData, we refer to the Appendix.

4. Finally, the **validation** phase investigates the model's quality. Ahlemann and Gastl recommend approaching the same group of expert interview partners for validation purposes, as this may facilitate comparability of the results. The modeler should refine the model based on proposed corrections, until the model converges to a final solutions. In contrast to the approaches proposed by Schütte, Becker et al. and Ahlemann and Gastl, our approach does not include practical application of the model. Due to the scope of ProPerData as a reference process model that addresses GDPR compliance in an overall organizational context, a real world evaluation scenario is not feasible. Instead, our evaluation is based on the approach by Frank (2006).

   Chapter 6 presents the results of eleven qualitative interviews with three different stakeholder groups, which served as input for the inductive improvement of ProPerData. A subsequent quantitative evaluation with 28 experts provides further insights into the va-

lidity of our approach. We complement the results from the quantitative study with an analytical discussion of ProPerData with respect to the requirements that we derived in this thesis.

The thesis closes with a reflection of the research process, a critical discussion of the results and potential for future research in Chapter 7. Figure 1.3 summarizes the research approach and the structure of this work.

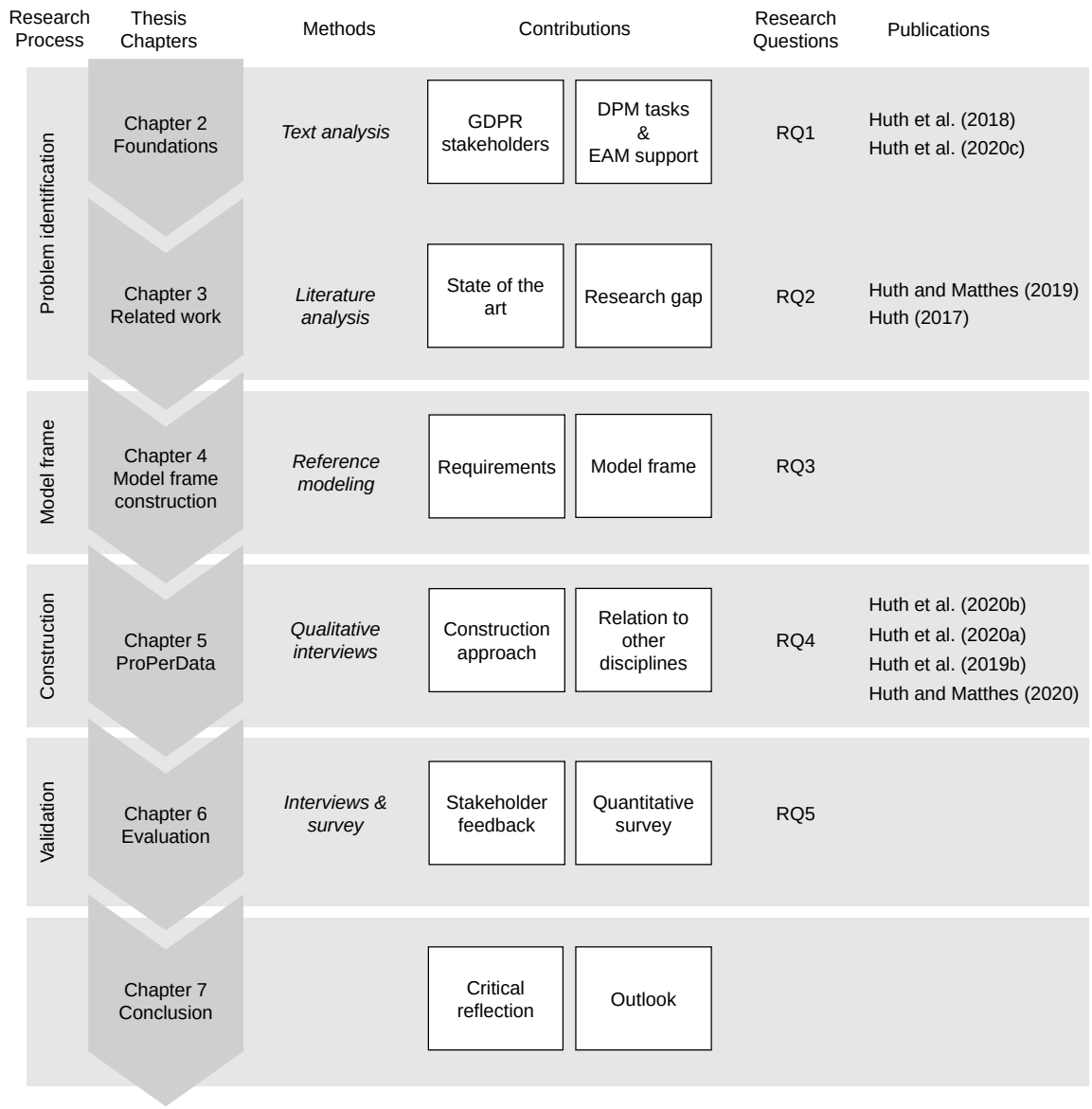

Figure 1.3.: Research approach and overview of this thesis

CHAPTER 2

Foundations

This chapter develops the foundations that ProPerData builds on. We first present key definitions from the GDPR to develop an understanding of personal data and the terminology in the regulation. Next, we investigate which stakeholders the GDPR covers in the legal text and which stakeholders are involved in GDPR compliance projects. Finally, we turn to empirical results on the tasks that DPM comprises, which we obtain from a survey among data protection professionals. Taking into account multiple scientific publications that promote the conceptual intersection between DPM and EAM, the survey also explores the support that EAM provided during the GDPR implementation from the perspective of data protection professionals.

2.1. GDPR definitions

The GDPR introduced some updated definitions regarding the processing of personal data, which better reflect the technological progress in the more than 20 years since the 1995 directive was passed. To be able to clearly describe the activities of data protection management, we will discuss key definitions of DPM related terms in this section.

A key definition of the GDPR concerns the data subject and personal data:

Definition: Data subject and personal data
Personal data 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 (1))

Article 29 Data Protection Working Party (2007)[1] discusses this definition in detail, based on the four elements *any information*, *relating to*, *identified or identifiable* and *natural person*.

The term *any information* clearly signals the intention to define the concept of personal data very broadly. Examples given by Article 29 Data Protection Working Party (2007) are data that concern the individual in a strict sense, but also professional habits and voice recordings. Biometric information, such as DNA or fingerprints, are also considered personal information, even though blood samples themselves are not.

Relating to again covers information that relates to individuals in a strict sense, but also information that can become personal, depending on the purpose (e.g. using home values to determine individual tax payments) or the result of data processing (e.g. locating taxis to optimize availability results in personal information about the drivers).

Identifiable refers to the possibility of singling out an individual in a group, even though the individual is not identified yet. This could either take place directly via a name, or via other identifiers, e.g. passport numbers. In particular, Article 29 Data Protection Working Party cites a European Court of Justice ruling that classifies IP addresses as personal data, because the internet service providers can identify individuals with manageable effort.

Fourth, a *natural person* is in effect any living human being, but can include deceased or unborn individuals as well, if information about them makes it possible to infer information about another living human being (e.g. hereditary diseases).

Table 2.1 provides examples of personal data and non-personal data, as presented by European Commission (2020).

The central concept in the GDPR for processing personal data is the processing activity, as further defined in Article 4:

Definition: Processing activity
Processing means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction. (Article 4 (2))

This broad definition, referring to *any operation* and specifying a non-exhaustive list, illustrates the breadth of the regulation. We will discuss affected internal processes in Section 5.5.

[1] The Article 29 Working Party (A29WP) was an independent advisory body to the European Commission that was established in the 1995 directive. It ceased to exist when the GDPR entered into force and was replaced by the European Data Protection Board (EDPB). The EDPB adopted the publications by the A29WP and continues to publish guidelines on single implementation issues of the GDPR.

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.: Examples of personal data and non-personal data (European Commission, 2020)

2.2. GDPR stakeholders

An important initial consideration is to asses who is affected by the regulation. In Article 2, the GDPR defines the material scope as any processing of personal data that is wholly or partly processed by automated means, but excludes the following cases:

(a) activities outside the scope of Union law;

(b) cases under specific provisions for common foreign and security policy;

(c) processing by natural persons or as a household activity;

(d) processing by competent authorities;

(e) Union institutions, bodies, offices and agencies.

Consequently, any organization that operates in Europe and processes the data of natural persons must comply with the GDPR.

However, this does not clarify the roles and relationships among the stakeholders. In this section, we first investigate the GDPR stakeholders that are defined explicitly in the GDPR (see section 2.2.1). This will help us in gaining an overview of the context that companies face. Then, as the basis for the roles of our process model, we will investigate which stakeholders are involved in GDPR implementation projects within an organization (see section 2.2.2).

2.2.1. Roles in the GDPR

There are various stakeholders that are mentioned explicitly within the GDPR. The following description is based on the study in (Huth et al., 2018), where we conducted a literature analysis to identify all actors within the 99 GDPR articles. For each of the identified actors, we extracted the relationships that appear within the regulation. Thereby, a relationship is any statement that connects an actor A (subject) with another actor B (object), e.g. "A cooperates with B".

Our search resulted in 17 entities that are named explicitly within the GDPR, as well as 33 relationships. To ensure significance of the actors, we regarded actors as main actors of the GDPR if they had at least three relationships with other GDPR actors. These main actors are:

A **Data subject** is *"an identifiable natural person (data subject) 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"* (Art. 4 (1)).

The data subject is the creator and owner of personal data and the *natural person* that the GDPR relates to. Data subjects can be customers, employees or business partners. As a special category of data subjects, the GDPR provides an extended set of rules for children (see Recital 38). We do not cover any particularities for handling personal data of children in this work.

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"* (Art. 4 (7)). The controller is in direct interaction with the data subject and must lay out how personal data is processed in the privacy statement. This thesis assumes the perspective of an organization, i.e. a legal person, as the data controller.

Processor designates *"a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller"* (Art. 4 (8)). The processor does not decide how personal data is processed. Rather, the legal agreement between the data controller and the data processor obliges the processor to adhere to the conditions of data processing that the controller stated towards the data subject. The term processor applies e.g. to cloud providers or employees of the data controller.

The **DPO** is a person that must be appointed by the controller or processor if their primary activities relate to the processing of personal data. The DPO may be internal or external and needs to have expert knowledge of data protection law (Article 37 - 39). The DPO is responsible for orchestrating the data protection activities and must report directly to the executive management.

A **supervisory authority** is an independent public authority in the member states that is responsible for monitoring the application of the GDPR (Art. 51). In the German case, there are 16 state supervisory authorities who enact the rules of the GDPR and/or their own state legislation which is based on the GDPR.

The stakeholders and the relationships between them are represented in Figure 2.1. From these main stakeholders in the GDPR, an important distinction we have to make is between the

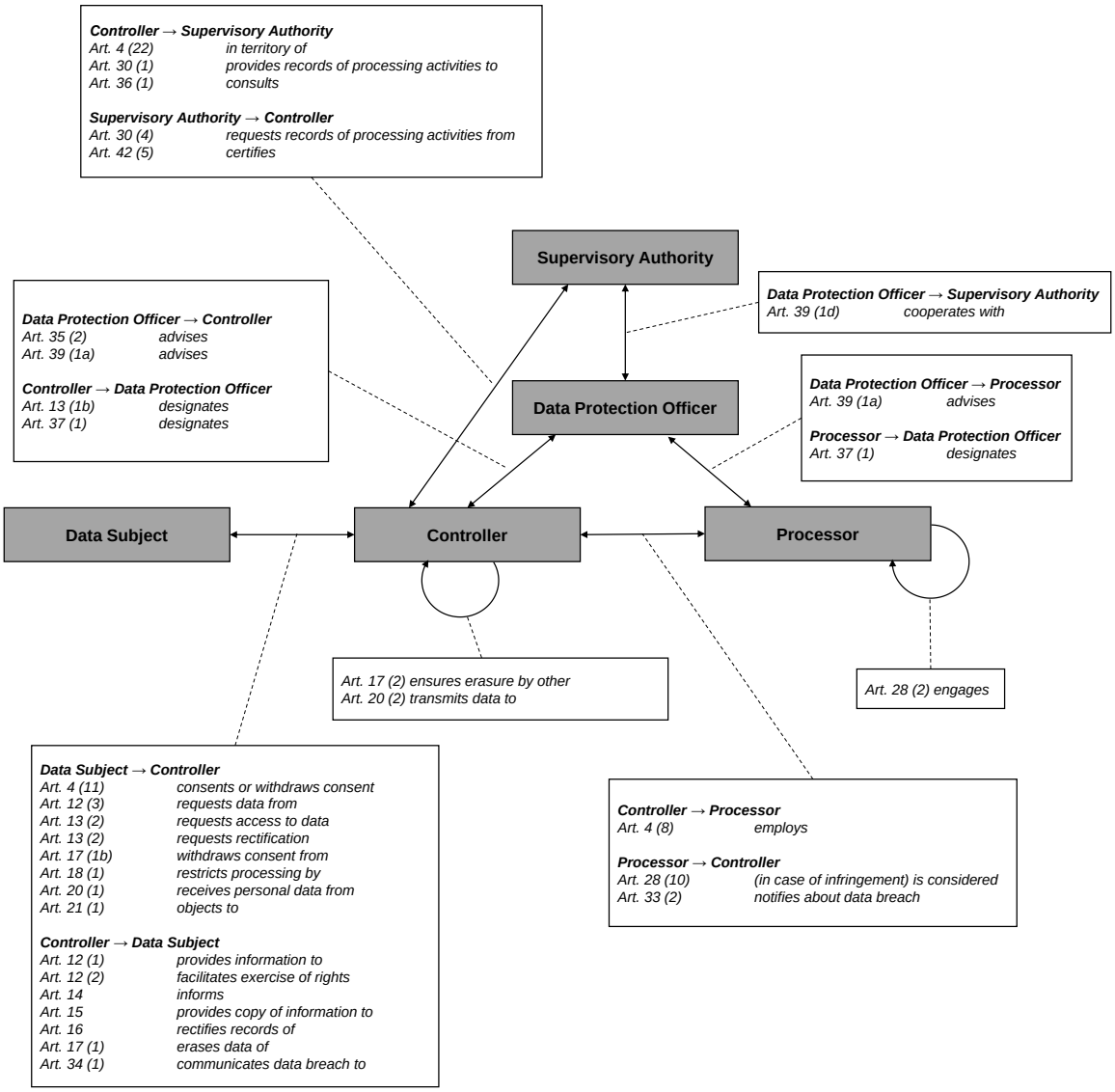

Figure 2.1.: Stakeholders and relationships between them (Huth et al., 2018)

controller and the processor. The data processor does not interact directly with the data subjects. As stated before, we assume the perspective of the data controller in this thesis. The data processor must adhere to the same security requirements as the data controller and appoint a DPO. In contrast to the controller, the processor is bound by the processing agreements and does not determine the purposes of the processing activities.

2.2.2. Roles in GDPR implementation projects

Literature sources identify the GDPR as a highly interdisciplinary challenge that involves business, legal, security, IT, as well as cross-organizational departments (Kabanov, 2016; CIPL, 2018). We investigated multiple GDPR implementation projects through interviews with enterprise architects, data protection officers and software developers (Huth et al., 2020b; Burmeister et al., 2020; Huth et al., 2019b), as well as additional discussions with stakeholders from these groups (cf. Section 5.2 for details). The following list of roles in GDPR implementation projects represents abstractions of roles that we identified in these projects. A more detailed discussion is provided in Section 5.3.2.

Although management is accountable for compliance with the GDPR, we deliberately do not include management as a role here. The reason is that if management takes an active role in GDPR projects, it can be identified within our list. The roles we present here are not strictly separated in all cases: by definition, DevOps roles combine development and operation of applications. Further, depending on the instance of the implementation project, one person can assume multiple roles at once.

DPM expert: Especially in larger organizations, the DPO is supported by a group of people, so we extend the role to the more general term *DPM expert*. The data protection expert has to command expert knowledge of data protection law and practices (Art. 37 (6)). According to the IAPP and TrustArc (2019), 81% of privacy professionals work in privacy, legal or compliance functions, while only 11% work in IT or information security departments and the remainder in some other function.

A **process owner** is the responsible role or person for a business process. A processing activity as per the GDPR is a business process that involves personal data. The process owner can provide a holistic view of the activity, especially with respect to the reasons for processing, the categories of processed data and the recipients of the data. The process owner is typically part of a business department.

The **application owner** is the person or role that ensures that the application meets the expectations of its business users. This includes its fit to customer needs and security. The application owner is especially clearly defined in the context of enterprise architecture, which we discuss in Section 2.4. However, the role exists in less explicit forms in any organization that employs software to support business processes.

To harmonize processing operations on single data objects, e.g. a customer address, some companies established the role of **data owner**. The data owner ensures that processing is in accordance with the purposes that were communicated upon data collection. A further

responsibility of the data owner is to identify and manage conflicting legislation, such as retention and immutability requirements by tax legislation vs. the data subject rights in the GDPR.

Software developers are transferring business requirements into IT solutions. We found that the role of software architects, who design the architecture of IT solutions, and programmers, who write the code that enacts the requirements, are often inseparable. Hence, we combine these two roles.

An **enterprise architect** has the task to strategically develop the enterprise architecture of an organization, i.e. the "fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and in the principles of its design and evolution" (ISO, 2011). We discuss EAM in more detail in Section 2.4.

The **IT security** department ensures the properties that make up the concept of information security. According to ISO (2018), this includes confidentiality, integrity, availability, authenticity, accountability, non-repudiation and reliability.

IT operations oversees and manages the IT services and IT infrastructure. The role is not responsible for the processing activities that are supported by these IT artifacts.

2.3. DPM Tasks

In this section, we discuss DPM tasks with respect to the GDPR. We first survey existing literature to provide the list of tasks of DPM in Section 2.3.1. The next section develops our consolidated list of tasks in DPM, based on results in Vilser (2019); Huth et al. (2020c) and Huth et al. (2020b). Finally, Section 2.3.2 presents results from a survey among 38 DPOs that assessed the difficulty of the tasks, based on Vilser (2019) and Huth et al. (2020c).

2.3.1. DPM tasks in other publications

As discussed in Section 2.2.2, we denote the group that addresses the tasks of the DPO as data protection management and use the term interchangeably within this thesis.

Recital 97 states that *"a person with expert knowledge of data protection law and practices should assist the controller or processor to monitor internal compliance with this Regulation"*. According to Article 38, this person must be able to execute these tasks in an independent manner.

The official responsibilities of the DPO are defined in Article 39: The DPO should *inform & advise* the employees of controller or processor, *monitor compliance with the regulation* by assigning responsibilities, training employees and conducting audits, support *data protection impact assessments (DPIAs)* and *cooperate with the supervisory authority*. A clarification by the Article 29 Data Protection Working Party (2016) explains that the DPO should follow a risk-based approach, i.e. prioritize the efforts on higher-risk areas.

According to Article 29 Data Protection Working Party (2016), the tasks defined in Article 39 represent the minimum set of tasks that should be assigned to the DPO, but does not limit or specify which other tasks the DPO should have. Tikkinen-Piri et al. (2017) compared the GDPR

provisions in detail with the previous directive and identified 12 key implications for companies to comply with the GDPR:

- Specifying data needs and usage
- Considering conditions for data processing in international context
- Building privacy through data protection by design and default
- Demonstrating compliance with GDPR requirements
- Developing processes to deal with data breaches
- Reckoning with sanctions for non-compliance
- Designating a DPO
- Providing information to data subjects
- Obtaining consent on personal data usage
- Ensuring individuals' right to be forgotten
- Ensuring individuals' right to data portability
- Maintaining documentation

In a small interview series to develop a questionnaire for DPOs, we assessed an initial list based on the official provisions and Tikkinen-Piri et al. (2017) and developed it iteratively together with four DPM experts (see Table 2.2). The iterative development stopped when no meaningful additions were made in the last two interviews.

Respondent	**Position**
R1	DPM, large organization
R2	external DPO
R3	external DPO
R4	external DPO

Table 2.2.: Interview partners for developing the list of DPM tasks (Vilser, 2019; Huth et al., 2020c)

The respondents added tasks for *audits* and *management reporting* and suggested summarizing the tasks that describe the execution of *data subject rights*. Further, the experts proposed separating between assessing existing processing activities and creating new processing activities, since the two tasks vary widely.

The subsequent survey (see next Section) provided the opportunity for the participants to comment on the list of tasks that we specified. Three out of the 38 participants left a comment, of which two explained general challenges in their work, and one suggested creating a certification

company. Since these comments did not refer to the task list itself, we consider the list of tasks in Table 2.3 to be validated by 38 experts in the field.

Task
Inform & educate
Verify existing processing activities
Create new processing activities
Conduct DPIA
Cooperate with supervisory authority
Maintain RoPA
Conduct Audits
Interact with data subjects
Report to management

Table 2.3.: List of DPM tasks (Vilser, 2019; Huth et al., 2020c)

Inform & educate: The employees of an organization have to be informed about data protection provisions on a regular basis. This task includes the creation of informational material, as well as organizing regular data protection trainings for employees of an organization.

Verify existing processing activities: Major changes in data protection regulation make it necessary to assess existing processing activities for compliance with the updated or changed requirements. If inconsistencies are identified, the processing activity has to be modified or retired.

Create new processing activities: In the development of new processing activities, it is crucial to include data protection considerations early on to avoid time-consuming readjustments. This includes the provisions for data protection by design and by default.

Conduct DPIA: A DPIA must be performed if the processing of personal data potentially threatens the rights of the data subject. It is the responsibility of DPM to support assessing the need for a DPIA and then advise the process owner when executing the DPIA.

Cooperate with supervisory authority: DPM is the first point of contact for the supervisory authority. Additionally, DPM may contact the supervisory authority if questions arise that concern the processing of personal information.

Maintain RoPA: The RoPA describes all processing activities of the organization in a systematic way. The tasks summarizes the interactions that are necessary to identify the processing activities and gather information about them.

Conduct Audits refers to assessing the compliance of the organization or parts thereof, including data protection and accountability, as well as the state of data protection trainings and risk management.

Interact with data subjects combines the activities that are necessary to comply with the data subject right in the GDPR (Article 12 to 22), e.g. requests for information or for deletion.

Report to management: Management is accountable for enforcing the implementation of the GDPR. It is therefore an essential responsibility of DPM to keep management informed about the implementation efforts and the compliance status.

Additionally, two more tasks emerged through interviews with enterprise architects and data protection experts (see Chapter 5): The execution of organizational tasks and leveraging data protection efforts for business impact.

Execute organizational tasks: Since all the previous tasks were focused on single processing activities that affect the same data subjects, i.e. the same individuals, there needs to be a function to combine the information of multiple different processing activities into one unified privacy statement.

Leverage data protection efforts for business impact: Organizations exist to pursue a business purpose. We observed initiatives to leverage the efforts that were associated with the GDPR implementation - e.g. data collection, process documentation, consideration of purposes of processing - for additional business insights or for strategic planning of the digitalization strategy.

2.3.2. Survey results on complexity and challenges of DPM tasks

For the study that is described in Vilser (2019) and Huth et al. (2020c), we approached the experts via the professional networks Xing and LinkedIn, based on their stated job description. Further, we reached out to data protection associations and interest groups[2] and asked to distribute the survey within their networks. In total, 38 data protection experts completed the survey. The experience in the field of data protection ranged from one to 25 years, with the 25% quantile at 3, the median at 6.5 and the 75% quantile at 10 years. The survey was conducted in August and September 2019.

As shown in Figure 2.2, the respondents spent more than 20% of their working time on the *verification of existing processing activities*. As companies advance in their compliance efforts, this proportion is likely to decrease. Support in the *creation of new processing activities* took up around 15% of working time. As with *information & education*, *maintenance of RoPA* and *conducting audits* and *DPIAs*, this task is motivated by changes in the underlying organization and should therefore persist.

The respondents spent the least amount of time on *management reporting*, *interaction with data subjects* and *cooperation with the supervisory authority*. A recent report confirms that the workload for data subject requests is less than what experts expected before the GDPR entered into force (IAPP and TrustArc, 2019, p.3).

When investigating the complexity of the DPM tasks (see Figure 2.3), we identified three groups of activities:

[2] e.g. BvDnet

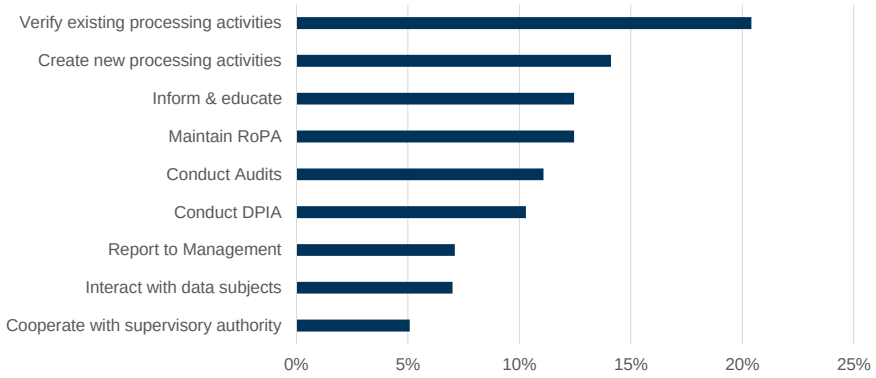

Figure 2.2.: Mean time consumption of data protection tasks (n=37)

- Predominantly **high to very high complexity**: *Conduct DPIA*, *verify existing processing activities* and *create new processing activities*.
- Predominantly **medium complexity**: *Maintain RoPA*, *interact with data subjects* and *conduct audits*.
- Predominantly **low to very low complexity**: *Inform & educate*, *report to management* and *cooperate with supervisory authority*.

As we will see in the the construction of the reference process model in Chapter 5, the wide array of organizational parts that are affected by the GDPR intensifies this complexity.

The study also asked the participants to select up to two factors that contributed most to the

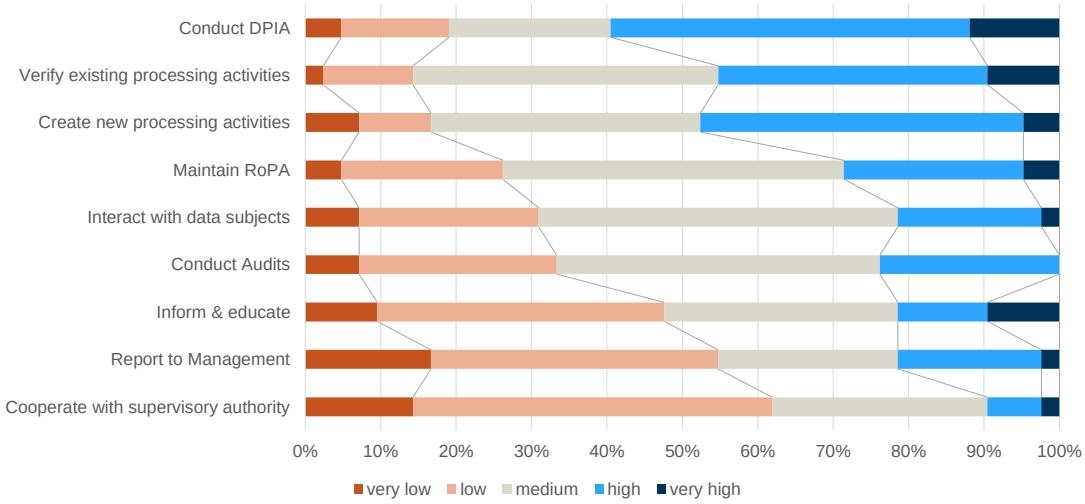

Figure 2.3.: Complexity distribution of data protection tasks (n=38)

complexity of each of the DPM tasks. The possible influence factors were *inaccuracy of legislation*, *missing tools & technology*, *lack of personnel*, *missing guidelines for practical application*, *finding the right contact person*, *lack of authority*, *missing holistic view on system landscape* and *insufficient information on single processing activity.*

The survey results (Figure 2.4) show that lack of personnel seems to be a major problem in most tasks. However, this is a problem that can only be resolved by a change in the organizational position towards data protection. Likewise, the lack of authority, especially for reporting purposes, can only be overcome by a change in the organizational structure.

Seven tasks suffer from a lack of clear guidelines and practical knowledge how to apply the data protection regulation. R1 (see Table 2.2) remarked that the knowledge will evolve as the regulation is interpreted by courts and as companies gain experience with the implementation.

Finding the right contact person, i.e. the process owner, application owner or data owner, seems to be particularly difficult in the verification of new processing activities. Assuming that the contact persons for supporting the creation of new processing activities are already known, the high value for this task suggests that our respondents interpreted this combination differently.

The experts identify other major difficulties in gaining a holistic view of the organization, especially in maintaining the RoPA, and in gaining insights into single processing activities. Lack of insights into single processes especially has an effect on the verification and creation of processing activities, as well as capturing the knowledge about this process in the RoPA and in communicating to data subjects how their data is being processed.

2.4. EAM for supporting DPM

This section focuses on how EA supports DPM in the GDPR related tasks. Architecture, according to ISO (2011), is defined as:

> **Definition:** Architecture
> The fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and in the principles of its design and evolution. (ISO, 2011)

Consequently, the EA describes all dimensions of an enterprise and their relationships (Hauder et al., 2013). Winter and Fischer (2007) list business architecture, process architecture, integration architecture, software architecture and technology architecture as essential layers that make up the EA (see Figure 2.5). The interfaces between the layers should be supported by further aggregation levels, such as products, applications or services.

EAM goals include business-IT alignment, strategic development of the EA, and increased interoperability (Winter et al., 2010). Winter and Fischer (2007) also name support for compliance as an EAM goal, because documented dependencies allow for analyses of multi-step dependencies, e.g. between server, software service and process deliverable.

Foorthuis et al. (2009) distinguish between the *descriptive overview* function and the *prescriptive framework* function of EAM. The descriptive overview provides insights that support high-level

Figure 2.4.: Most severe problems in data protection tasks

management decisions, while the prescriptive function focuses on future states of the EA (*'to be'*). The *to be* architecture guides the development of subsequent solutions.

Multiple researchers have addressed the support of compliance concerns with EAM. Acknowledging that regulations and laws have enormous impact on how organizations conduct their business, Cleven and Winter (2009) argue that holistic compliance cannot be achieved in isolated projects. The authors locate the results of a literature survey on compliance in information systems research (ISR) in a conceptual model of an enterprise architecture. At the time of the study, the Sarbanes Oxley Act, which regulates financial reporting of American public companies, was a regulation that was widely discussed by ISR. While the authors assessed the organizational influence factors as thoroughly investigated, the study identified a lack of methods and models for holistic compliance implementation. Abdullah et al. (2009) support the conclusion that proposed solutions were underrepresented at the time.

Such a solution artifact is presented by Timm and Sandkuhl (2018), who develop a reference model for a compliance organization. The model captures compliance domains of financial institutions (anti money-laundering, know your customer and fraud) and presents viewpoints for each concern, such as the fraud prevention processes. Similarly, Jugel et al. (2018) present a metamodel to integrate control objectives in EA viewpoints. Timm and Sandkuhl (2018) conclude that a reference model is an important tool to overcome challenges that arise from the background of different stakeholders (p.14). According to the authors, EA is a suitable

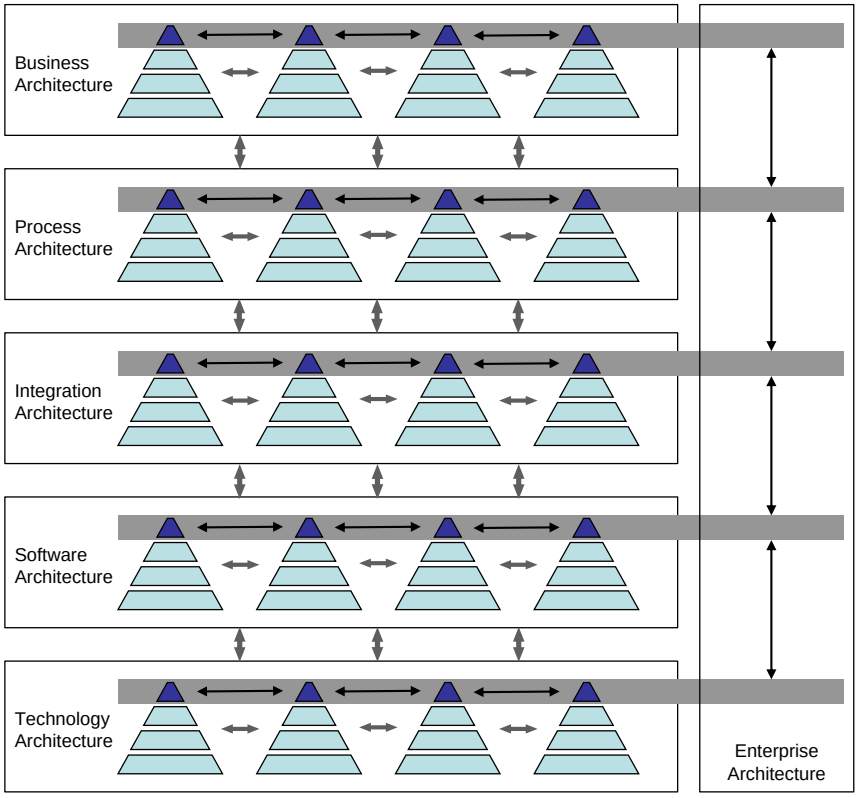

Figure 2.5.: Enterprise Architecture as Cross-layer view (Winter and Fischer, 2007)

way to capture the different elements of such a management system, including the mutual dependencies.

When considering the tasks for DPM, there is an apparent connection to elements of an EA model (cf. Figure 2.6):

- The task **inform & educate** affects the entire architecture, because many different people across the organization develop these elements and therefore need to be aware of legal obligations. Similarly, to **cooperate with the supervisory authorities**, the DPO needs to have a high level overview of the entire architecture in order to fulfil requests for information. Also, the GDPR makes it mandatory to document aspects on all levels of the enterprise.
- **Verify existing processing activities**, **create new processing activities**, **conduct DPIAs** for processing activities and **maintain the RoPA** require a deep insight into the processes that the organization engages in, as well as an understanding of the business capabilities that structure the organization. Further EA aspects for these tasks are *adherence to principles & standards* and *security of data processing*.
- To **conduct audits** and **interact with data subjects**, the DPO again has to be informed about the processes, the employed applications and the responsible people behind them.

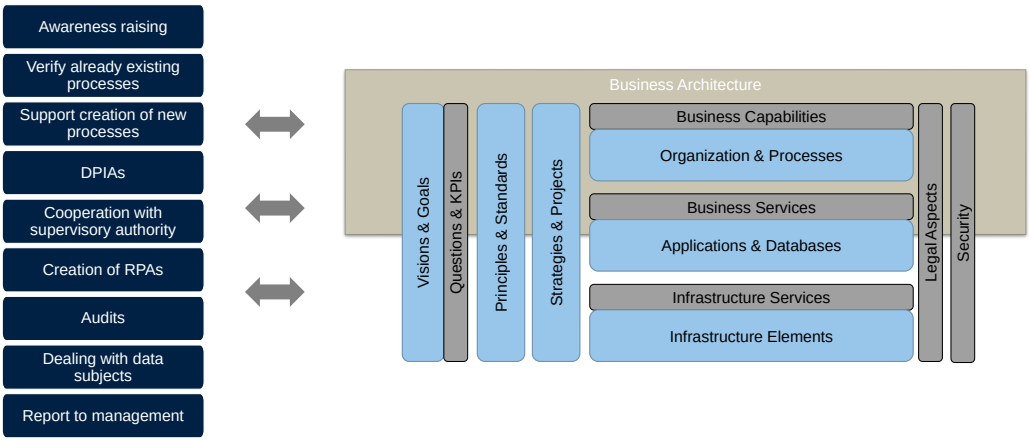

Figure 2.6.: Relation of DPM tasks to EAM elements, based on (Huth et al., 2020c) and (Buckl, 2011, p.3)

- For **reporting to management**, the DPO requires aggregated KPIs and the context of data protection challenges.

The influence of GDPR requirements on EA elements is also discussed in academic literature (Wichmann et al., 2020; Rozehnal and Novak, 2018). As the regulation affects processes, applications, data and people, knowledge about their relationships can facilitate the implementation of specific measures (Rozehnal and Novak, 2018). Wichmann et al. (2020) analyze three established EA frameworks (ToGAF, DoDAF, NAF) for their applicability towards GDPR implementation and conclude that the metamodels of these frameworks are capable of capturing GDPR-related information. Burmeister et al. (2019) elaborate this relational knowledge in an EA metamodel that captures privacy and security related information. This transparency provides the informational basis for GDPR measures, e.g. which systems process which kind of data and which security measures are already in place.

Labadie and Legner (2019) propose a framework that translates the GDPR requirements into business capabilities. The authors distinguish between organizational capabilities and system capabilities, reflecting the technical measures and the organizational measures that the GDPR addresses. We will further discuss this approach in chapter 3.

These existing publications focusing on GDPR implementation from an EA perspective motivate the question how this conceptual overlap has been leveraged in GDPR implementation approaches. Our study in (Huth et al., 2020c) investigates how DPM experts were supported by EAM during the implementation of the GDPR. From the 38 DPM experts who participated in the study, 12 (32%) stated that they received support from EAM.

As Figure 2.7 shows, the proportion of organizations where DPM and EAM collaborated increased with the organization's size. In organizations with 1.000 to 10.000 employees, 50% of DPM experts relied on support from EAM. Although none of the respondents worked for an organization in the largest category, responses from enterprise architects suggest that the

proportion in organizations with more than 10.000 employees is even higher (cf. Huth et al. (2020b)).

Independent of the organization size, the main reason for not collaborating with EAM was that it does not exist in the organization of the respondent. Less frequent reasons for not collaborating were unawareness of EAM or the right contact persons, the view that it does not provide any benefits or the objectives do not match, or simply that there was no time to invest in the collaboration.

To identify the DPM tasks that can be supported, our study in (Huth et al., 2020c) asked the DPM experts to assess the usefulness of EAM for DPM tasks, as shown in Figure 2.8. The results show the answers to the question *how likely is it that you would recommend EAM to a colleague?*, which was answered by 38 DPM experts. Since only 12 of the respondents relied on the support of EAM and 26 did not, the results are represented as percentages.

Applying the terminology of the NPS (Reichheld, 2003), there were two *detractors* (value ≤ 6; 17%), 4 *passively satisfied* (values 7 and 8; 33%) and 6 *promoters* (value ≥ 9; 50%) in the group who had collaborated with EAM. Overall, this group had an NPS of 33%. Conversely, the group who had not collaborated with EAM had an NPS of -38%.

Figure 2.9 shows the assessment of the usefulness of EAM for each individual DPM task. As the red bars exhibit, the tasks *conduct audits*, *report to management*, *inform & educate*, *interact with data subjects* and *cooperate with supervisory authority* were not supported by EAM in all cases.

At least half of the experts considered EAM support to be very helpful or extremely helpful in the creation and verification of processing activities and in maintaining the RoPA, which summarizes the processing activities. A possible explanation is that the relational knowledge that is captured by the EA provides assistance for identifying the stakeholders and applications that are involved in the processing activities.

Only three respondents used the opportunity to leave a free text comment after filling out the survey. Two of the responses referred to the lack of certification standards, which makes it difficult to audit the working methods of data protection managers. One expert also pointed out the challenges of collaboration within the organization:

> *"The cooperation in the association for example is incredibly difficult. Interest groups are often divided on basic issues."*

Another expert gave a valuable insight into the contrast between the work as an internal DPO and work as an external DPO:

> *"When working for a large organization, I had free space to perform my function as a corporate DPO. Now as an external DPO for small companies that obtain their hard- and software from external providers this is much harder, because the providers charge for every request they receive."*

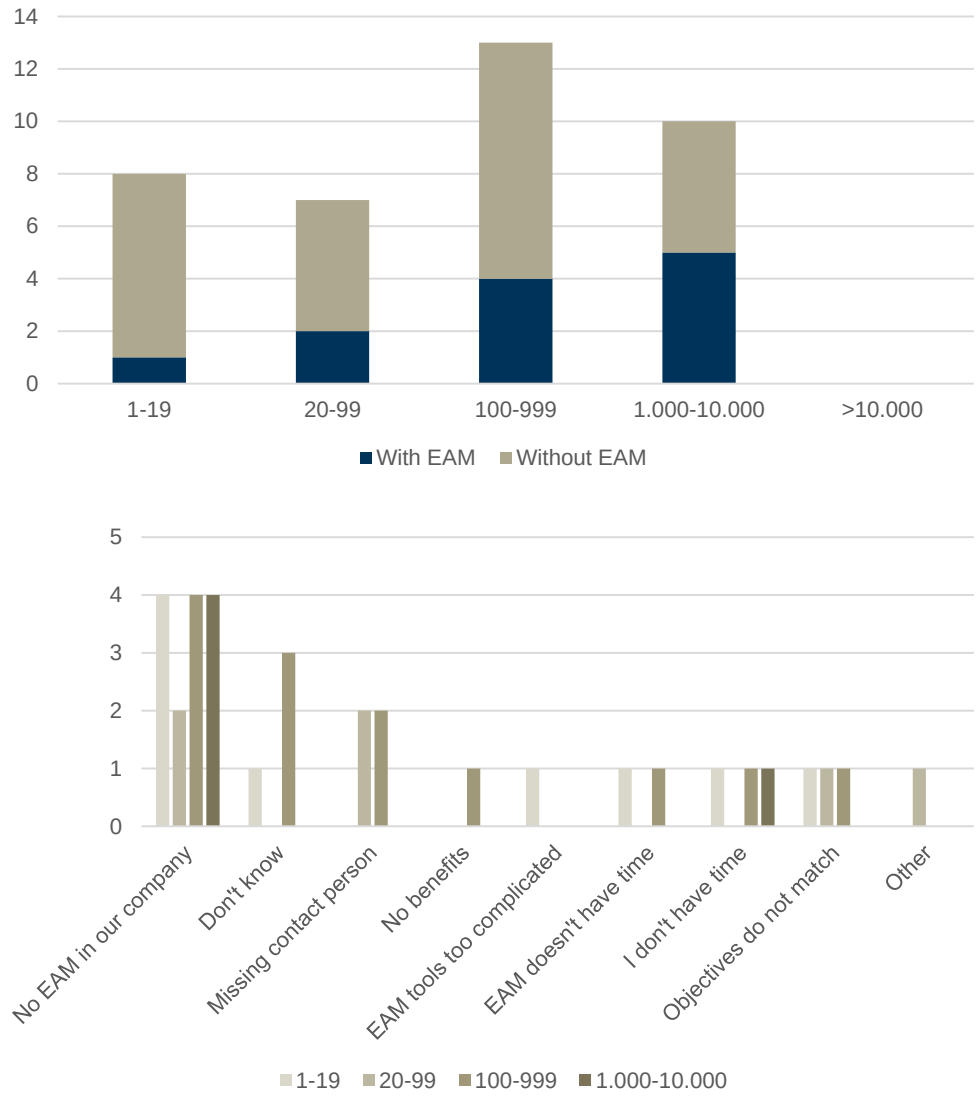

Figure 2.7.: Proportion of DPM experts who were supported by EAM and reasons for not collaborating with EAM, by organization size (Huth et al., 2020c)

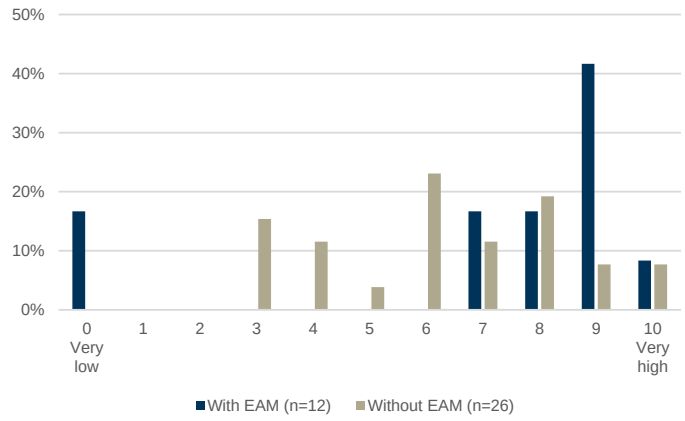

Figure 2.8.: Overall perceived usefulness of EAM support for DPM (Vilser, 2019)

2.5. Summary

In this chapter we presented the key definitions for personal data and processing activities, as well as examples for personal data. The main external stakeholders in the GDPR are data subjects, controllers, processors and their respective DPOs, and the supervisory authorities. Within an organization, we identified the roles DPM expert, process owner, application owner, data owner, software developer, enterprise architect, IT security and IT operations as essential in GDPR implementation approaches.

According to our survey in (Huth et al., 2020c), nine tasks describe the responsibilities in DPM. Out of these nine tasks, especially conducting DPIAs and verification and creation of process-

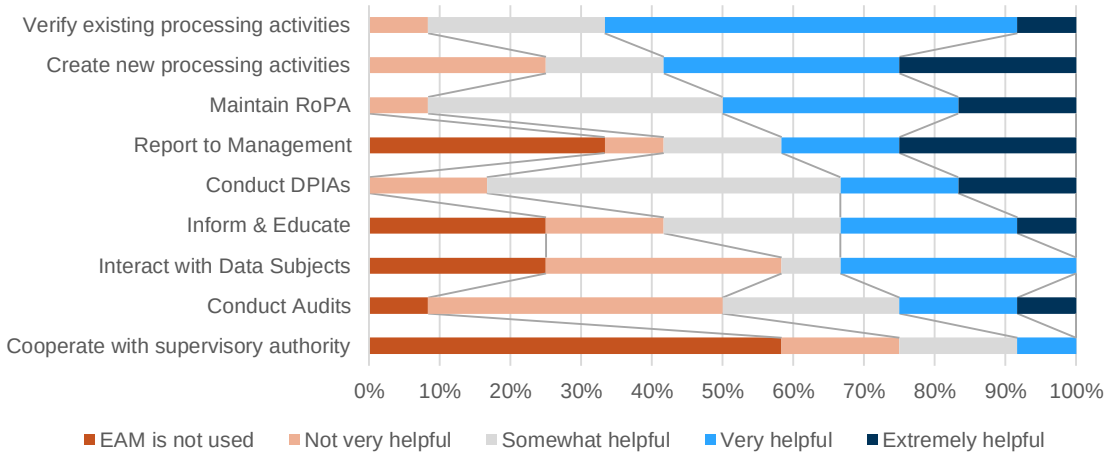

Figure 2.9.: Usefulness of EAM for each DPM task (n=12) (Huth et al., 2020c)

ing activities are particularly complex and time consuming. DPM experts identified missing personnel and the lack of clear guidelines as general problems that affect the majority of DPM tasks. Problems that affect particular DPM tasks include the knowledge about single processing activities and the holistic perspective on data processing activities in the organization. Further free text comments mentioned the difficulty in collaborating among different departments or, in case of external service providers, different companies.

To alleviate some of the mentioned problems with GDPR implementation, the IS literature has advanced the application of EAM concepts to address data protection challenges. The reasoning for this work is clear: Both EAM and DPM address the same enterprise and share concepts such as the organizational and the technological architecture. The important relational knowledge in EAM is equally important for DPM, because only the right contact partners can provide deep knowledge of architecture elements that are relevant for data protection.

Our survey therefore also investigated to which extent DPM experts relied on EAM support during the implementation of the GDPR. We found that the proportion of organizations where DPM and EAM collaborated increased with the organization size, up to 50% in the group of organizations with 1.000 to 10.000 employees. The main reason for not collaborating with EAM was that the function did not exist in the respective organization.

DPM experts considered EAM support as particularly helpful for verifying and creating new processing activities, as well as in maintaining the RoPA. Among the group of experts that collaborated with EAM, we observed strong endorsement for this collaboration (NPS 33%), compared to predominant scepticism among the group which had not (NPS -38%).

CHAPTER 3

Related Work

In this chapter, we present approaches that contribute to the goal of overall GDPR compliance. First, we introduce the scientific approaches that have been published so far to address the topic of overall GDPR compliance. Then, we discuss research work from the ISR discipline that is concerned with single aspects of GDPR compliance, but does not classify as holistic approach. The selection of methods partially draws from Huth (2017) and Huth and Matthes (2019).

Section 3.2 first covers the Standard Data Protection Model (SDM) of the German data protection authorities, which is the only self-contained holistic approach for GDPR compliance that we identified. Well-established holistic approaches for IT governance and management are presented next. These approaches do not address GDPR compliance specifically, but a broad range of IT management concerns, which also include data protection. Lastly, we shortly discuss other relevant approaches from industry publications.

3.1. Academia

The academic community has made many meaningful contributions to the field of privacy engineering, but there are only few approaches that we know of with the depth and breadth to be considered a full approach to support GDPR implementation. We will first present these approaches and then discuss other notable contributions to the field in this section.

3.1.1. PRIPARE

PRIPARE (Preparing Industry to Privacy-by-design by supporting its Application in Research), a multi-year project to prepare for the GDPR, published its privacy and security by design

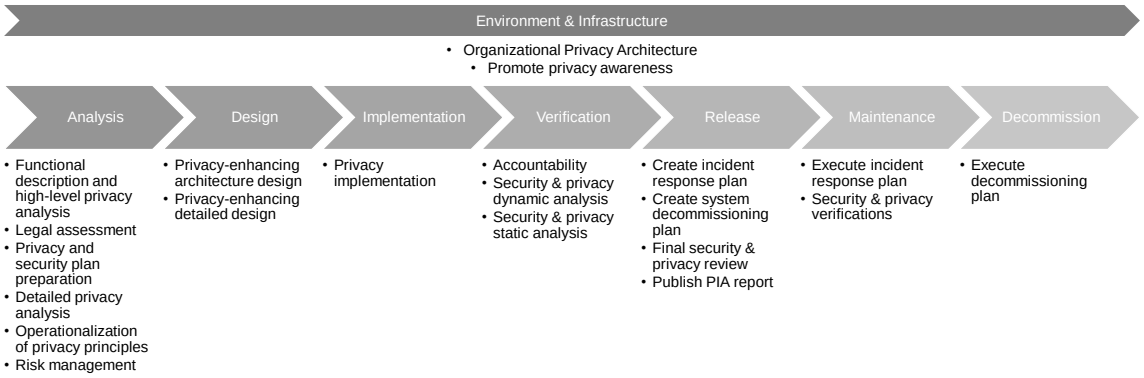

Figure 3.1.: The PRIPARE method (Crespo et al., 2015)

method to ensure compliant systems engineering in 2015 (Crespo et al., 2015; Notario et al., 2015).

The method describes six categories of roles, of which the first two are further subdivided:

System engineers are responsible for realizing successful system implementation, in which all relevant aspects are considered. The sub-roles include business analyst, system designer, system developer, UI designer, and tester.

Privacy & security managers & officers are the executive roles that are responsible for privacy risk management and include the sub-roles privacy & security engineers and privacy & security officers.

Data protection authorities are the independent bodies in charge of supervising the application of privacy regulation, advising companies in their implementation and serving as contact point for data subject complaints.

Data subjects are the individuals whose data is being processed.

Project managers are described as the senior executives in charge of the scope, costs and schedule of a development project.

End users are the employees who ultimately operate the implemented system.

Seven phases structure the PRIPARE method (cf. Figure 3.1). Each process is assigned to one of the phases and all processes are described as a SIPOC diagram (Supplier - Input - Process - Output - Consumers).

1. Within the **analysis** phase, a functional description of the planned system and its capabilities is created, including a high-level privacy analysis. Next, the privacy analysis is refined based on legal opinions. Then, tasks and responsibilities must be defined. In a detailed privacy analysis, stakeholders, roles, personal data and its flows, as well as the required privacy controls are captured. An important aspect is the operationalization of privacy principles. In a risk management process, possible threats, risks and treatments should be evaluated and balanced.

2. The second phase, **design**, is concerned with the design of a privacy enhancing architecture at a suitable level of detail. PRIPARE proposes formal methods as a suitable approach to ensure the desired properties.
3. Third, the **implementation** phase covers the actual programming of the planned system. This should include the selection and implementation of specific libraries and platforms. Of course, the programming should follow the privacy enhancing architecture from the previous step.
4. **Verification** of the privacy design is achieved through the documentation of measures that are implemented to ensure compliant processing. Another option is to conduct static analysis based on a formal representation of the system, or to execute test scenarios.
5. After the successful verification phase, the **release** phase addresses the creation of incident response plans, system decommissioning plans, the final security & privacy review and the publication of a privacy impact assessment (PIA) report, which includes the results of all previous steps.
6. During the **maintenance** phase, incident response plans must be executed as needed. Continuous security & privacy reviews ensure that privacy requirements are still met.
7. The **decommissioning** phase involves the execution of the decommissioning plan, which specifies how data should be deleted and how the deletion should be validated. The decommissioning plan is part of the system design.

As an encompassing phase of PRIPARE, **environment & infrastructure** addresses the factors that include the organizational aspects of implementing privacy regulation through the creation of a governance framework. PRIPARE hints at a few options on how to do this, but leaves it to the reader to investigate further. Lastly, the phase suggests promoting privacy awareness within the organization through codes of conduct, e-learning and internal communication.

Discussion

The PRIPARE method addresses the development process of a processing activity in detail. It includes the planning phases, as well as steps to ensure the compliant decommissioning of the processing activity. With respect to the full GDPR requirements, such as the RoPA, the DPIA, or the data subject rights, PRIPARE does not go into further detail. Therefore, the approach is mostly suitable for development of single processing activities and should be complemented by another approach for the organizational activities in regulatory compliance projects.

3.1.2. Capability-based approach by Labadie and Legner

Labadie and Legner (2019) adopt a data- and resource-based view to represent the capabilities that are necessary to comply with the GDPR. In particular, they distinguish between system capabilities, which reflect the technical measures required by the regulation, and organizational capabilities, which reflect the necessary organizational measures. With the use of capabilities,

the authors aim to clearly describe *what* the organization should do, as opposed to *how* the specific implementation should be carried out.

The main system capabilities are *defining the scope of protected data*, *managing consent*, and *enabling data processing rights*, whereas the main organizational capabilities include *orchestration of data protection activities*, *demonstration of compliant data processing* and *disclosing information*.

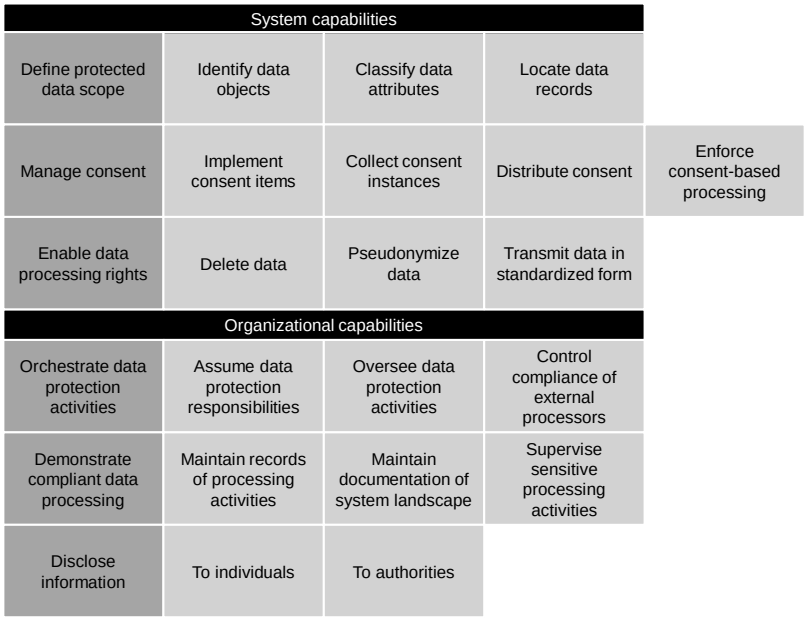

Figure 3.2.: Capability model for data management (Labadie and Legner, 2019)

Identification of the data objects for which the GDPR applies, the classification of data attributes by sensitivity levels and identifying the storage instances of personal data objects fall within the first category, *define protected data scope*.

Managing consent involves implementing consent items for processing activities, storing consent instances that are expressed by individuals, distributing this consent to all affected processing systems and making sure that all processing activities take the consent decision into account.

The category *enable data processing rights* is based on the data subject's rights and the principle of privacy by design. It encompasses the sub-capabilities deleting data of individuals, pseudonymizing personal data to minimize usage of identifiable data, and the capability to transmit personal data to another controller, as required by Art. 20 (cf. Huth et al. (2019a)).

Orchestration of data protection activities comprises assigning responsibilities for data protection tasks, including who should lead the overall data protection activities, and monitoring compliance of external processors with the GDPR. To *demonstrate compliant data processing*, the organization must possess the sub-capabilities to document processing activities in a RoPA, to document the system landscape, and to supervise data processing activities that pose special risks to the rights of the data subjects. *Disclosure of information* addresses the capability to

inform individuals about the data processing and to collaborate with supervisory authorities upon request.

To apply the capability model, Labadie and Legner propose assessing the compliance level for each sub-capability, defining a to-be state for the overall capability model and identifying the compliance gaps between the as-is and the to-be states of the capabilities. From these compliance gaps, it is possible to derive an action plan. The publication does not yet specify how to derive such an action plan.

Discussion

The approach subdivides the variety of GDPR requirements into six clearly cut capabilities and therefore provides a clear concept of the tasks that have to be fulfilled to achieve a state of compliance with the GDPR. However, the focus of the approach lies within the overview perspective, rather than how to transform a single capability from a unsatisfactory compliance level to compliance. Further, as the authors discuss, the approach currently does not take into account the ongoing nature of GDPR projects.

3.1.3. Method by Koç

Koç et al. (2018) identify the lack of actionable checklists or best practices for implementing the GDPR provisions and develop a method to initiate GDPR projects from an EA perspective. The method includes the roles *project team*, the *managing director*, *business unit managers*, *employees*, and *IT operations*.

Five phases define the method:

1. In the **project preparation**, requirements are collected, a project plan is created, awareness for the topic is raised, and the final project plan is communicated to the stakeholders.
2. For **collecting and categorizing systems**, the authors propose using applications because of two reasons: information about applications is easy to retrieve from EA documentation or a configuration management database (CMDB), and it facilitates the identification of data flows and technical and organizational measures. The GDPR project team then checks and documents the relevance of each application.
3. The objective of phase **definition of data categories** is to classify the data objects, e.g. addresses or financial information, and the data subjects, e.g. customers or employees.
4. During the **data collection and validation** phase, the information for the RoPA is collected and discussed in multiple workshops within the business units.
5. Finally, the **creation of the RoPA** takes place, which requires a consolidation of the collected activities.

Koç et al. also discuss the relationship of the RoPA to EA artifacts, e.g. applications, roles and organizational units, business processes, business objects, and risks and controls.

Discussion

The method was developed from practical experience and provides clear steps to take when initiating a GDPR implementation project. As the authors note, the published process for the creation of an RoPA is only one part of the method. Considering only the published artifact, the method considers one task of the GDPR from the perspective of its main stakeholders, which are the DPO and the enterprise architects.

3.1.4. Other contributions

Most scientific contributions concentrate on single aspects of the GDPR or, in publications prior to that landmark legislation, on privacy regulation and privacy goals in general. Bellotti and Sellen (1993) first developed a framework to assess and mitigate the impacts of a type of videoconferencing tool on user's privacy. After 2012, the number of published articles in the field of *privacy engineering* has increased sharply (Gürses and Del Alamo, 2016). As the field itself considers a multitude of aspects, ranging from technologies to formal representations of privacy problems to methodical aspects, we present only a selected set of publications that relate to the GDPR explicitly.

Kurtz et al. (2018) present a literature review on scientific contributions to operationalize the principle of transparency, which is one of the principles of privacy by design and therefore one of the guiding principles of the GDPR. The authors analyze which type of artifact the scientific contributions provide - concept, model, method and/or instatiation - and which steps of the design science process by Peffers et al. (2007) the contributions cover. As a result of the study, Kurtz et al. (2018) identify a lack of research that contributes in establishing privacy by design in organizations, and especially for third-party data processors.

Rösch et al. (2019) make use of the well-established concept of patterns as *a recurring solution to a reoccuring problem.* The authors develop a privacy pattern catalog of 13 patterns in three categories:

(a) general privacy control patterns;

(b) patterns for data subject rights;

(c) patterns that reflect the obligations of data controllers and data processors.

Each of the patterns relates to specific GDPR requirements, explains the resulting challenge and proposes a technical solution to the problem. An example for a pattern is 'storage limitation', which proposes a data lifecycle. The approach, as exemplified in the publication, consists in checking and documenting the 13 patterns one by one. It does not consider different roles.

Ayala-Rivera and Pasquale (2018) present a six-step approach to support software developers in eliciting solution requirements that fulfill the GDPR requirements. The approach considers the roles data processor, data privacy expert, IT professional, legal expert and governance. In the approach, a data audit form serves for the creation of a data inventory (step 1). A gap analysis leads to possible areas of improvement (step 2). The IT professional and the data privacy expert collaboratively determine corrective actions to the identified gaps from the previous step

(step 3). This plan is reviewed by legal experts and the governance team (step 4). The IT professional implements these corrective actions (step 5), which are subsequently reviewed by the data privacy expert and the legal expert (step 6). The authors propose using a catalog of suitable privacy controls that maps the controls to privacy requirement, e.g. access control fulfills the requirements lawfulness, purpose limitation, confidentiality and accountability. Similar to PRIPARE (Notario et al., 2015), this approach focuses on system development and does not discuss further organizational aspects in detail. The approach is validated through an analytical discussion.

Burmeister et al. (2019) argue that EA modeling, with the goals of transparency, consistency and measurability of business and IT components, is suitable to address the challenge of ensuring GDPR compliance. Therefore, the authors develop a privacy-driven EA meta-model that captures the elements and relations that are necessary to cover the GDPR-related stakeholder concerns. The authors divide the GDPR requirements into the four categories *compliance with superior principles*, *information obligations*, *satisfaction of data subject's rights* and *implementation and verification of technical and organizational measures*. For each of the four categories, the corresponding EA artifacts are presented, resulting in a meta-model that, if instantiated in an organization, can supply the relevant GDPR information requirements. Examples in the publication are the purposes for processing personal data or which applications process personal data. Despite the valuable conceptual contribution in the publication, the authors do not propose an approach how to use the meta-model to attain GDPR compliance.

3.2. Industry

3.2.1. Standard data protection model (SDM) of the German data protection authorities

The SDM of the German data protection authorities was released as version 2.0 in November 2019. It is intended as a tool to ensure and prove that processing of personal data follows the requirements of the GDPR (Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder, 2019, p.7). The focus groups of the SDM are data protection authorities and the people that are responsible for processing personal data. For the latter group, it should serve as support in planning, implementing and operating processing activities.

Seven data protection goals build the basis of the SDM. These data protection goals include the well-established security goals confidentiality, integrity and availability. In contrast to information security, these protection goals are not considered from an organizational perspective, but from the perspective of the data subject.

Data minimization Processing must be restricted to the extent that is appropriate and necessary for the specified purpose. The minimization principle holds for the amount of data that is processed, the extent of processing operations, the limitation of storage periods, and the limitation of accessibility by employees.

Availability Personal data and the information about processing must be accessible to the data

subject. Further, the processing systems must be resilient to ensure access and recoverability in case of failure.

Integrity Personal data must remain unchanged, complete, correct and up to date. Automated decision processes must be non-discriminatory.

Confidentiality It must be ensured that access to personal data is only granted to individuals who need it to fulfill their personal function.

Unlinkability Combining or linking personal data that has been collected for separate purposes is prohibited.

Transparency Data subjects, data controllers and supervisory authorities must be able to understand the basic information about processing activities, such as the purpose, the process and the responsibilities.

Intervenability The data subject must be able to influence the processing of personal data, e.g. object to processing, request rectification or request deletion.

These seven protection goals systemize the more granular requirements that the SDM derives directly from the GDPR. The SDM also uses them to systemize generic protection measures for practical implementation.

For discussing the documentation of processing activities, the SDM uses three conceptual layers: (1) the business layer, which addresses the purpose of data processing; (2) the application layer, which must ensure that the purpose limitation is respected; and (3) the IT infrastructure layer, which provides the technical measures to protect the purpose limitation. Processing activities that are likely to result in risks to the rights and freedoms of data subjects must be further assessed and adequate measures have to be taken.

To apply the SDM in the context of DPM, the publication defines the DPM process as a plan-do-check-act (PDCA) cycle, as shown in Figure 3.3.

The SDM also discusses the relation to IT security, using the example of BSI Grundschutz[1]. The two approaches share the protection goals confidentiality, integrity and availability, as well as the modeling approach for processing activities. However, the SDM is designed to protect the individual, while IT security aims to protect the organization. Various cross-references between the SDM and BSI Grundschutz ensure that these two standards complement each other.

Discussion

When the GDPR entered into force in 2018, the SDM had not been published in the version that specifically targets the GDPR. Its predecessor, version 1.1, was not mentioned by any of our interview partners, nor was it cited in the many publications that addressed GDPR compliance in the ISR field.

Granular protection goals originate from the seven higher-level principles in the GDPR. Single requirements (Articles) are assigned to these protection goals. For each protection goal, possible

[1] IT Security Guidelines by the German Federal Office for Information Security

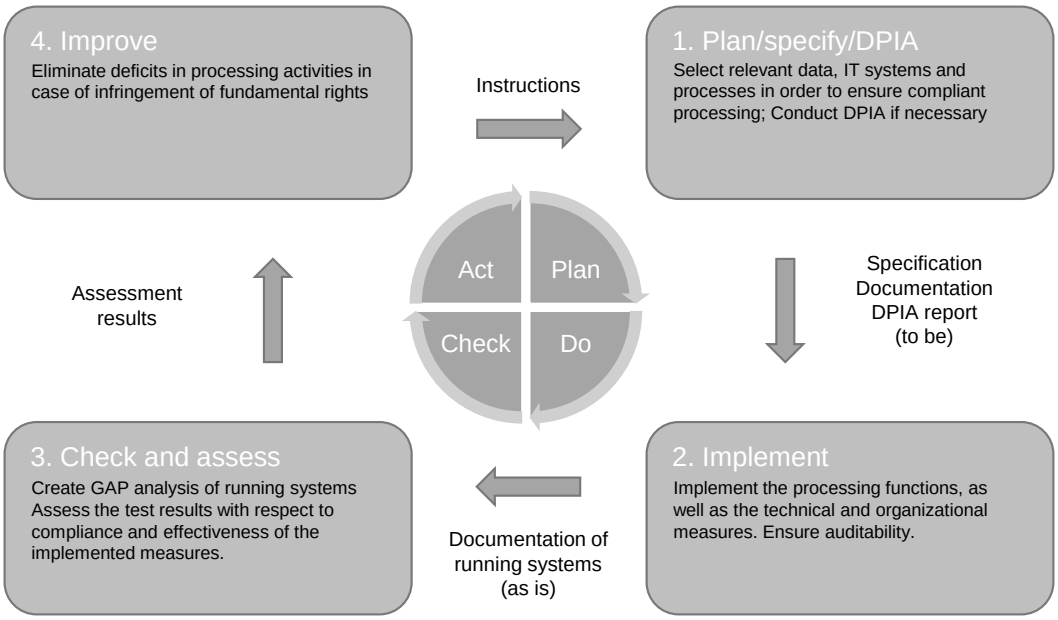

Figure 3.3.: The DPM process (Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder, 2019, p.54)

solution approaches are listed. The fact that the approach was published by the 'conference of the German data protection authorities' makes it a reliable source for avoiding penalties from supervisory authorities. Due to the list format of the protection goals, the approach supports checking off all aspects.

The approach is clearly addressed at the DPO as the main stakeholder. Other stakeholders receive only limited support, e.g. software developers or product owners through the examples for possible solutions to cited requirements.

3.2.2. ISO 27001

ISO2700x is a family of international standards for establishing an information security management system (ISMS). In the public discussion that accompanied the GDPR, this standard was frequently mentioned. We describe the approach of ISO27001 (requirements for an ISMS) based on Brenner et al. (2011) and Klipper (2015).

The basic protection goals of information security are confidentiality, integrity and availability of information, complemented by the properties *authenticity*, *accountability*, *non-repudiation* and *reliability* (Klipper, 2015, p.17). According to Brenner et al. (2011), the process to define, implement, monitor, check, maintain and improve a documented ISMS is as follows (cf. Figure 3.4):

1. **Defining the scope** of the solution, considering the characteristics of the business, the organization, its location, assets and technology.
2. **Creating an ISMS guideline**, which lays out the general guidelines for information security. The guidelines should reflect the economic, legal and contractual obligations of the organization and state general criteria for the assessment of risks.
3. **Defining the process** for risk assessment: The process should lead to reproducible results and include acceptance criteria for residual risks.
4. **Identifying risks** encompasses the structured elicitation of potential risks for the information security of the organization.
5. **Analyzing risks** includes the assessment of the business impact and probability of security incidents.
6. Identifying and assessing possible courses are activities for **defining risk management options**. The responsible person then has to decide on possible courses of action, which include countermeasures for risk mitigation, risk acceptance, risk prevention, or risk transfer to a third party.
7. For **selecting appropriate protection measures**, an exemplary, non-exhaustive catalog of countermeasures is provided with the standard.
8. Since security risks cannot be avoided completely, management has to **approve residual risks**. The standard requires management to explicitly approve the risks that have not been mitigated.
9. **Management approves the ISMS**, e.g. the scope, the guidelines or risk management options.
10. Finally, a **statement of applicability** is issued. It should include the selected goals and protection measures, the measures that are implemented and the ones that are not (and why they are not implemented).

Discussion

ISO2700x serves as a protection against security incidents from a company perspective, and ensures the secure processing of data, whether personal or not. However, it does not reflect, nor intend to reflect, the perspective of the individual who's data is being handled. Thus, the suitability of ISO2700x for GDPR implementation is limited to the requirements that relate to the security of processing. This deficiency is partly addressed by ISO27701 (ISO, 2019), which represents an extension to the ISO2700x family. Even though it does not cover all aspects of the GDPR, Weiß and Strauß (2019) point out that it could be a step towards certifiable compliance with the regulation.

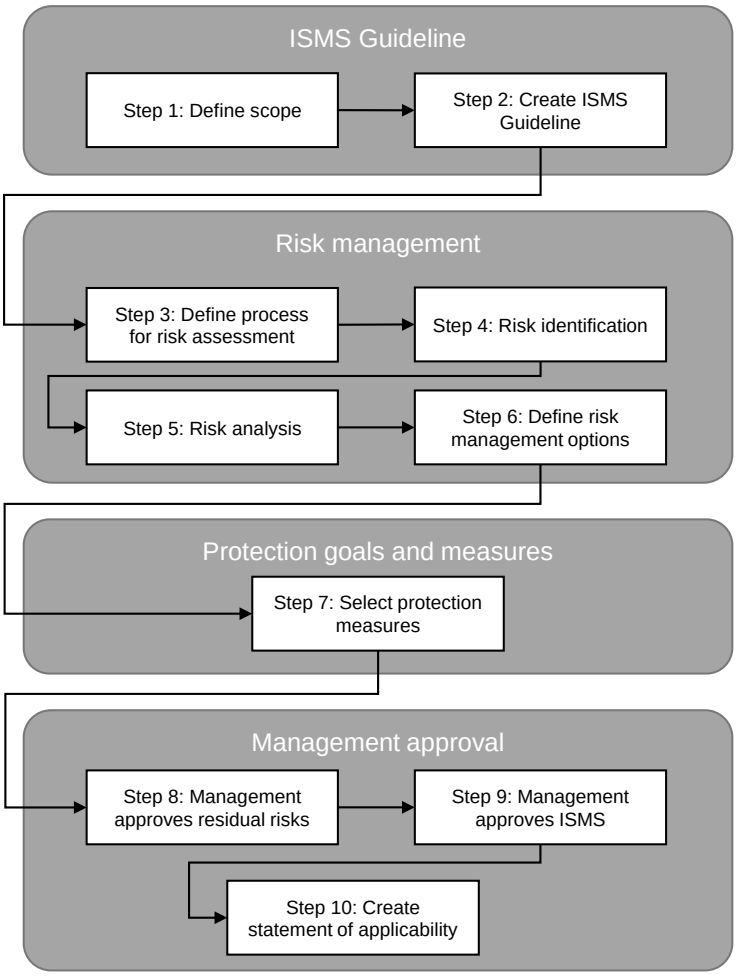

Figure 3.4.: ISO27001 implementation process (Brenner et al., 2011)

3.2.3. COBIT

COBIT (Control Objectives for Information and Related Technology) is a framework for IT Governance that is published by the Information Systems Audit and Control Association (ISACA). We describe the core concepts of the most recent version COBIT 2019 based on Gaulke (2019) and Steuperaert (2019).

COBIT was originally developed in 1996 and is designed to be independent of technology and industry. The framework structures all processes that take place in an IT-function within an organization and provides a holistic process reference model. COBIT focuses on the *what* rather than the *how* of IT governance.

The six core principles of COBIT are:

- Providing added value for stakeholders

- Holistic approach
- Dynamic governance system
- Governance separated from management
- Tailored to the needs of an organization
- End-to-end governance system

Additionally, COBIT 2019 introduced three principles that governance frameworks should adhere to:

- Based on a conceptual model: A conceptual model should represent the core concepts and their relationships.
- Open and flexible: Adding new elements should be possible and the framework should allow for flexibility when addressing new challenges.
- Aligned with important standards: The framework should integrate the ideas of important and relevant standards.

The core model (cf. Figure 3.5) comprises five domains, of which the first (*evaluate, direct, monitor*) is targeted at governance goals, while the other four pursue management goals. The process reference model is intended as a common language between the different stakeholders, especially business and IT. It does not prescribe how to implement the described processes, but rather suggests their consideration.

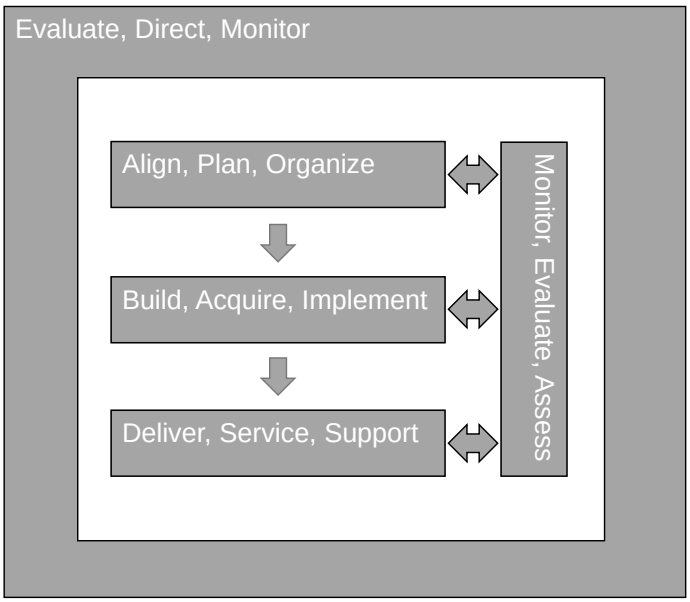

Figure 3.5.: COBIT process reference model (Gaulke, 2019)

The five domains encompass 40 management goals (see Figure 3.6), for which governance or management practices, example metrics, activities, capability levels, and further references are

defined. Further, each governance or management goal describes the organizational responsibilities for the goal with a RACI-Matrix[2]. There are inputs, outputs and work products for each practice that belongs to a governance or management goal.

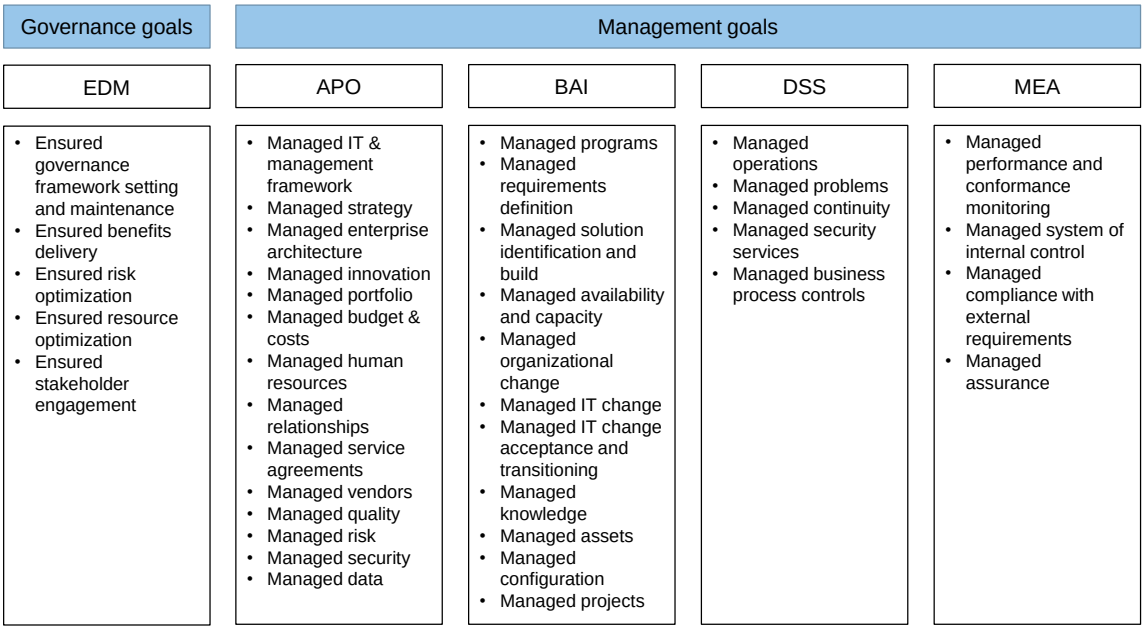

Figure 3.6.: COBIT domains with management goals (Gaulke, 2019)

Discussion

COBIT is focused on developing an encompassing governance model for IT. As such, it defines the roles and activities on this holistic level. Privacy & security criteria are not defined specifically, but are rather part of management goals of COBIT. Management goals APO13 (managed security) and APO14 (managed data) include topics that relate to data protection.

Even though COBIT is not directly applicable to the implementation of specific data protection regulation, it provides a conceptual framework that accommodates these activities. Additionally, the structure of COBIT - goal domains, goals, subgoals, activities to fulfill these goals, responsibilities, as well as their input and output - is adaptable to the domain of data protection as well.

3.2.4. IT4IT

The IT4IT standard by The Open Group defines a reference architecture for managing the business of IT, which draws from ITIL and COBIT (The Open Group, 2017, p.145). It interprets

[2]Responsible/Accountable/Consulted/Informed

the business of managing IT itself as a value chain, where each step adds value to the product *IT services*. The *IT value chain* summarizes all the key activities in IT.

The IT value chain is separated into primary activities and supporting activities (cf. Figure 3.7). Primary activities are typically the direct responsibility of the IT departments, while supporting activities are corporate activities that can be hosted in the lines of business or IT.

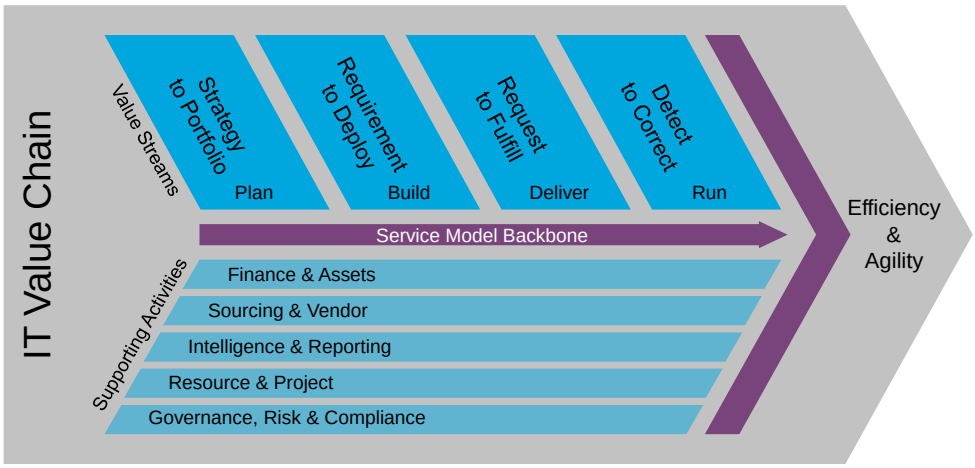

Figure 3.7.: The *IT value chain* in IT4IT (The Open Group, 2017)

The primary activities or *value streams* that are defined in IT4IT are (The Open Group, 2017, pp.8-15):

Strategy to portfolio (S2P) The strategy to portfolio value stream aims to transfer strategic demands into conceptual services. Activities include defining strategy objectives, aligning business and IT roadmaps, setting up standards and policies, service portfolio rationalization, enterprise architecture, service blueprint, demand consolidation, priority and impact analysis, and the selection based on business value, costs, benefits, resources and governance.

Requirement to deploy (R2D) In the requirement to deploy value stream, the logical service is designed from a conceptual service. This involves planning & designing the component (functional & technical requirements, logical service model, IT project plan), development, testing and deployment.

Request to fulfill (R2F) Request to fulfill leads to service catalog entries, including a price and service contract. If the service is ordered, the value stream is responsible for transitioning the service to production environments. The value stream includes setting pricing options and publishing services, managing service subscriptions in a subscription portal, managing fulfillment and deployment via internal and external providers, and measuring service usage, customer satisfaction and costs.

Detect to correct (D2C) Detect to correct integrates operational aspects like monitoring, management and remediation. In detail, the activities include detecting events, understanding the relationships between events, conducting root cause analyses on the events, analyzing

the impact and defining escalation paths, managing change requests and finally implementing the change and closing the records.

It is important to note that IT4IT considers itself as independent from process or capability models, and is therefore suitable for lean, agile and waterfall scenarios (The Open Group, 2017, p.27). The concepts of IT4IT are defined on five different levels of abstraction (cf. Figure 3.8), which each define their own conceptual information models.

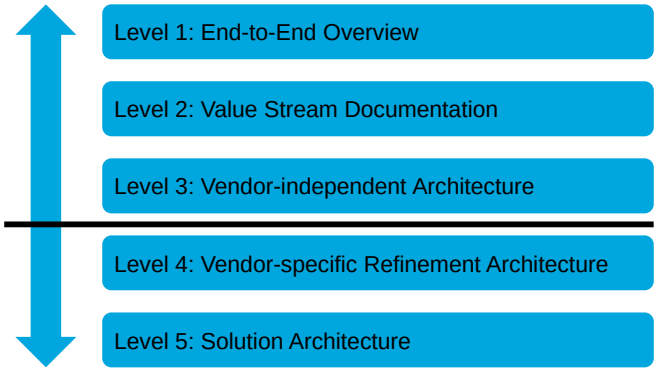

Figure 3.8.: Levels of abstraction in IT4IT (The Open Group, 2017)

Discussion

Since processing activities in the sense of the GDPR are often based on IT services, the IT4IT reference architecture is a tool that can contribute to the compliant development and operation of processing activities. The clear breakdown into the four phases of service development, as well as the separation into five different levels of abstraction, support the assessment of IT service development tasks in these categories. Further, the representation of concurrent organizational tasks, including governance, risk & compliance, visualizes the interdependency of IT service development with other processes.

We consider the R2D, R2F and D2C value streams as relevant in the context of GDPR compliant IT service development. The conceptual models that describe the relationships between elements (e.g. EA component to Service Portfolio Component) support an understanding of the overall organization, but do not relate to data protection in particular.

IT4IT reveals some further significant drawbacks when analyzing its suitability for addressing GDPR compliance tasks. The standard focuses on systems of record only, while neglecting systems of engagement or systems of insight. As IT4IT's goal is to support the provision of IT services, it only cites compliance criteria as secondary tasks without further explanation. Thus, we consider IT4IT as a valuable conceptual resource for GDPR compliance, but not as immediate process support.

3.2.5. Other approaches

Various other publications outline approaches for GDPR compliance or for establishing DPM within the organization.

Cavoukian and Dixon (2013) analyze the relationship between EA and Privacy by Design (PbD). Their approach, which rather focuses on security than privacy, consists in understanding business objectives (step 1), evaluating gaps between the current and future state of the security architecture based on a capability maturity model (steps 2-4), and defining the enabling architecture, strategic roadmap, business case and governance process (steps 5-8). The authors place particular emphasis on establishing the governance process for an EA security strategy, as the implementation of an EA security (and - assumably - privacy) strategy is a long-term process.

The practical guidebook by Voigt and von dem Bussche (2018) addresses the topic from a legal perspective. The authors propose a five-step process for implementation consisting of a gap analysis, risk analysis, project planning, implementation and realization of national peculiarities. The implementation step entails the appointment of a DPO, the establishment of a data protection management system (DPMS) and a RoPA, and the monitoring of these measures. The establishment of DPMS and RoPA is described in detail.

The challenge of compliance with the GDPR and other privacy regulation brought forth the field of privacy tech, which is characterized in IAPP (2020). Further, many consultancy companies have developed proprietary approaches within their business offering, but there is not sufficient information on these approaches to assess them.

3.3. Summary and Research Gap

Our review of related work presented core contributions to support the implementation of the GDPR from the academic community and industry frameworks. While some publications were developed before the GDPR entered into force, a large number of contributions has been released after the year 2018 and has added considerably to the body of knowledge during our research endeavor.

A number of publications point out the relationship between EA and DPM. Burmeister et al. (2019) develop a detailed EA metamodel that captures the information requirements of the GDPR. Labadie and Legner (2019) transfer the GDPR tasks into EA capabilities. Cavoukian and Dixon (2013) outline a method to combine privacy by design and EAM.

Multiple approaches detail on the stakeholders that are involved in regulatory compliance efforts. Crespo et al. (2015) specifically mention six roles that are involved in GDPR compliance projects. Koç et al. (2018) develop their method for five roles. COBIT, as a very elaborate and detailed example, comprises 33 roles for IT governance.

The ongoing and iterative nature of GDPR compliance projects is an essential component of the SDM, but also incorporated into PRIPARE and approaches by other authors. IT4IT makes the different levels of abstraction in IT related (and, for our purpose, data protection related) projects explicit.

With respect to the research goal - to develop a reference process model that can serve as a blueprint for GDPR compliance approaches - we identified the following research gap:

- Practice-based contributions are rare in the current body of knowledge. Koç et al. (2018) report from a method that was developed from scratch and validated in practice. Labadie and Legner (2019) derive the necessary capabilities from both literature and expert knowledge, but adopt a resource-based view of GDPR compliance. To the best of the author's knowledge, contributions that cover the practical implementation do not exist.
- Practitioner frameworks for IT governance and IT management describe different roles and acknowledge the distribution of responsibilities. The SDM, however, only addresses the perspective of DPM.
- The GDPR defines events, such as data breaches or data subject requests. These temporal aspects can be identified within the framework of Labadie and Legner (2019), but are not represented explicitly. PRIPARE defines an operational phase for a processing activity, in which the possibility of data subject requests is included. The more general practitioner frameworks do not incorporate GDPR provisions.

CHAPTER 4

Construction of the Reference Model Frame

The next step in the research process is defining the reference model frame. We first collect the set of requirements that shall guide the construction of ProPerData. Then, we construct the reference model frame by selecting a suitable metamodel and a suitable modeling language.

4.1. Requirements for a reference process model that supports GDPR implementation

Requirements of a reference model define the goals that the model should fulfil. They are therefore an essential and important preparation for the development of a reference process model. This section discusses the requirements that guided the development of our reference process model. They originate from the following sources:

- Empirical results on successful practices in GDPR implementation projects, as presented in Huth et al. (2020b), Burmeister et al. (2020), Huth et al. (2019b), and Huth et al. (2020c).
- Requirements proposed by work addressing regulatory requirements in the financial sector (Timm and Sandkuhl, 2018).
- The general requirements for reference models presented by Frank et al. (2007) and Vom Brocke (2003), as well as the guidelines of modeling by Schütte (1998).

We subdivide the requirements into three groups: the first group discusses and states the external modeling requirements for the reference process model in Section 4.1.1. Section 4.1.2 deals with requirements that arise from the regulatory framework, i.e. the GDPR. In Sec-

tion 4.1.3, we elaborate topics that emerged as additional benefits in interviews with different GDPR stakeholders.

4.1.1. General requirements for a reference process model

According to Frank et al. (2007), conceptual modeling aims at clear representations of systems. Clarity and understandability must be understood from the perspective of the target group, because the target group should be able to use it (Schütte, 1998, p.117). The target groups we address are the stakeholders that are involved in a GDPR implementation approach in an organization. Due to the large scope of the GDPR, the model should provide a clear and concise visualization of what the essential elements of a GDPR implementation approach are and how they relate to each other. This visual language facilitates communication about GDPR approaches. Therefore, the first requirement is stated as:

R1: A clear conceptual visualization of GDPR implementation approaches.

Processes are sequences of activities that are executed in order to attain a specified goal. In GDPR implementation approaches, the spectrum and diversity of activities make it impossible to define a rigid sequence. However, the temporal units and the dependencies between work units must be included for the process model to be able to support its goal:

R2: Ability to capture temporal units and dependencies between work units.

The reference process model should be applicable for a similar problem in a class of organizations (Schütte, 1998, p.69) and can be interpreted as the abstract description of instantiated processes in each organization. Conversely, it should be possible to transfer the reference process model to the individual context of each organization that is addressed. Especially the organizational and informational flexibility are important (Schütte, 1998, p.128). Thus, the next requirement is:

R3: Adaptability to the context of the organization as a socio-technical system.

In line with the previous requirement, individual parts of the reference model should provide value, without the immediate obligation to instantiate the full reference model. Depending on the specific organizational context of a company, not all aspects of the GDPR provisions are considered equally important and a risk-based approach is followed (Huth et al., 2020b). This means that the reference model should allow for an iterative and incremental application to address the identified issues in a consecutive manner:

R4: Incremental and iterative applicability.

4.1.2. Requirements to support compliance with the GDPR

While ***R1*** refers to a clear and understandable visualization of the model, we must also focus on the content of the model itself. All necessary elements must be included in the model at a suitable and consistent level of detail. (Schütte, 1998, p.159) remarks that completeness merely

refers to the problems of the model user. In this sense, the model can only provide completeness at the level of abstraction that it represents. For our reference process model, completeness requires coverage of the regulatory requirements. Further, the reference process model should be representative for GDPR implementation approaches from the perspective of EA:

> ***R5:*** Correspondence with regulatory requirements and representation of GDPR implementation approaches from the perspective of EA.

Further, practical guidelines and practical knowledge are crucial for efficient implementation of the GDPR provisions. The reference process model must combine the regulatory requirements with practical insights and recommendations from GDPR implementation projects, as well as the rich body of knowledge from academia and industry reports. By showing possible implementation approaches for the work units, the reference process model will foster reuse of successful practices and ultimately support the implementation of the regulatory provisions:

> ***R6:*** Provision of practical insights for implementation of single GDPR work units.

4.1.3. Requirements that originate from empirical opportunities and barriers in GDPR implementation projects

In the course of investigating the GDPR implementation approaches, various respondents referred to organizational opportunities and challenges that emerged in the efforts for regulatory compliance.

DPM experts referred to the substantial support of the DPM tasks by EAM (see Chapter 2), while enterprise architects mentioned the higher level and increased sustainability of documentation, a higher level of collaboration and the opportunity to discover optimization and consolidation potentials as particularly helpful. Therefore, the reference process model should emphasize the shared concepts and possible shared resources:

> ***R7:*** Foster reuse and value of the established artifacts and processes.

Lastly, the model should illustrate the different roles that are relevant in GDPR implementation projects. A recurring topic in the interviews was the value of collaboration and the difficulties that arise from a lack of exchange between the different stakeholders (Huth et al., 2020b, 2019b; Burmeister et al., 2020). The identification and organization of tasks is an important initial step from the technical perspective, but the major challenge in creating successful and sustainable compliance efforts is the integration into the everyday operation of an organization. Therefore, the reference model should reveal the overall complexity by clearly representing the involved stakeholders and responsibilities:

> ***R8:*** Account for different stakeholders and emphasize the value of collaboration between departments.

4.2. Reference model frame

According to (Bichler et al., 2016, p.315), *"a model comes with its background, e.g., with paradigms, assumptions, postulates, language, thought community, etc."*. This chapter defines the model frame for our reference process model. Based on Schütte (1998), we address the following steps:

- Classification of information objects in the model and description of a master or metamodel in Section 4.2.1.
- Selection of the representation and suitable modeling language in Section 4.2.2.

4.2.1. Metamodel for the reference process model

A master reference model provides a blueprint for the construction of reference models (Schütte, 1998, p.212). Such a master reference model defines elementary standard components from which reference models can be constructed: Schütte (1998) lists *information objects* that are processed, *tasks* that are defined for these objects, and the *context* in which the tasks are executed. In other words, a master reference model is a model that is used to describe the reference model itself. Such models are also called metamodels:

Definition: Metamodel
A metamodel is an information model that modelers develop or use to support the construction of application models, where the relation between metamodel and application model is characterized by the fact that the metamodel, as opposed to the application model, describes specific aspects in the construction process of the application model according to a meta principle. (Vom Brocke, 2003, p.83)

Martin and Del Alamo (2017b) describe metamodels as abstract sets of elements and relationships that describe the shared characteristics of multiple instances of their subject matter. The authors present the Metamodel for Privacy Engineering Methods (MPEM), which is based on the existing software engineering metamodel for development methods (ISO, 2014) and provides a sound master reference model for privacy engineering approaches. The MPEM assumes three layers of abstraction: The metamodel itself, methods that instantiate the metamodel, and projects, i.e. instantiations of the approach (cf. Figure 4.1).

We will describe the metamodel of our reference process model in this section. Its main elements, which are based on the MPEM (Caiza et al., 2019), are:

Definition: Role
A business role in which a stakeholder performs work units and produces work products.

This definition is slightly different form the definition by Caiza et al. (2019), who define a *producer*. A producer in the sense of the MPEM does not necessarily have to be a person, but

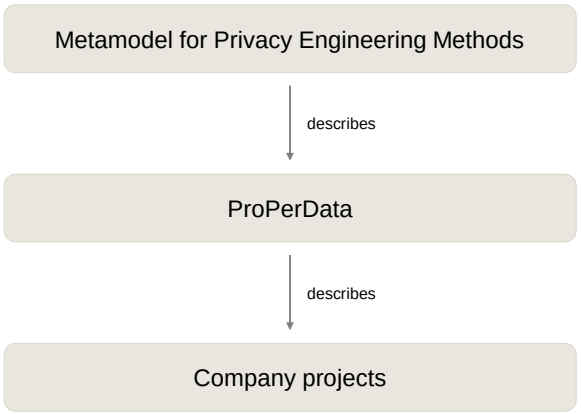

Figure 4.1.: Metamodel, model and instance levels, as used for ProPerData

could just as well be a technical system. Since we use the MPEM to describe an organizational effort of privacy engineering, rather than an effort that concentrates on a single process or system, we restrict the notion to business roles, which are assumed by people.

The role defines the *who* of a process. The *what* is defined by work units:

Definition: Work unit
A work unit represents a job performed as a part of endeavor-specific processes. (Caiza et al., 2019)

Each work unit can be broken down to more granular instructions. Work units are the central element of a process model, because they represent the elementary steps that have to be carried out. To support the work units, the actor can use resources within the organization:

Definition: Resource
An abstract representation for reusable elements used 'as is'. (Caiza et al., 2019)

Together with the instructions given in a work unit and usage of the resources, the producer creates the work product. A work product is typically a document or repository that captures the results of a work unit:

Definition: Work product
A work product represents an artifact of value that can be created, modified, used or destroyed within one or more work units. (Caiza et al., 2019)

An important aspect of process models is the specification of activity sequences. Since the DPM process is non-linear, we represent the temporal organization and dependencies of work units through stages:

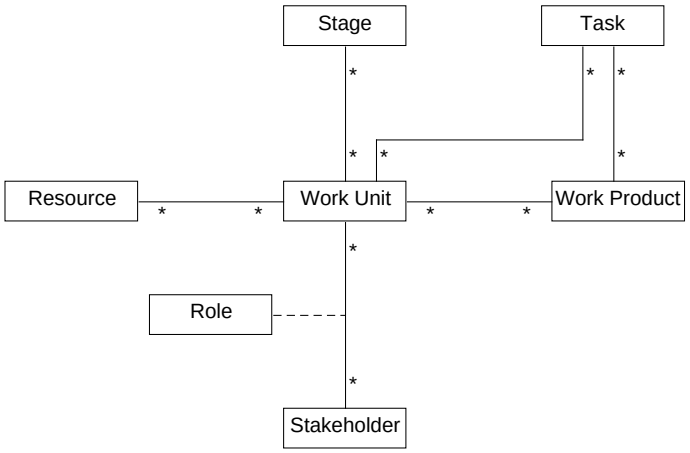

Figure 4.2.: ProPerData metamodel

Definition: Stage
A time frame during which a temporally cohesive set of work units are performed. (Caiza et al., 2019)

These five elements form the basis for the conceptual metamodel of ProPerData, which is shown in Figure 4.2. We adopted the definitions for work units, work products and stages directly from the definitions of the MPEM (Caiza et al., 2019). To allow focusing on the responsibilities for the execution, we restrict the definition of producers to stakeholder groups. Since stakeholder groups can assume different responsibilities in each work unit, we added an association class *role*.

Even though the MPEM refers to *methods*, we intentionally refrain from designating ProPerData as a method. A method includes a step by step sequence of instructions, which we consider implausible to be defined in a rigid sequence for an effort with the scope of DPM. Instead, we regard ProPerData as a demonstration of elements that allows for a company-specific selection of an implementation sequence. In our opinion, this does not change the validity of the MPEM as the underlying metamodel of our reference process model.

In addition to the work units that have to be executed by the respective stakeholders, we also consider the overarching tasks that have to be addressed to achieve overall GDPR compliance. Therefore, we introduce tasks as an additional element to be included in the metamodel:

Definition: Task
A generalized obligation for DPM within a large organization that serves as an abstract aggregation level for work units and work products.

Tasks group both work unit and work products and thus provide an additional structuring element. The overall metamodel of ProPerData is depicted in Figure 4.2.

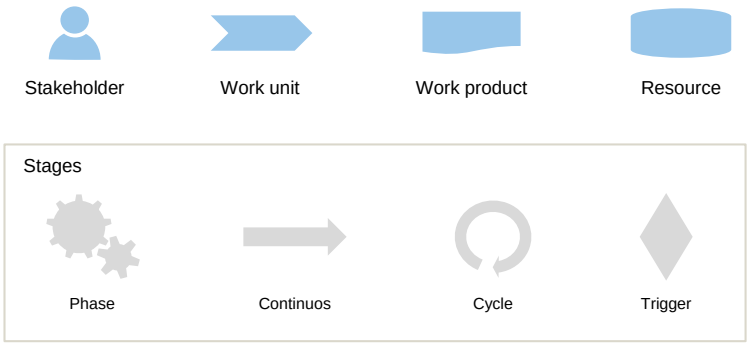

Figure 4.3.: Symbols in the reference model

4.2.2. Modeling language

Another important aspect in the construction of reference models is the selection of the appropriate modeling language and structure of the reference model (Schütte, 1998, p.220). The modeling language is crucial to convey the information of a reference model in an understandable manner, while the structure of a reference model determines its practicality.

We chose to use well known visual elements for the overview image of the reference model (see Figure 4.3). These elements support an understanding of the concepts without the complex semantics that would be necessary to capture the variety in the work units.

Further, we decided to create a single overview page (cf. 5.4) that captures parts of the relational knowledge that is inherent in the reference process model. Such an overview provides a clearly defined entry point to a process model and supports the selection of the right work units. As recommended by Schütte (1998), the work units are arranged in a matrix structure that supports the identification of the right work units.

Each work unit addresses a different set of stakeholders and thus requires flexibility in the language that describes the process steps that have to be taken. To account for this variety in the spectrum of expression, we opted to use the written language as the most basic and versatile modeling language. Where appropriate, we relied on commonly used modeling approaches, such as conceptual diagrams, to support single work unit descriptions.

4.2.3. Summary

This chapter presented the requirements that form the foundation for the development of ProPerData. The metamodel of ProPerData, which is based on an established privacy engineering metamodel, comprises the five main elements *roles*, *work units*, *resources*, *stages* and *work products*. *roles* are assumed by stakeholders, and *tasks* structure *work units* and *work products*. The modeling approach is based on semi-formal notations to support understanding for all stakeholders.

CHAPTER 5

ProPerData - a reference process model for GDPR compliance management based on EA

This chapter describes the construction of the main artifact *ProPerData*[1] within the research approach presented in Chapter 1. Our approach is based on Schütte (1998) and Ahlemann and Gastl (2007). We already *defined the problem* (Chapters 2 and 3) and *constructed the model frame* (Chapter 4). This chapter explains the *construction of the reference model*.

To underline the scientific adequacy of a reference model, the author must explain the underlying abstractions and design decisions that led to the final artifact (Frank et al., 2007). This section elaborates on the approach for the development of our reference process model *ProPerData* based on the reference modeling approach by Schütte (1998).

5.1. Construction approach

We interpret company-specific process models as the configured instances of a reference process model, which is yet to be discovered from empirical evidence. As shown in the exemplary visualization in Figure 5.1, the company-specific instances represent partial instantiations of the reference process model, whereby the characteristics for the instance configuration are determined by the underlying characteristics of the organization, e.g. the company size, the organizational structure, the processing activities or the IT landscape. The consolidated union of the company-specific instances (i.e., the maximum set of observed elements that qualify as good practice, as per self-assessment of the interviewees) constitutes the underlying reference process model.

The number of elements and the number of possible combinations ultimately determine the

[1] Process model for the **Pro**tection of **Per**sonal **Data**

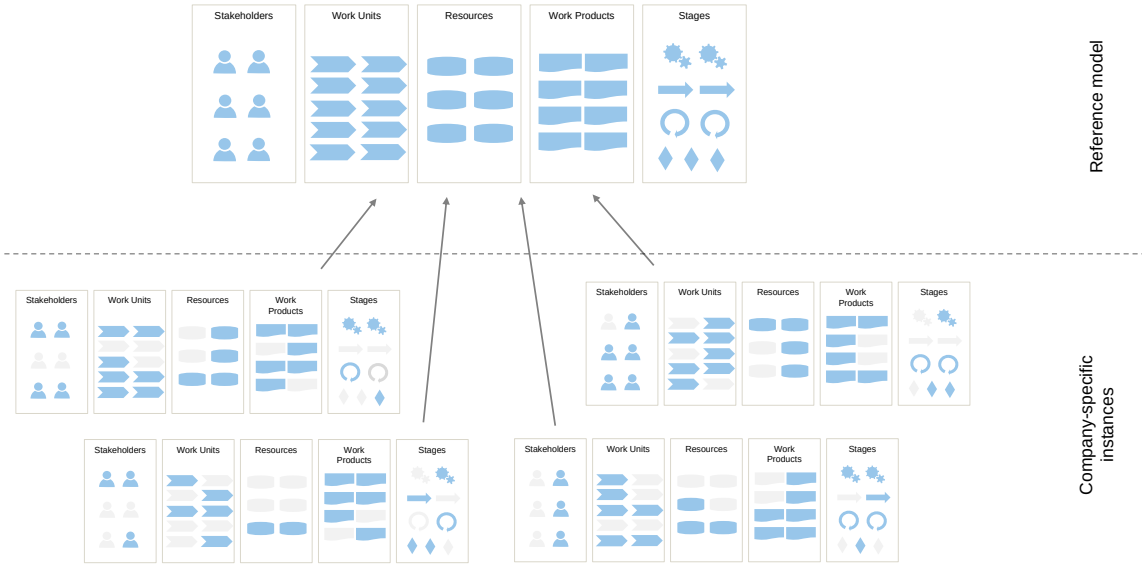

Figure 5.1.: Elements in single company instances as partial sets of the elements in the reference process model

complexity of the process model. Due to the large number of requirements and the large number of affected processes, a reference process model for the implementation of the GDPR is inherently complex.

To avoid the complexity that is introduced by modeling sequences, Schütte (1998) recommends omitting sequences completely and offering only the basic elements. Specifying the sequence is then the responsibility of the model user (Schütte, 1998, p.255). This is particularly true for environments where the underlying structure of the target organizations (i.e., the business processes) change frequently. Such a reduction to basic modeling elements is called a *compositional approach* (Schütte, 1998, p.257). First, we iteratively identified the initial elements of ProPerData, based on the interview results from Huth et al. (2020b), which we will detail further in Section 5.3. After the initial reference process model from the interview results, we added missing elements to create the final set of ProPerData elements. Finally, after the evaluation discussion, we incorporated useful suggestions, e.g. to harmonize the granularity of the work units.

The **selection of a suitable level of abstraction** plays an essential role: It must be able to display generality of the concepts, but take into account the different perspectives of the stakeholders. A basic requirement is that the different levels of abstraction must remain understandable for the respective model users. If the variety of organizations in the target group is too extensive and would make the model too heterogeneous, the modeler should not further refine the models and instead focus on the commonalities among the target group. The level of detail largely depends on the application scenario of the model. (Schütte, 1998, p.236) refers

Figure 5.2.: Construction approach for ProPerData

to survey results that criticized overly detailed reference models. Yet, a lack of detail makes it difficult to actually instantiate the model. Thus, we aim for a compromise between the abstract representation and the implementation details.

Next, structural analogies should be identified to fulfil the basic requirement for generality of the model. Structural analogy does not necessarily mean structural equivalence. It is rather necessary to fulfill two requirements: firstly, matching of the majority of information objects, and secondly, alternative implementations of the same concepts. The focus on structural analogies allows to investigate the subject matter independently of the particular economic context in which the project is conducted, again contributing to the claim of generality of reference models.

Process models represent an isolated instance of a process execution. Thereby, interrelations with other processes and resources are ignored (Schütte, 1998, p.240). In our case, we consider the importance of interrelations as essential: Establishing regulatory compliance is an overarching meta process within an organization that touches a great number of internal processes, and hence we cannot neglect these relationships. We will cover these aspects in the following sections.

The construction approach of ProPerData is visualized in Figure 5.2. The *identification of basic elements* is discussed in Chapter 4. This section will detail on the steps *data collection* in Section 5.2 and *construction* in Section 5.3. The final step *iterative refinements* is discussed in Section 6.2.

5.2. Data collection

The empirical evidence that we collected constitutes the basis of our reference model. Since we defined the target group of our reference model as medium to large organizations that employ enterprise architecture models, we consider each individual GDPR implementation approach as an instance of our reference model. Therefore, we interpret each project description by an EAM expert as a (possibly incomplete) set of elements from our reference model. Figure 5.1 shows this perception of the set of elements (according to the metamodel) in each project instance as an incomplete set of elements of the overall process model.

5.2.1. Interview series with enterprise architects

As recommended by Ahlemann and Gastl (2007), we used interviews as empirical inquiries to gather practical knowledge about GDPR implementation projects. We approached enterprise architects because of their holistic perspective on the organization as a socio-technical system of

people, processes, applications and technology. The interview partners from 24 interviews (cf. Table A.1) include a wide range of industries and company sizes from 300 employees to more than 50.000 in the German speaking area. All interviews were held in German.

The exploratory interview guideline was discussed iteratively among four researchers and covered the following topics:

Enterprise architecture and collaboration aspects in the organization

- For how many years has your organization engaged in EAM?
- How many people are involved in EAM in your organization?
- How would you characterize the role of EAM in your organization?
- How do you obtain the information for your models?
- Who uses your information?
- How did you / do you collaborate with other departments, in particular during the implementation of the GDPR?

EA models

- Which elements of your enterprise architecture do you capture in your model?
- How do you model personal data and the processing of personal data?
- Do you model data protection aspects, such as security measures or anonymization?
- Which advantages did you observe because of transparency about data processing?
- Did you also observe any disadvantages?

Tooling

- How do your EA tools support data protection aspects?
- Do you use any dedicated tools to achieve GDPR compliance?
- For which DPM tasks did your EA tools support GDPR compliance? E.g. RoPA, DPIA, data subject requests, notification processes for data breaches, etc.

EA frameworks

- Do you follow an EA framework?
- Do these frameworks address management of personal data in particular?

Open discussion

- Please discuss the statement 'EA is a key to compliance'.
- Which role does data exchange between data processors play in your view?
- What is your personal opinion of the GDPR?

Each interview was conducted by one researcher between March and May 2019. All of the interviews were recorded and lasted between 36 and 72 minutes.

5.2.2. Other empirical inquiries

During our research endeavor, we discussed various proposals that contributed valuable practical insights into GDPR implementation projects and their challenges. Some of these inquiries happened in the context of publications, such as eight DPM experts (Huth et al., 2019b) or seven IT experts from the data security and privacy field (Huth et al., 2019a) (cf. Table A.3 and A.2). Other relevant information was extracted from notes that were taken in conversations with lawyers (two discussions), software developers (two discussions), one cybersecurity expert and the discussions with fellow researchers at direct meetings and at scientific conferences.

5.2.3. Academic publications and industry guidelines

The body of knowledge provides extensive descriptions of single activities, such as the engineering of privacy aware systems. We included these references in the description of the respective work units in the Appendix. Further, we adopted notions of the approaches that we presented in Chapter 3. We discuss these notions in Section 5.4.2.

5.3. Construction of *ProPerData*

The actual construction step of the construction process started with transcribing the interview audio into text. As suggested by Mayring (2000) and Saldaña (2013), we marked *GDPR tasks*, *benefits*, *EA models*, *collaboration aspects* and *barriers* in three collaborative coding cycles with two researchers, using the qualitative analysis tool MAXQDA. The first coding cycle assigned initial codes, which were further refined in the second cycle and then consolidated in the third. Overall, we assigned 51 different codes to more than 1600 text segments.

Then, the model elements were extracted from the coded segments and developed in two iterations. The preliminary set of model elements was arranged into an initial overview canvas of ProPerData. The initial version of the ProPerData technical report included the overview canvas, as well as detailed descriptions of the ProPerData elements. The construction step is visualized in Figure 5.3 and explained in detail in this section.

We selected the ProPerData elements based on the work units, i.e. the work units represent the central elements of ProPerData. We applied the following process in the selection of ProPerData work unit candidates:

1. First, we analyzed the coded text segments from 24 interviews with enterprise architects for occurrences of work unit candidates. The work unit candidates were documented according to a pattern documentation template, which we adapted from Aleatrati Khosroshahi et al. (2015), and the MPEM structure that we introduce in Chapter 4. Thereby, the inclusion criteria for ProPerData work units were:

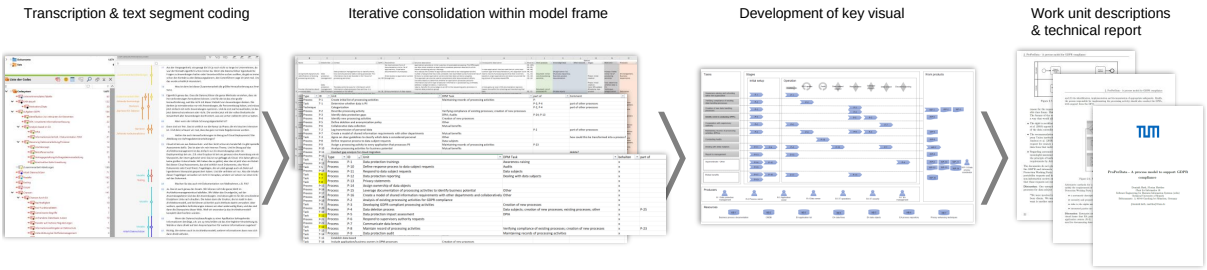

Figure 5.3.: Construction step of the initial version of ProPerData

- Self-assessment as *good practice* by the interview partner or positive context in which the practice was mentioned.
- Generalizability of the element and relevance for the target group, as assessed by the researcher.

The first iteration resulted in 48 documented work unit candidates. For each work unit candidate, we documented a name, the involved stakeholders, the context of the work unit candidate, the relevant GDPR reference, possible preconditions for the organization, and the description of the solution and its consequences. Each of the work unit candidates occurred between once and fourteen times, where the occurrences were non-exclusive. Additionally, we examined the knowledge bases that were necessary for the candidate, as well as the producers and the resulting work products. Each work unit was assigned to at least one stage in which it occurs. Table 5.1 shows one exemplary documentation of a work unit candidate.

2. In the next step, we examined the prospective work units regarding their relationships with other work unit candidates and the categorization according to the DPM tasks. The goal in this step was to reduce and consolidate the list of work unit candidates, since the practices that we identified in the interviews often shared significant traits and only differed in aspects, such as the CMDB or the EA application repository as the data source for applications that might process personal data. Inclusion criteria, which resulted in 19 work unit candidates were:

   - Possibility to combine separate work unit candidates into one, due to similarity in the most important traits. For example, the work unit candidates *create RoPA from CMDB* and *create RoPA from EA application list* were merged.
   - Matching granularity among the work unit candidates: as the example in Table 5.2 illustrates, *determine the worst case scenario in data breach* is included in the work unit candidate *conduct DPIA*.
   - Direct relationship to one of the DPM tasks in Chapter 2, e.g. *creation of processing activities*. If we could not identify such a relationship, we reconsidered the list of tasks that we developed in Chapter 2. Ultimately, this resulted in two additional tasks *execute organizational tasks* and *leverage data protection efforts for business impact*.

Pattern template	ID	3
	Name	Use IT operations data for identification and characterization of processing activities
	Stakeholder	Data protection group, IT operations
	Context	Data protection management has to identify where, how and why personal data is being processed. This information has to be recorded in the "record of processing activities". However, it is difficult to identify possible processes and the responsible contact partners.
	GDPR concern	Art. 30
	Precondition	IT operations team with a CMDB (configuration management database)
	Solution	Collect information about running applications from the IT operations team (virtual machines, applications, instances, business owner, technical owner). If information about the application owners exists, contact the application owners for more detailed information, e.g. the reasons for processing and the type of processed data.
	Consequence	Processing activities are not always conducted via dedicated applications. Also, the view from single applications might be too granular to reflect overarching processing activities that rely on multiple applications
	Observed in	I02
MPEM elements	Work products	Document: Initial List of processing activities
	Knowledge base	CMDB
	Stages Work Unit Producers	Phase: initial setup Create initial list of processing activities IT operations

Table 5.1.: Exemplary initial documentation of practices, following a pattern documentation template and the MPEM structure

3. Following the consolidation of the work unit candidates, we extracted and consolidated the other elements within our model frame.

   - Eight producers remained as stakeholders in the identified work units: EA management, process owner, data owner, application owner, IT operations, data protection management, software developers and IT security.
   - Similar to the work units, some work products included other work products, such that we consolidated the work products as well. The list then comprised 16 work products, which were reduced to 13 after further discussion and consideration.
   - Twelve resources were identified initially, which were later consolidated to seven resources in ProPerData.
   - Two phases (initial setup and operation), two cycles (RoPA cycle and audit cycle) and three trigger events (new process, changed process and data breach) form the

Candidate ID	P-14
Description	Determine worst case scenarios in data breach
DPM tasks	DPIA; interact with data subjects
Part of candidate	P-13 (conduct DPIA)
Keep	no

Table 5.2.: Example of ProPerData work unit candidate in work unit consolidation

initial stages of ProPerData. After the qualitative interviews, the review phase, the continuous stage, and the decommissioning trigger event were added.

4. The initial elements were arranged in a canvas to support visualization of the relationships between the elements. The canvas was developed in multiple iterations and discussions, using the visual elements that we presented in Section 4.2.2.

   Figure 5.4 depicts the ProPerData overview canvas. On the upper left side, the eleven DPM tasks provide the basic structure of the canvas. They span from left to right to illustrate the relationship between tasks, work units and work products.

   At the center, the three phases *initial setup*, *operation* and *review* structure the time dependencies. The operations phase is subdivided further into continuous, cyclical and event-triggered activities. Stages and DPM tasks form a matrix-like frame in which the work units are located, indicating both the temporal relationship and the relationship to the DPM tasks.

   Work products are separated into internal and external work products. External work products are official documents that are required by the regulation.

   Roles and resources are placed below the main elements, but do not bear relationship information in the visual structure.

5. To complete the initial version of ProPerData, we formulated the work unit descriptions, as well as descriptions of the other identified elements. The textual descriptions follow the basic structure:

   - Statement of the **rationale**, which describes the underlying considerations within the GDPR and the GDPR recitals.
   - The **description** of the basic process, supported by additional conceptual diagrams.
   - A **discussion** of the implications and difficulties, if such information was available.

   This initial version of ProPerData was published as a technical report (Huth and Matthes, 2020) and served as the basis for the qualitative evaluation interviews.

6. Finally, after the qualitative evaluation interviews that we present in Chapter 5, we incorporated the feedback of 14 experts into ProPerData and updated the technical report. The updated technical report served as the basis for the expert survey to assess the validity of a reference process model to support GDPR compliance.

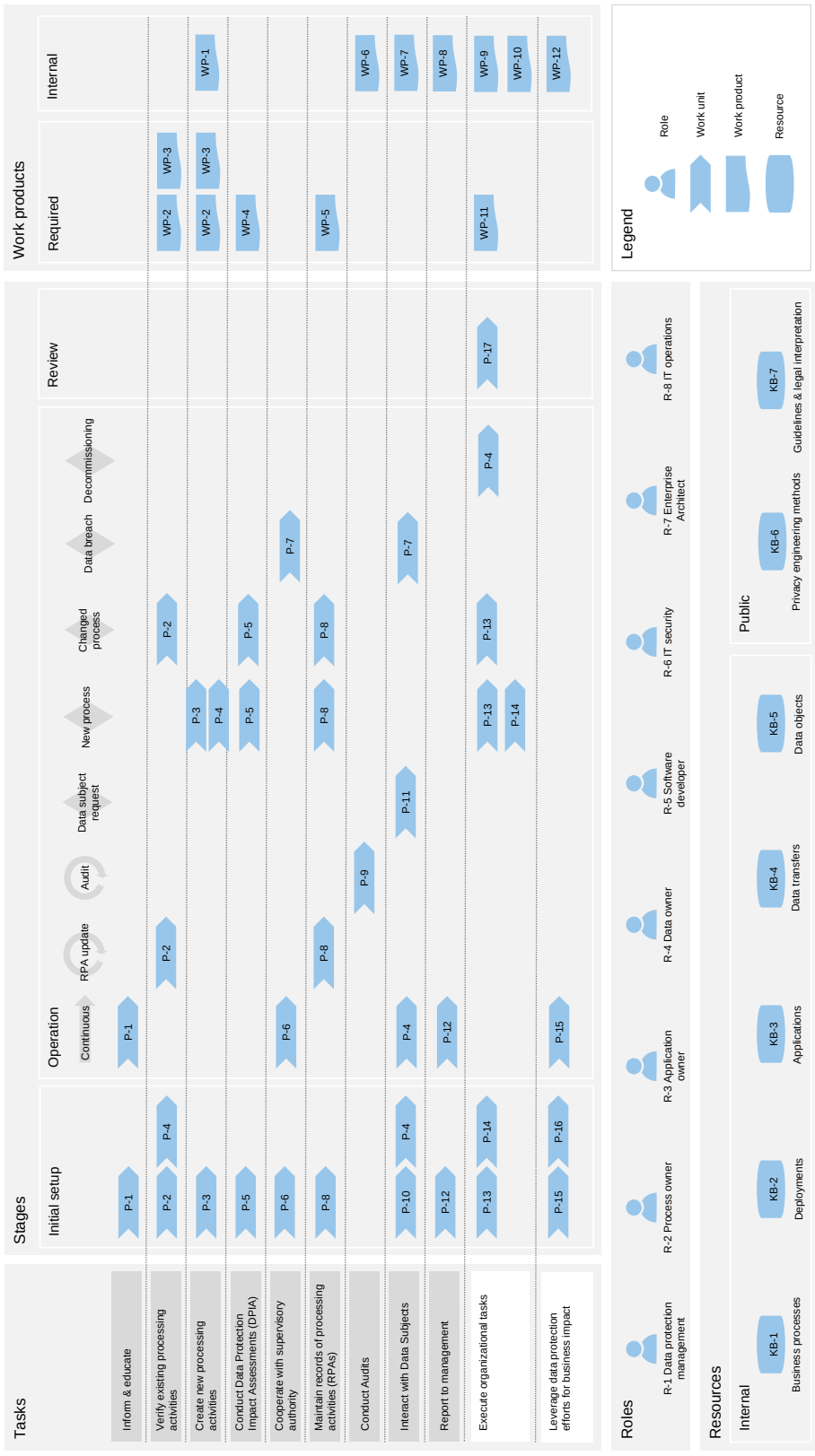

Figure 5.4.: The ProPerData overview canvas

ID	Work unit	Data protection management	Process owner	Application owner	Data owner	Software Developer	IT security	EA management	IT operations
P-1	Data protection trainings	R	I	I	I	I	I	I	I
P-2	Analysis of existing processing activities for GDPR compliance	C	A	C	C	C	C	C	C
P-3	Developing GDPR-compliant processing activities	C	A	R	C	R	C	C	C
P-4	Data deletion process	C	R	R	A				
P-5	Data protection impact assessment	C	A, R	C		C			
P-6	Respond to supervisory authority requests	A, R	R	C	C		C		C
P-7	Communicate data breach	R	R	A, R	R	C	C		C
P-8	Maintain record of processing activities	A, R	R	C	C	C	C	R	C
P-9	Data protection audit	A, R	C	C	C	C	C	C	C
P-10	Respond to data subject requests	R	A	C	C			C	
P-11	Data protection reporting	A, R	C		C			R	
P-12	Update privacy statements	A, R	C		C			C	
P-13	Harmonize processing activities for data objects	C	R	I	A	I		I	
P-14	Reflect and adapt GDPR implementation practices	A, R	R	R	R	R	C	R	I
P-15	Leverage documentation of processing activities to identify business potential		R	C	C			A, R	
P-16	Align information requirements and collection processes with other departments	R					R	A, R	R

Figure 5.5.: ProPerData work units with responsibilities (RACI-Matrix)

5.3.1. Interrelations of model elements

ProPerData is a reference process model to support achieving and maintaining GDPR compliance from an organizational perspective. It is intended to structure compliance projects, identify work units, stakeholders and resources, and check for completeness in the attained results. We present the overview canvas of ProPerData in Figure 5.4 and describe the respective elements in detail in this chapter. The overview is presented as a canvas that visualizes the dependency between data protection management (DPM) task categories, the respective work units and their temporal relationships, as well as the outcome of these work units.

The ProPerData canvas shows the stakeholder roles and resources of the model, but does not incorporate relationships that involve these two groups. These relationships can be identified from Figure 5.5 or from the work unit descriptions.

In a similar fashion, the ID numbers of the work units do not carry any information other than the approximate position within the canvas and are only intended as support in navigation.

Figure 5.5 defines the responsibilities for the respective ProPerData roles. Figure 5.6 shows the dependencies among the work units, i.e. which work units are prerequisites for others.

5.3.2. Roles and collaboration

In this section, we elaborate on the collaboration between the different ProPerData roles with a focus on the EA perspective. The results are based on the interview series that we published in Huth et al. (2020b) and Burmeister et al. (2020).

Most interview partners described an organizational dependency with the **DPM expert** teams. A12 reported a clear separation between the responsibilities of EAM and DPM:

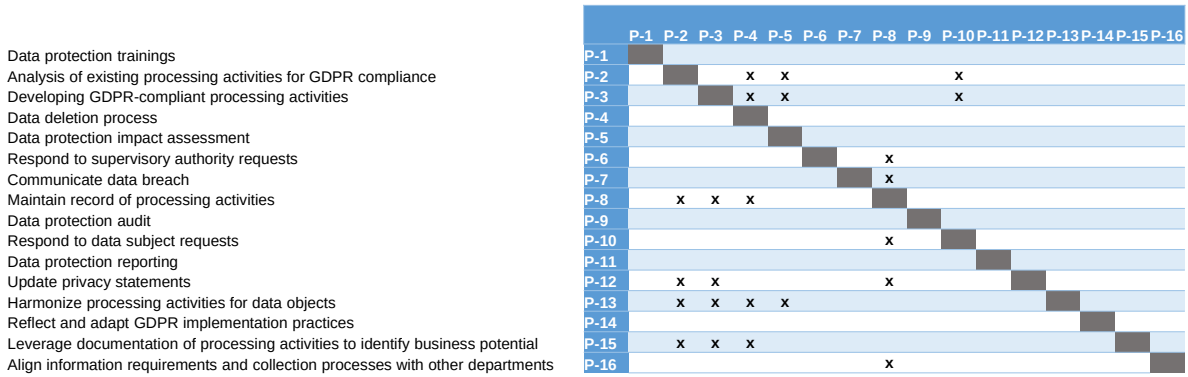

Figure 5.6.: Dependencies among ProPerData work units

> *"But collecting this (data protection related) information, that is nothing we do in EA, but the DPM team. They do that. And they have to do that."*

There was a general awareness of the need for collaboration between multiple departments, as expressed by A18:

> *"We have to bring together various instances. The DPO is one of them, then information security. Another one is the identity and access management group. And strictly speaking, we all have to collaborate to make it work."*

However, this intention often proved to be difficult in practice. Multiple interviewees reported a lack of exchange, e.g. A01:

> *"We are in contact with the data protection team, but that is expandable. [...] In principle, DPM is in our department as well."*

A very positive and enthusiastic account of a fruitful collaboration was given by interviewee A07:

> *"[We get a lot of information] through the integration of 'befriended' departments or responsibilities, like DPM. They work directly with the architecture management tool. And when they get information about new applications that we haven't recorded yet, they enter them into the tool."*

The **application owners** supply large parts of the information that is necessary for the data protection documentation, such as whether or not an application processes personal data or details on the technical and organizational measures for protecting the personal data from unauthorized processing. A01 identified an unclear responsibility for delivering this information:

> *"The solution owners themselves - those that are responsible for the applications - are typically not aware. Currently, this is driven by DPM. That's is not right in my view. It should be driven by the application owners."*

A common theme and good practice was that application owners were identified through EA

documentation and then approached directly with a questionnaire, as interview partner A04 described:

> *"We classified the applications and asked ourselves, where personal data is even processed. The ones that were relevant received a questionnaire from me that asked all questions that concerned the GDPR, roughly 40 questions. I got this questionnaire from the information security and compliance people."*

Ultimately, the enterprise architects stated the goal to involve the application owners in the data collection process. According to A08, the application owners already contributed large parts of this effort. However, as A12 pointed out, this approach still works less than optimal.

Product owners are typically employees of the business units that are responsible for business processes. As A14 explained, there is an inherent motivation to become active:

> *"Let's put it this way: there is a large intrinsic motivation. [...] The business units are very active based on this intrinsic motivation. It is in their own interest to have [GDPR measures] implemented. We as architects don't have to act as motivators. But we do have to point out dependencies and support managing them."*

Since personal data is processed for a business reason, the product owners are the responsible group for specifying the purposes of processing activities. Nonetheless, the distribution of responsibilities is not always clear. Especially for the RoPA, multiple actors have to collaborate in compiling the information. A22 reported:

> *"The business units specified which applications are used for which purposes, and DPM approved that. [...] Within IT, process development is not our focus. It is a purely business-driven topic. [...] Sometimes it should really stay like that."*

To ensure consideration of data protection topics within the business departments, A16 reported the establishment of data protection contact points within each department, who must ensure that audits are executed, that processing activities are kept up to date, etc.

Another side role is the role of **data owner**. According to A18, the data owner is assigned to data objects, such as address data of data subjects or transactional data. A17 describes the role as follows:

> *"It is the responsibility of the respective data owner to ensure compliance with the data protection requirements. The data owner is part of the respective business departments. And, as far as I know, the DPO discussed and classified all data objects with the data owners in each business unit."*

As we sourced our information mainly from enterprise architects, it is important to highlight this one-sided representation. In some instances, the **IT operations** group supported the identification of relevant applications. Enterprise architects, e.g. A08, pointed at the problems that may arise through one-time exports from repositories such as the CMDB:

> *"What they did was to take an export from the CMDB and collect the information in a separate tool, because they didn't know better. They probably didn't know our*

> *EA tool, it's really a lot of work to bring in all the people that collect architectural information. [...] Of course they could have done the same thing in our EA tool, then they could use more recent lists instead of a one time export from the CMDB."*

The **IT security** team is involved in the definition of technical and organizational measures. Interviewee A03 described an informal exchange between EA and IT security, where the colleague from IT security reports current topics. Interviewee A18 described the strong coupling of IT security and DPM as a result of the German provisions for critical infrastructures[2], to which interviewee A15 also related:

> *"IT security implements [these requirements], and one of these aspects is deletion of personal data. [...] We have a model of the technical systems that is used for the purposes of IT security."*

A07 reported a successful collaboration between IT security, DPM and EA:

> *"The interesting thing about our setup is that we as enterprise architects are assigned to the IT security department, which closely collaborates with DPM. This goes very well together."*

Another stakeholder in GDPR implementation approaches is the **software development** team. As we assume the perspective of enterprise architects, the discussion on software development is mostly focused on solution architecture. Ths solution architecture represents the link to the actual implementation. A10 referred to a bilateral exchange of information, where the involvement of solution architects in the development work ensures that architecture recommendations are taken into account. A18 talked about checkpoints for developers to determine the relevance of data protection in a specific development project:

> *"We have a process that is followed in any project initialization. This process has a checkpoint to assess the relevance for data protection. This assessment is forwarded to DPM for further examination. DPM, security management and EA are involved in this process."*

A DPM expert that we interviewed in the context of (Huth et al., 2020c) described this as *guard rails* and pointed out the importance of considering them early on in the development process, as this front-loading can significantly decrease the amount of support and clarification that DPM has to provide when the development has progressed further.

5.4. Design decisions

5.4.1. Level of abstraction

The level of abstraction is an important consideration when designing an information model. As our goal is two-fold - (1) to facilitate communication and planning and (2) to provide support

[2]KRITIS, `https://www.kritis.bund.de/`

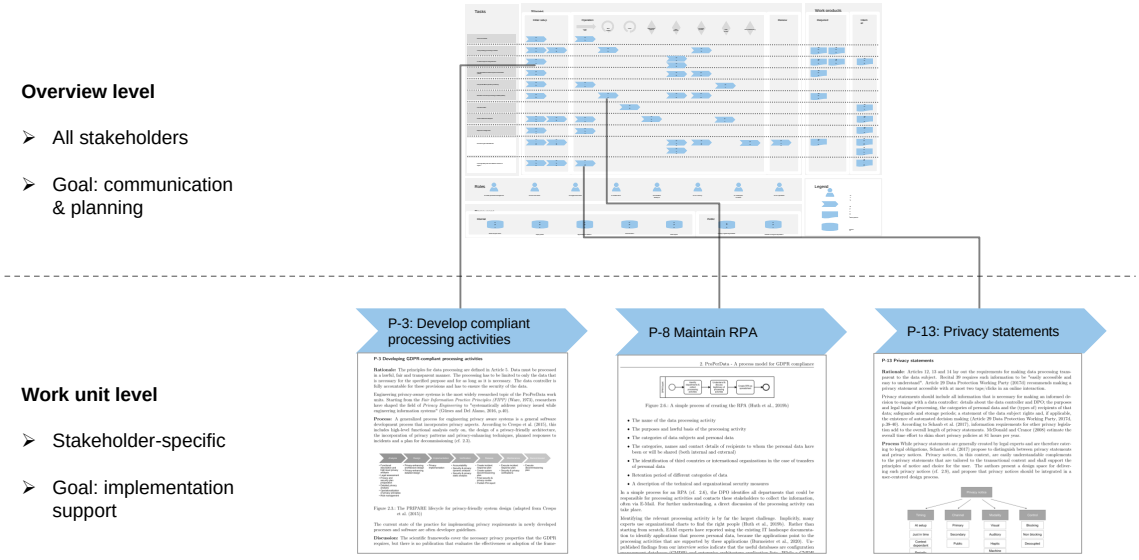

Figure 5.7.: The two abstraction levels of ProPerData

for performing the work units - we have to provide two levels of abstraction. Figure 5.7 shows a schematic representation of the two levels of abstraction.

The first level is the **project overview** level. It addresses all stakeholders within the model, since its goal is to support communication about GDPR implementation projects and identify responsibilities among these stakeholders. By using the overview canvas as a map, the stakeholders can identify necessary work units and their interrelationships based on the time units (stages) or tasks.

At the **work unit** level, the single work units each address target stakeholders. Here, the goal is to support the execution of the identified work unit. Each work unit description is structured in a similar way. First, we include key considerations with respect to the work unit, e.g. the privacy objective that the work unit target and the respective text passage in the GDPR. Next, we summarize our insights from publications in academia and industry, as well as the official guideline documents by the A29WP to describe the elementary steps and considerations when executing the work unit. Lastly, we added practical insights that we obtained during our research endeavor in the discussion. These insights are intended to illustrate possible difficulties or opportunities when engaging in the work unit.

The choice of stakeholder-targeted work units implies that portions of the technical report are addressed at these stakeholder groups only. As an example, conceptual Unified Modeling Language (UML) diagrams are very familiar to software developers, but possibly less expressive for some data protection experts or business users.

5.4.2. Relation to existing approaches

The approaches we presented in Chapter 3 propose concepts that we adopted in the construction of ProPerData as well.

- Most approaches incorporate categorizations as structuring elements: The SDM uses overarching protection goals as the structuring elements for single requirements, COBIT uses five overarching governance and management goals for the 40 sub-goals, and Labadie and Legner use six categories to structure the necessary capabilities for GDPR compliant data protection. We adopt this proven concept with the tasks that we identified in Section 2.3. The tasks provide a structure for the work units and work products.
- Different levels of abstraction are used in PRIPARE, where a general organizational process complements the (dominant) system lifecycle process. IT4IT defines five different levels of abstraction, which range from the end-to-end perspective to specific solution architectures. We understand the two ProPerData levels of abstraction - the canvas overview level and the stakeholder-specific level - similar to the levels in PRIPARE, with a clearer separation between the single work units at the implementation level.
- PRIPARE encompasses six roles, two of which are subdivided further. COBIT describes 33 roles, which include decision boards and top management roles. In ProPerData, we include the roles that surfaced in our empirical inquiries. However, in some cases the roles we define summarize various roles of COBIT: as an example, we summarize the legal function that is concerned with data protection topics with other, company wide data protection roles, whereas they represent two separate roles in COBIT.
- Among the approaches we presented in Chapter 3, we only located the concept of a RACI matrix in COBIT. We chose the RACI matrix as a concise way to represent responsibilities in ProPerData as well.
- Especially the SDM stresses the necessity to understand DPM as an ongoing, iterative endeavor. This property is reflected in ProPerData through the phase *review* in the ProPerData overview.
- The terminology of EA, as described by e.g. Koç et al. (2018), proved to be very helpful in describing the holistic nature of DPM.
- Another concept that influenced the development of ProPerData are patterns. Patterns have been used in the privacy engineering field for more than a decade (Kalloniatis et al., 2008; Hoepman, 2014) and are widely accepted in EAM (Aleatrati Khosroshahi et al., 2015). In Huth (2017), we discuss the applicability of patterns to address GDPR requirements, an idea that was adopted by other researchers (Rösch et al., 2019).

5.4.3. Design principles to enhance information governance

Burmeister et al. (2020) derived seven design principles for EAM to enhance information governance, based on benefits and barriers that were identified from GDPR implementation projects (cf. Table 5.3).

Type	No.	Design principle
Structural	**DP1**	Identify the decision-makers within information governance to prioritize the consumers of EAM
	DP2	Define roles and responsibilities in each department that collaborate with EAM on managing the information artifact
Procedural	**DP3**	Foster strategy development regarding information usage by providing valuable insights into architectural relations and potential synergies
	DP4	Proactively advise all business and IT departments in realizing effective information governance
Relational	**DP5**	Ensure a shared terminology and unified definitions of the EA in the context of information governance
	DP6	Create and use a lean and intelligible EA meta-model that covers information artifacts, data flows and data processing
	DP7	Initiate a routine for information exchange and the use of a shared EA repository for information governance

Table 5.3.: Design principles for EAM to enhance information governance (partial table from (Burmeister et al., 2020))

- The two *structural design principles* relate to the roles and responsibilities of the stakeholders. This enforces the accountability principle of the GDPR and ensures that business and data protection are balanced adequately.
- Two *procedural design principles* make sure that clarity about processing activities is established. This helps in minimizing data collection and demonstrating key interrelationships between data objects.
- Three *relational design principles* contribute to a shared terminology and understanding for IT and business artifacts. This shared understanding facilitates the communication about processing activities from the IT, business, and legal perspective.

5.5. Interdependencies with internal processes

The data protection and GDPR compliance process, which we cover in this thesis, is highly interrelated with other processes. As illustrated by the exemplary process map in Figure 5.8, we understand the data protection and compliance process as a supporting process, which affects the core business process, management functions and other supporting processes. In the following, we will discuss these interrelations.

5.5.1. Interdependencies with EAM

The EAM process is aimed at developing the EA towards a to-be state that is aligned with the business and IT strategy of the organization. As described by Hauder et al. (2014), an

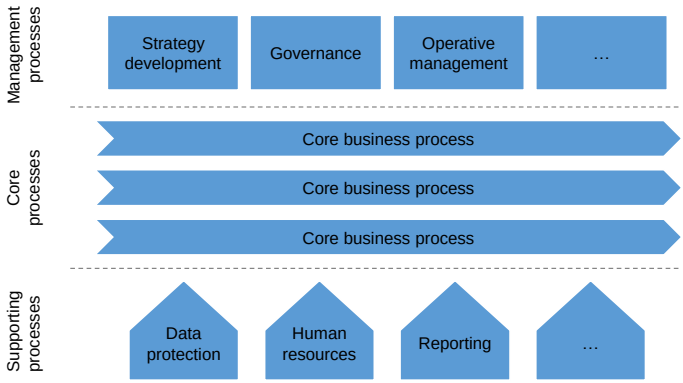

Figure 5.8.: The data protection process in an exemplary process map (representation based on (Gadatsch, 2017, p.85))

agile EAM process consists of the steps (1) information collection and modeling, (2) interaction with the respective stakeholders via metrics, visualizations and reports, and (3) reflecting and adapting the approach. Thereby, EAM addresses concerns of different stakeholders, such as top management or, as we observed in the interviews, DPM.

EA models are particularly suited to start DPM endeavors. As Figure 5.9 shows, EA tools (in this case ArchiMate) provide the modeling concepts to represent all data that is relevant for data protection documentation. We show how to query such a model in order to generate a RoPA from an EA model in Huth et al. (2019b).

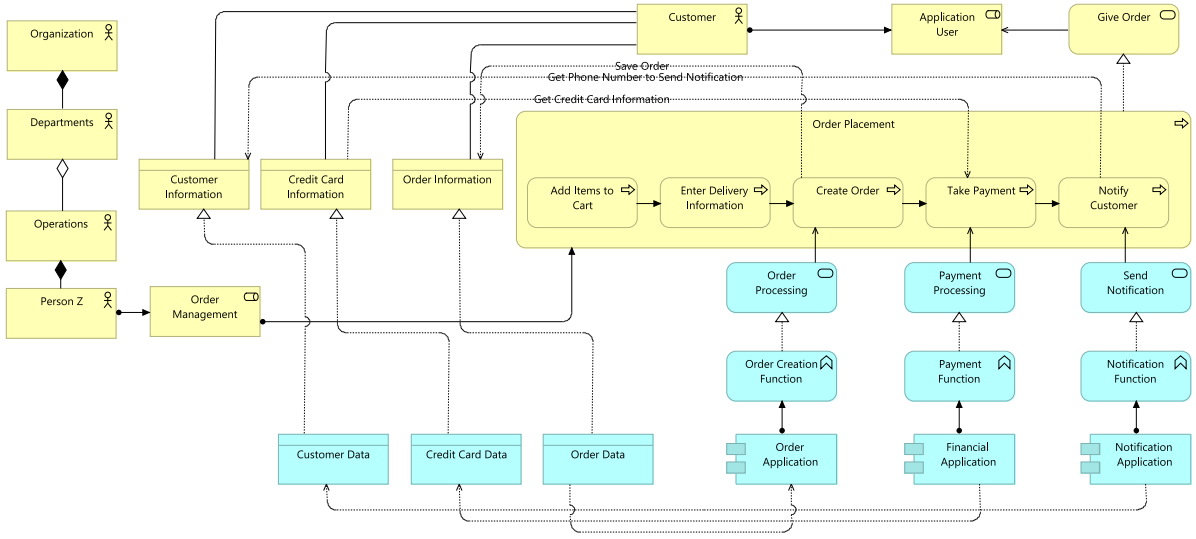

Figure 5.9.: An exemplary EA model of an order placement process, its supporting applications, the processed data and the responsible business units (Huth et al., 2019b)

In Huth et al. (2020b), we develop a four-level framework for how EAM supported GDPR implementation projects (cf. Figure 5.10).

At the first level, EAM provided existing information to DPM, and DPM used this information independently to identify applications, processing activities and responsible persons. This level required only basic EA documentation, such as an application list.

While also focused on information support, EAM took an active role in information collection at the second level. Integrated EA tools support sending out surveys and including data protection related information in the models. Multiple respondents reported representing the full RoPA in the EA tool.

The third level aimed at supporting DPIAs and information security analysis. In DPIAs, the relational knowledge within EA documentation helps to identify possible risks for the rights of data subjects, e.g. if personal data is transferred to a third party. Information security analysis benefits from documentation about technical protection measures or access rights.

Finally, the fourth level describes the active role in influencing processing activities, such as defining processes for data subject rights, developing data deletion policies or documenting data processing agreements. Especially notable was the holistic assessment and planning support to balance data protection, business and EA requirements.

5.5.2. Interdependencies with software development

Closely associated with the development of business processes is the software development process. Software developers, as Hadar et al. (2018) point out, are often very familiar with information security, but lack the required terminology for effective consideration of privacy requirements. The DPM process relates to software development in the specification of the privacy requirements that need to be considered.

During the development process, software developers are free to choose the approach that fits best. In Huth and Matthes (2019), we present the relevant requirements for technical and organizational measures (TOM) in the GDPR and analyze a selection of approaches that aim to incorporate privacy aspects in software engineering. Once the development is finished, documenting the TOM is an activity that again intersects with the DPM process.

We propose an approach for incorporating the privacy & security requirements in the agile development process of a business software company in Huth et al. (2020a). The approach involves

- Managing the organizational privacy & security requirements in a project management tool.
- Assigning individual items out of the requirements (e.g. *revoke consent*) to single business requirements. Based on the text description of the business requirement, the tool proposes privacy requirements from similar business requirements in different development projects.
- Browsing proposed solutions and/or rejection comments to the assigned requirements from previously documented solutions.

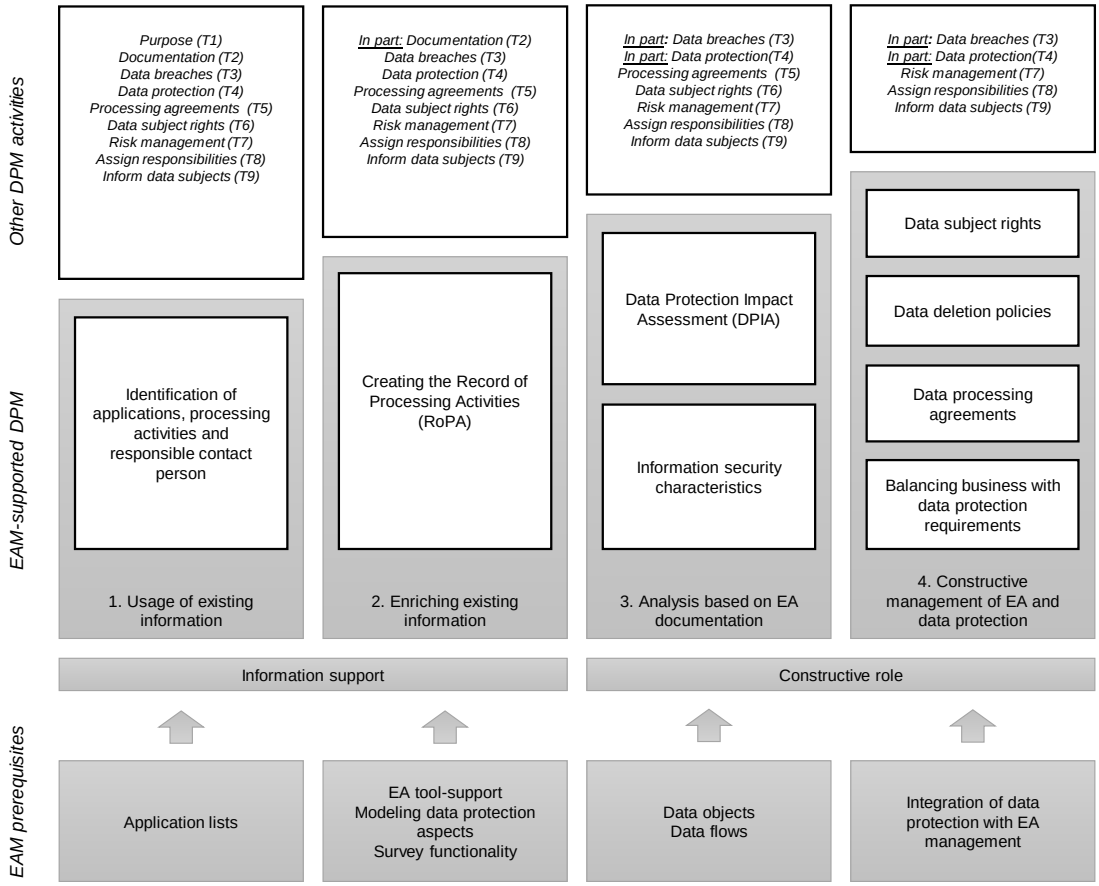

Figure 5.10.: Four levels of EAM support for DPM (Huth et al., 2020b)

- Documenting the solution draft (e.g. used libraries, code snippets) or rejection comments (e.g. *"consent not necessary, processing necessary to fulfill contract"*). Thereby, the TOM are documented and the internal knowledge base is expanded.

A prototype implements the approach and serves as the basis for preliminary feedback and ideas. Within a small focus group discussion, three key insights evolved (Huth et al., 2020a):

1. Implementation support for technical solutions to privacy challenges is most important for developers who are unfamiliar with the regulation. As the experience with data protection measures increases, supporting tools might not be needed any longer.
2. Product owners are concerned with business-driven features, but are accountable for the implementation of data protection measures as well. Bridging this gap can free up resources that can be invested in developing business-driven features.
3. Privacy checkpoints could be integrated into the regular development workflow, instead of checklists on the side.

5.5.3. Interdependencies with other internal processes

The discipline of IT governance, as defined by De Haes and Van Grembergen (2004), is

> **Definition:** IT governance
> ... the organizational capacity exercised by the Board, executive management and IT management to control the formulation and implementation of IT strategy and in this way ensure the fusion of business and IT. (De Haes and Van Grembergen, 2004)

IT governance has the tasks to define structures, processes and relational mechanisms (De Haes and Van Grembergen, 2004, p.28). Organizational structures include assigning responsibilities and decision-making authority. The implementation of processes is achieved by invoking strategic planning committees who oversee major IT projects and priorities. To establish relational mechanisms, possible tactics are creating business/IT partnerships or raising stakeholder awareness. COBIT, which we presented in Section 3.2.3, is a well-known framework for IT governance.

The DPM process influences IT governance and vice versa. Data protection regulation influences the responsibilities that must be assigned for processing activities. Especially the GDPR constitutes a major driver for large change projects, as we observed in the in-depth interviews and many industry publications.

Next, the DPM process affects the core business processes of an organization, because there is a direct interaction with the customers. Depending on the type of customers (private or business) and the type of product or service, there is a considerable amount of personal data that is processed: customer name, address, birthdate/age, financial information, personal interests, maybe even health information. Therefore, it is crucial for these core business processes to strictly consider the regulations in the GDPR, i.e. the legal basis for processing, the purpose limitation and the designation of a responsible person within the controller.

Lastly, other supporting processes intersect with the DPM process. Human resources processes personal data of applicants and manages addresses and bank account information of employees, as well as vacation and working times, sick leaves and possibly information about disabilities and medical conditions. Data that is processed in business intelligence (BI) can easily be aggregated to analyze working behavior of single employees, thereby creating a new processing activity which would have to be documented and covered by an appropriate legal basis.

5.6. Usage of ProPerData

To apply ProPerData (cf. Figure 5.11), stakeholders should first identify their role-specific responsibilities in the provided RACI matrix. The main responsibility type we address with ProPerData is **R - responsible**. Once the relevant work units are identified, the stakeholder can locate these work units and their context in the ProPerData overview canvas. The assignment to a DPM task and the temporal assignment describe the context of the respective work units.

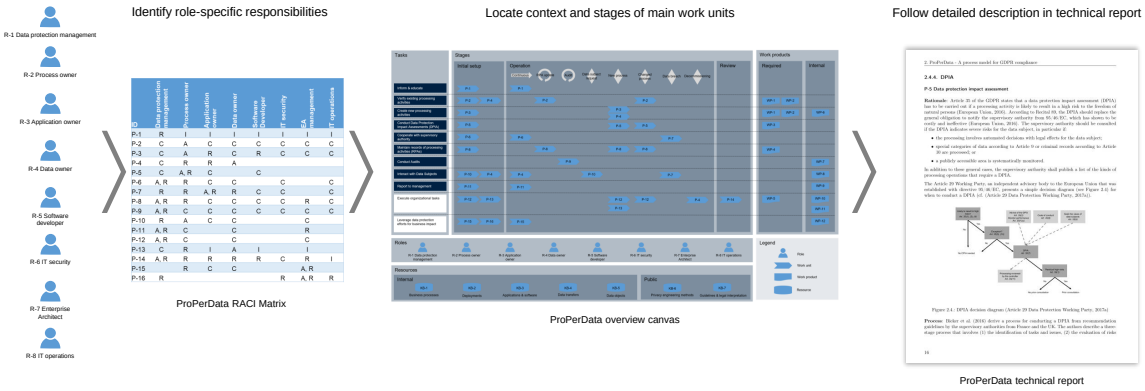

Figure 5.11.: Usage scenario of ProPerData

If the stage shows the need to take action (e.g. if a new business process is developed), the stakeholder can then refer to the detailed textual description in the technical report.

However, this is only the straightforward description of the usage scenario. The overview canvas of ProPerData may also be used as a communication and planning tool or to promote understanding of the regulation and DPM.

CHAPTER 6

Evaluation

Design science research aims at the evaluation of IT artifacts that address the identified organizational problems (Hevner et al., 2004, p.77). This chapter presents our validation results for *ProPerData*.

We will first describe our evaluation approach in Section 6.1. Next, we present qualitative results from in-depth interviews with ProPerData stakeholders in Section 6.2 and survey results in Section 6.3. We evaluate the reference process model against the research gap in Section 6.4 and summarize the evaluation results in Section 6.5.

6.1. Evaluation approach

The evaluation of reference models is particularly difficult due to methodological and practical reasons (Schermann et al., 2007). In addition to the challenges that are presented by the evaluation of conceptual models, the evaluation of reference models - since they claim generality for a range of related problems - must take into account the variety of requirements and constraints within this group of problems (Frank, 2006, p.119).

The range of problems that are addressed by ProPerData are GDPR implementation approaches in organizations that engage in EAM. There is no restriction of the business model or organization size, which implies a wide variety of requirements that should be validated. Also, GDPR implementation projects last multiple months or even years, and involve almost the entire organization. Due to this large scope of ProPerData, we consider it impractical to evaluate the artifact through real world application.

Frank (2006) acknowledges the impracticality of real-world evaluation of reference models and stresses the necessity of a multi-perspective approach. By considering multiple perspectives, the

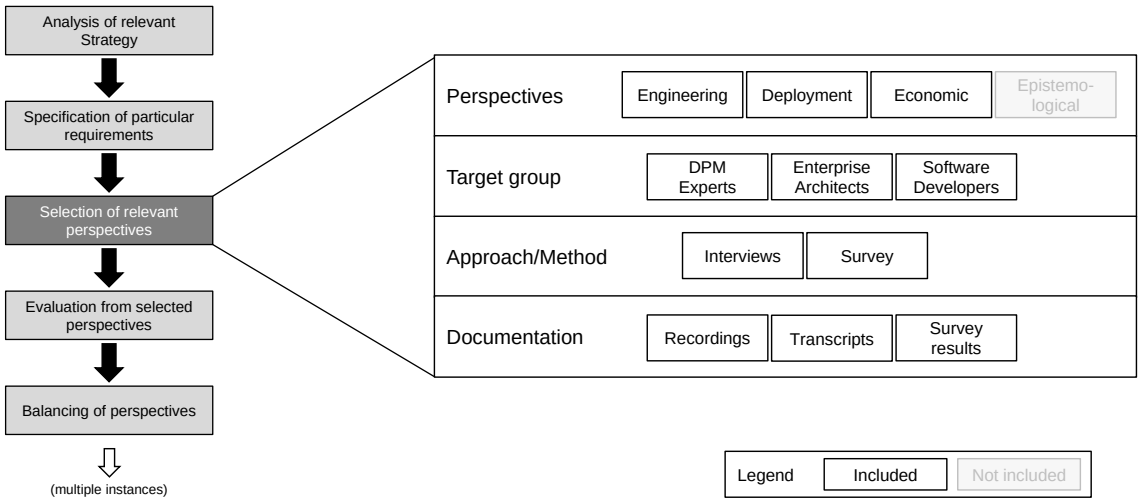

Figure 6.1.: Evaluation framework for reference models (Frank, 2006), configured for this evaluation

evaluation gets closer to an objective and balanced judgement (Frank, 2006). To structure these perspectives, Frank provides a conceptual framework, which includes the following aspects:

Economic perspective Criteria that are relevant for judging the costs and benefits that arise from the use of the reference model.

Deployment perspective Criteria that are relevant for those who work with the model, e.g. comprehensibility or compatibility.

Engineering perspective Evaluation of the reference model as a design artifact that has to fulfill the specified criteria.

Epistemological perspective Evaluation of the reference model as the result of scientific research.

To apply the conceptual framework, Frank specifies a process model for evaluating reference models. The process model takes into account the four perspectives and the underlying considerations for the conceptual framework (see Figure 6.1). As stated by the author, the proposed method (i.e., the conceptual framework and the process model) should serve as structuring guidelines, rather than clear directions. We will follow these guidelines in our analytical (Section 6.4), qualitative (Section 6.2) and quantitative evaluation (Section 6.3).

The process starts with the strategic analysis of the reference model. Only if a benefit can be expected from the model, the development should continue. We determined the prospect of a potential benefit in Chapter 2 and therefore developed ProPerData. In addition to generic criteria for reference models, specific requirements must be defined, whereby the level of detail depends on the specific case. We specified eight explicit requirements in Chapter 4.

In the actual evaluation, the relevant perspectives must be considered. We compose our evaluation of three different approaches, each of which focus on a different combination of the per-

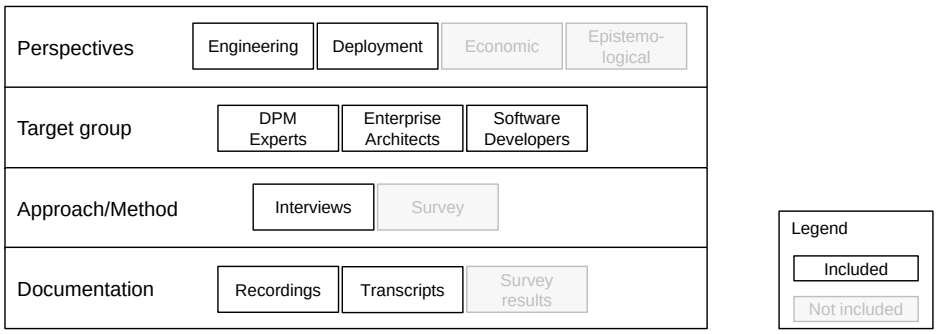

Figure 6.2.: Parameters for the qualitative evaluation

spectives. According to Frank, not all perspectives are mandatory in every evaluation project. As an example, the necessary theoretical background for a full discussion on the epistemological perspective might be missing. We elaborate on the combination of the perspectives in each evaluation approach.

6.2. Qualitative evaluation

6.2.1. Approach

For the qualitative evaluation, we conducted eleven interviews with 14 experts from target stakeholder groups of ProPerData. We held the interviews via video calls. Nine of the interviews were recorded and transcribed for thorough analysis, two were documented with notes.

In the interviews, we first explained the motivation behind ProPerData and the construction approach. Then, we presented the ProPerData overview canvas and an exemplary work unit description that addressed the target group of the interview partners (e.g. software developer or enterprise architect). Subsequently, we asked open-ended questions from a prepared questionnaire.

The interview guideline included the following topics and questions:

Interview partner background

- What is your industry / role / experience?
- Does your company use a method or framework to maintain GDPR compliance?

Correctness

- Does ProPerData match your experience?
- Do you consider ProPerData as relevant for attaining and maintaining GDPR compliance?

- Are the model elements (roles, work units, work products, stages and resources) represented correctly?

Applicability

- Are the work units applicable in practice?
- Is ProPerData applicable in small organizations, large organizations, or both?
- Can you imagine using ProPerData or parts of ProPerData in your work?

Comprehensiveness / Completeness

- Can you make a statement about completeness of ProPerData?
- Is the level of detail adequate for the problem at hand?

Formality

- Is the formality adequate to address GDPR compliance?
- Is the representation clear?

Related approaches

- Do you know any other approaches with similar objectives?

Further remarks and suggestions

- Participants were asked to share further comments about ProPerData or other topics that they regarded as relevant.

We will group and discuss the responses in the following section.

Table 6.1 lists the participants in the qualitative interview. We specifically addressed interview partners that made significant contributions in the development phase to increase the significance of the feedback.

6.2.2. Results

As shown in Table 6.1, we had in-depth discussions with the stakeholder groups *enterprise architect*, *software developer* and *data protection expert*. To clearly illustrate possible differences between the groups, we present the evaluation results separately.

Enterprise architects

This section summarizes the results obtained from the enterprise architects in interviews E1, E8, E9, and E11. We discussed the construction approach, the ProPerData canvas and EA support in DPM in the presentation part, followed by the set of open questions that addressed the following topics.

Deployment perspective

ID	Role	Experience	Industry	Length
E1	Enterprise Architect	12 years	Manufacturing	1:03
E2	Software Developer	20 years	Software	0:35
E3	Data protection expert	3 years	Software	0:43
E4	Software Developer	9 years	Manufacturing	0:41
E5	Data protection expert	1 year	Software	0:41
E6	Software Developer	25 years	Consulting	0:27
E7	Senior Consultant	16 years	Software	0:54
E8	Enterprise Architect	3 years	Software	1:03
E9	Enterprise Architects (3)	13 / 8 / 4 years	Research	1:37
E10	Software Developer	3 years	Software	0:28
E11	Enterprise Architect	3+ years	Banking	0:52

Table 6.1.: Participants in the qualitative evaluation interviews

Applicability in practice Interviewee E1 assessed the work units themselves as applicable, since they specify the goal, the process, and the involved parties. While the text descriptions contained the relevant information, such as the information sources, the expert suggested adding context diagrams with boxes and arrows to have a visual representation of the relationships. Interviewee E8 supported this assessment, as the descriptions provide a basis for communication about the work units. However, a simple step-by-step instruction would further support this applicability.

Suitability for different organization sizes The target organizations of our reference process model are large organizations that engage in EAM, so we derived ProPerData from organizations with EA departments. However, interview partner E1 pointed out another perspective:

> *"I would say it is also suitable for small organizations. In the end, we are talking about the same requirements, independently of whether it is a small organization or a large one. Of course, some roles or work units might be combined. [...] But the process should always be the same."*

In response, we adapted the resource descriptions to more general concepts, such as *deployments* instead of *CMDB* or *Applications / Software* instead of *EA application list*. Enterprise architect E8 shared this opinion. Regarding the suitability to large organizations, both experts fully agreed.

Possibility of using ProPerData or parts thereof in practice Interview partners E1, E8 and E11 could imagine using ProPerData as a reference when initiating a GDPR implementation project. E8, who developed an internal approach to address the GDPR challenges, emphasized the potential of a reference process model in the initial stages:

> *"When I was searching for a method in the beginning, I would have tried that. Absolutely. Because we didn't find anything. We found a few white papers, some recommendations from consultancies. But what I needed was something like this, what are the roles, the resources, the activities, ..."*

Respondent E1 liked the formality of the model in comparison to the rather textual descriptions that are used by the company. As a highly integrative process, DPM must be understandable to a large audience and the visual representation, accompanied by textual explanations, could contribute to that goal. Thereby, ProPerData could also serve as a reference to support project managers in the consideration of data protection aspects.

Enterprise architect E11 was sceptical whether an organization would replace the DPM systems that are already in place, but could imagine a use case where the measures that are already in place are matched to the reference model to verify compliance.

Formality and level of detail Enterprise architects E1 and E8 considered the formality of the model as adequate, because it employs clear graphical elements and a structured approach.

Regarding the level of detail, E8 remarked that the required details depend a great deal on the role. As an example, the DPO would probably need more material than what is included in ProPerData. As the work units are directed at specific stakeholders, the respondent regarded this requirement as fulfilled.

A different opinion was stated by E11. Recalling the amount of workshops and people that were involved in the company's GDPR project, e.g. in P-8, the interviewee did not consider the complexity to be represented sufficiently.

One respondent in interview E9 regarded the overview as a core contribution to the topic of GDPR implementation, but added that it is equally important to illustrate the possibilities for collaboration and the interrelations with other internal processes.

Engineering perspective

Relation of own professional experience to ProPerData According to E1 and E8, the presented model matched their professional experience with the GDPR implementation projects. E11 identified similarities with the company's approach, but noted that the tasks are specified differently. E8 recommended extending a method description for applying the reference model:

> *"I think you have asked the right questions and developed the right solution. What I can add from practical experience is that the people will ask: Ok, we have the roles, the work units and so on. What do we do now? [...] If at all, I would think about the procedural aspects."*

Coverage of GDPR aspects Interview partner E8 confirmed the relevance of ProPerData for achieving GDPR compliance, noting that if only a fraction of the necessary measures had been implemented, companies would have reduced the risk and amount of fines significantly.

E1 and E11 described the two aspects that ProPerData has for GDPR compliance - first

for setting up GDPR compliance measures, and then checking the established measures continually for compliance. Both had already finished the initial setup part, and thus focused more on the compliance checks.

> *"I believe for comparing companies that is very helpful. Because it is not only a model, but also a language. A common understanding. So I think that helps, absolutely." (E1)*

Roles, responsibilities, work units, resources and work products To complement the stages *setup* and *operation*, expert E1 suggested adding another stage for improving the measures that are in place. Thus, we added the stage *review* to the overview canvas, as well as a corresponding work unit.

As E8 commented, there is no model of roles that accommodates every situation that occurs in practice. Consequently, the respondent did not object to the roles as we specified them in ProPerData. However, some roles might be combined in many instances, such as the data owner and application owner.

Neither E1 nor E8 noticed missing work units, work products or resources, and thus approved of the respective elements that are included in ProPerData.

Additional remarks

An interview partner in interview E9 suggested enhancing the perspective of ProPerData to not only describe the internal work units of DPM, but to also illustrate the relationship to IT governance and management. It is crucial to address the questions of responsibility and risk management more thoroughly. Further, the interviewee recommended adopting an EA-based view on the data objects to facilitate tracking the data flows across the organization.

Software Developers

This section discusses the responses of software developers E2 and E10, who work in medium sized companies, developer E6, who works as technical consultant, as well as E4, who works at a large enterprise. After explaining the construction approach and the ProPerData overview canvas, we showed the work unit that concerns the development of new processing activities. Since none of the developers signaled expert's knowledge in data protection, we do not present statements on completeness with respect to GDPR tasks.

Deployment perspective

Applicability in practice Respondents E4 and E6 saw a direct applicability of ProPerData, especially because the model communicates the legal text in a language that is readable and understandable for developers. The other developers pointed to some practical issues when adapting a framework for planning to the agile development process, as discussed by E10:

> *"It's applicable in some cases. Then again, in agile software development you can't really paint by numbers and go through issues one by one. You have to*

> *advance little by little. And if you spot an issue, you can involve a data protection expert. [...] Maybe at a smaller scale, you can have the process models."*

Suitability for different organization sizes While E4 discussed that in general, ProPerData should also apply to smaller organizations, all respondents agreed that such a reference model is most suitable for large organizations. Large organizations have to scale and make information available to a larger audience, whereas smaller organizations can usually achieve the same result in small meetings.

E6 considered the development and use of tailored software solutions as the criterion for the suitability of ProPerData. This criterion is typically met starting from medium sized solutions. Further, as E4 noted, the relevance of data protection depends heavily on the number of organizations and the type of product that the company offers. In the context of a large industrial enterprise, the focus is on processing activities of employee data.

Possibility of using ProPerData or parts thereof in practice According to interviewee E6, developers would not use the overview canvas, but

> *"something more concrete, something like a checklist. A simplified version, basically just the track that touches them."*

This appreciation is shown by E4, who stated:

> *"I would say I wouldn't like it in general, but since I have to be compliant I would be glad to know there is some kind of support, this kind of support. Especially because it is pretty structured. You don't need to go through a large number of texts and you have some visual guide to it. So you understand the process and also the concepts you need and so on."*

As an employee of a growing medium-sized company, developer E10 reported clearly assigned responsibilities, but otherwise considerable freedom in how the respective tasks are executed. With increasing company size, the respondent noted, more documentation has to be created. This development is also driven by various certifications, such as SPICE[1]. Thus, application of the overview canvas would rather be an issue for decision makers, as confirmed by E6. In any case, usage of ProPerData would require adaptation to the company context, since it is only a reference model.

Formality and level of detail Participant E4 pointed out that there are different types of developers with different tastes. In the participant's perception the usage of UML, BPMN and block schemata was easy to grasp. This makes the notation an adequate choice for experienced software developers in larger organizations. The formality in the overview canvas, especially the temporal relations, appealed to developer E10 as well:

> *"In the end, using swimlanes and a matrix type representation is not a revolutionary thing. And that's a good thing. You've seen it before, and that's why you can handle it quite well."*

E1 remarked the challenge to transform an academic formalization to the practical context,

[1] ISO/IEC 15504 - Software Process Improvement and Capability Determination

where a formal model is sometimes difficult to explain. As a result, the representation of the overview canvas would have to be adapted, which, as E6 stated when asked about the applicability, would not be of much use for developers anyway.

Engineering perspective

Coverage of GDPR aspects All four respondents were convinced that a process model like ProPerData contributes to GDPR compliance. In particular the complete overview picture and its structure were regarded as helpful, because it allows a company to check its current efforts and identify possible gaps. According to developer E2, the structure can also support communication between different stakeholders:

> *"I realized that when we do the ad hoc process and check with the legal department, especially in our legal department, there is no structure how they do it or how they address it. So that's why I consider that having a process model and to clearly define the roles and the tasks could help the communication, because sometimes I see that it is very unstructured."*

Further, the overview picture creates an understanding of the involved roles and the necessity of them communicating well, as E6 added.

Relation of own professional experience to ProPerData The big picture approach of the ProPerData overview canvas was shared by the company of developer E6:

> *"Some of our customers have been very very concentrated on just one tiny detail, and we want to make sure that rather the big picture is met with some controls, than just that some concern is met with very hard controls. So the emphasis is what is changing, and the big picture, the map is very important. And we see it quite similarly."*

The developers in medium-sized organization, especially E1 and E6, pointed to the difference between the agile and the cascading development approaches, which also affects how data protection is enacted. As E1 expressed it,

> *"In the cascading approach it was easier to adapt, and to link the steps to change the process. But the way we are working now, we do Scrum adapted to us, and in this time it's very difficult. The process is more ad hoc, it's more like a policy instead of a formalized process how to do it."*

However, referring to the incremental improvements, interviewee E6 noted that the increments contain essentially the same phases, and especially the events (such as data breach) were very familiar.

Developer E4 reported routine security clearing of software, where the IT security department provides guidance on how to develop secure software. Due to the limited amount of personal data that is processed by the company the security aspect is prevalent in software development. The respondent described how it is typical that software is built to demonstrate an innovative idea first, and made to comply with security and privacy rules at a later time. Considering compliance rules early on in software development could save

significant amounts of work later on. However, according to E4, it is uncertain whether such a way of thinking will ever be adopted.

Stages, roles, responsibilities, work units, resources and work products Similar to enterprise architect E1, developer E2 suggested adding a stage to prescribe how the process can evolve. Subsequently, we introduced a review stage to the reference process model.

The importance of the ongoing nature of GDPR compliance projects was addressed by E6:

> *"So as I mentioned my worries that people often tend to see this as a one time thing, they do it as a project and then GDPR is taken care of. But in this model I see that we have changed process and new process, which would be triggers to check something or do something. So that's a good thing."*

The further substages of the operations stage, such as the decommissioning phase, new process or audits, received approval from E6 as well.

Respondent E2 identified an exact mapping between the ProPerData roles and the roles within the company. Software developer E10 mentioned that, while the software developer would not have come to mind as stakeholder of a GDPR compliance effort at first, the inclusion in an overall approach makes perfect sense.

Some uncertainty related to the representation of the legal department (raised by E6), which we intent to cover with the role DPM, and the interpretation of the DevOps role (raised by E4). We consider the DevOps role as a combination of the software development and IT operations roles in one person or team. Another possible specialized role that we did not include in ProPerData could be the quality assurance engineer, who tests the final software product. We interpret this role as a part of software development as well.

An additional discussion concerned a possible role of data steward, which was suggested by E6. The role would be concerned with harmonizing work across different services and projects, with a focus on the data objects. This role description could be assigned to either the EAM role, or to the data owner role, as the data owner has to supervise which processing activities are conducted on a particular data object (e.g. data subject address).

Contrary to the privacy-aware software development process we describe in P-3, interviewee E10 explained that the product owners at his employer use common sense to determine whether a further consideration of data protection is necessary in software development. When developing or adapting software to a customer's context, the role of data processor (cf. Section 2.2.1) becomes relevant. This situation, which applies to E2, E6 and E10, adds another layer of complexity, because the roles within the ProPerData canvas could be assumed by internal or external employees, as interviewee E6 discussed. As an example, test data might have to be anonymized as well, implying further work units and responsibilities. However, since ProPerData is aimed at data controllers, we factored in this aspect with the processing agreements, which are listed as a work product.

Respondent E4 suggested distinguishing between internal and external work products. We adopted this idea to show more clearly which work products have to be reported to the supervisory authority.

Concerning the resources, E2 pointed out that similarly to roles, two different concepts - such as the business process repository and the application repository - could be combined in a single physical entity, e.g. the EA repository.

Additional remarks

Agile development A topic that triggered valuable discussions was the question of how to align the requirements of data protection with agile development. E6 and E10 described similar approaches of increasing awareness of developers about data protection so they can recognize situations that require further scrutiny:

> *"So we are using autonomous agile teams. And the idea is that instead of pushing the design so much from outside in or from upside down to the teams, the idea is that teams would be empowered and they would be also educated as much as possible to also be aware and cautious about these things." (E6)*

Increasingly, automated tools support checking for security issues in agile development, as E2 commented.

Method support for applying ProPerData Another suggestion for extension of the ProPerData reference process model was to add concrete steps that would make it easier to apply the model. As one possibility, E4 proposed visualizing the work products. A further suggestion by E10 consisted in the creation of role-oriented overview canvases.

Other drivers Legislation is not the only driver for implementing data protection measures. E10 explained one instance where the worker's council prohibited analyses on employee productivity in a project management tool. As a result, the company was forced to disable the feature. E6 argued that it is also important to maintain efficiency and balance when aiming for privacy:

> *"[...] and then they buy the heaviest encryption you can have. Often for example just the encryption as a placebo. So it might be used in places where it is not actually increasing your privacy. So I'm a big fan of a holistic view, of taking a look at the big picture. And I'm not at all opposed to having some process to guide you there."*

Data Protection Experts

In this section we present the evaluation responses from interviews E3 and E5, who both work in the privacy tech industry. In the presentation, we went over the construction approach, the overview canvas, and then explained the work units for software development and EA-supported RoPA creation before turning to the prepared open questions.

Deployment perspective

Applicability in practice Speaking as a data privacy practitioner, interviewee E5 reported that it is typically not possible to follow the exact same pattern in each organization, because there are always particularities to consider. Especially the processes that interact with the

DPM process imply the need to tailor the approach. Despite the need for expert knowledge about internal processes,

> *"I do think that it can help to approach the topic with a model. Maybe you would have to adjust the model here and there to fit the internal business processes."*

I.e., the reference process model would have to be instantiated in the particular organizational context.

Expert E3 questioned the direct applicability, as ProPerData does not include instructions on how it should be instantiated. The respondent especially considered the UML diagram in P-3 as too abstract to be directly applicable.

Suitability for different organization sizes Respondent E3 discussed the general suitability for different organization sizes in detail. As the tasks are derived from the legal text, they are independent of organization size. The approaches differ between small and large organizations mainly in the stakeholders and the resources that are used for executing the work products:

> *"What happens in a small organization is that the stakeholders are not different people, but one person. The resources... sometimes they don't even exist in the large organizations, even less so in small organizations. Typically, the resources are the minds of the people [...]. That doesn't contradict the assumption that it fits for one organization size and doesn't for the other, but I think it collapses a bit for small organizations."*

Possibility of using ProPerData or parts thereof in practice Interview partner E5 doubted being able to use ProPerData in everyday work, but could imagine very well that data protection consultants would. E3 related ProPerData to the two different products that their company, a privacy tech company, offers: the solution for small companies prescribes a linear, step by step process that uses to do lists to guide the responsible person. The enterprise solution offers just the tools without any prescription on what to do next.

> *"I am thinking about the future of our software, whether there could be a link. To be able to say as a small organization, we can't do everything sequentially. Or as a large organization, I would like to have some more guidance. That's where I would consider this as a smart approach. In particular the temporal aspect."*

Level of detail The necessary level of detail strongly depends on the targeted stakeholders, as interviewee E5 replied. To non-experts on DPM, the ProPerData overview canvas would have a sufficient level of detail to serve as entry material to the topic, because it summarizes the most important aspects and how the work units should be conducted.

In contrast, respondent E3 considered data protection knowledge as essential, if the reader should not be overwhelmed by the tasks. However, investing 5 to 10 minutes in understanding the overview canvas would lead to sufficient understanding and therefore make it very feasible.

Engineering perspective

Coverage of GDPR aspects Regarding the overall coverage of GDPR aspects from a data controller perspective, both E3 and E5 stated that all the main aspects are included in ProPerData. Nonetheless, any attempt at completeness imposes the question for the level of detail. To check for completeness, E3 would use the SDM, which we described in Chapter 3, as a reference. We must note that the DPM experts we interviewed do not have a legal background, and thus mainly assess the procedural aspects of ProPerData.

Relation of own professional experience to ProPerData Respondent E5 could relate ProPerData very well to practical experiences with customers:

> *"I believe the things that are included in the model reflect the activities that have to be conducted for GDPR compliance in a very, very structured way. I believe that such a structured representation rarely exists. At our customers, I have rarely seen such structured, complete representations. Many times I see parts of that, where someone who is responsible for a subarea conceptualized and visualized it."*

Expert E3 also liked the structure of the overview canvas, noting that it helps to "dissect the monster". According to that same person, especially the operations stage is very familiar:

> *"Those are things where every DPO says: yes, yes, yes, that's exactly how it is."*

Roles, responsibilities, work units, resources and work products According to respondent E3, the roles require careful consideration. Since the ProPerData roles embody an abstract reference scenario, it is possible that an organization defines them differently. Aside from the role definition, the expert also referred to questions of governance, as not all the people that should be involved always are in a real world scenario.

Referring to the work units, respondent E5 criticized the definition of a particularly narrow work unit *define data owner*, which differed from the abstraction level of all other work units. Subsequently, we integrated the work unit with P-3 and aligned the abstraction level of the work units overall.

When asked whether the work products were represented correctly, both interview partners again issued a note of caution with respect to completeness. The decision about the satisfaction of legal requirements is ultimately taken by courts who interpret the law, so the goal should be to adhere to the legal text and the publications by official authorities. E3 commented that the main informational resource is in fact the regulation itself, possibly accompanied by A29WP or EDPB guidelines or the material that is provided in the context of data protection certifications by data protection associations, such as the International Association of Privacy Professionals (IAPP).

Regarding the internal resources that ProPerData describes, respondent E3 felt overwhelmed by the EA-driven terminology of the concepts due to missing experience in the field. The discussion about the concepts that are embodied by the terms, e.g. the EA application lists, led to a mutual understanding and ultimately to the inclusion of the general concepts rather than EA-specific terms.

Additional remarks

Interviewee E5 gave some additional insights from the perspective of a privacy tech company. Whereas ProPerData explicitly names the knowledge bases for business processes, applications or deployments, the expert stated that even in surprisingly large companies, no such documentation could be encountered. In such cases, it pays to be able to identify the responsible people within the organization. The strong focus on EA-based information was viewed sceptically, as it might limit the applicability of the reference model (or add to reservations against it). We adapted the reference model to state concepts rather than specific instances.

Despite the suggestions, the interview partner also acknowledged the multitude of opinions by different people. Since not everything can be acted on, they have to be merged on the basis of a common denominator. This holds for the level of detail as well. Each point within the model could be extended on, but the expert viewed the presented level of detail as reasonable.

Table 6.2 summarizes the discussion of the qualitative interviews with three stakeholder groups.

		Enterprise Architects	Software Developers	Data Protection Experts
Deployment	Applicability	visualizations support understanding	diagrams support applicability	missing instructions on how to apply or instantiate the reference model
	Organization size	high-level concepts applicable in any organization size, but the process model might conflate	helpful in large organizations, too extensive for small organizations	resources are at best modeled in large organizations
	Usage in practice	as a frame of reference	would mostly be interested in own tasks	only in right context
Engineering	Formality and level of detail	requirements heavily dependent on role necessary level of detail higher in practice	requirements depend on background of individual developer	contradicting opinions: a) only for introduction to DPM b) DPM knowledge necessary to understand ProPerData
	Relation to professional experience	matches professional experience, but specific instructions would be the next step	agile development not adequately represented data protection requirements contradict the imperative to 'build fast'	holistic, structured overviews do not exist in practice matches professional experience
	Coverage of GDPR requirements	no obvious blind spots	no statement possible	no obvious blind spots completeness would have to be checked against SDM
	Plausibility of ProPerData elements	iterative/ongoing nature of compliance approach important	agreed discussed aspects of the elements	correct, but EA concepts needed explanation
	Additional remarks	relationships should be illustrated representation of responsibilities important	relationship to agile development and tool support important pointed to workers union as other driver for data protection	ProPerData is limited to large companies which engage in EAM

Table 6.2.: Summary of qualitative interviews with three stakeholder groups

6.3. Quantitative Survey

6.3.1. Approach

After the qualitative interviews, we conducted a quantitative survey (cf. Figure 6.4). We addressed 170 potential ProPerData stakeholders in the German speaking area directly via

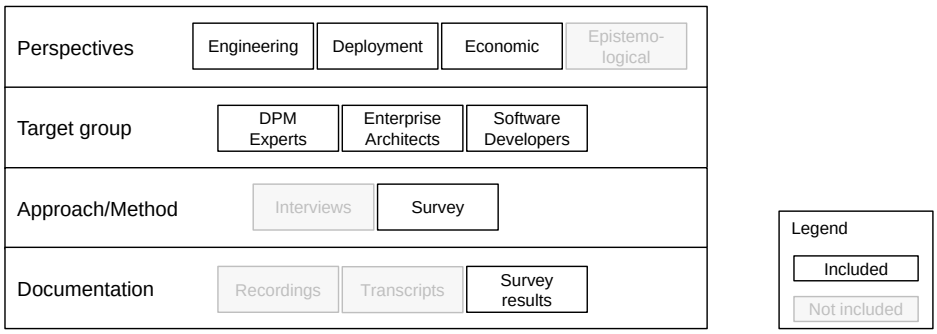

Figure 6.3.: Parameters for the quantitative survey

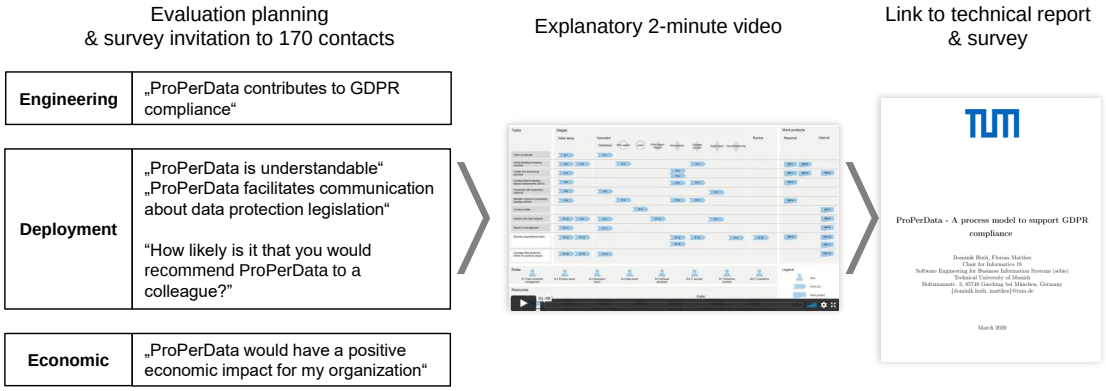

Figure 6.4.: Approach for the quantitative survey

email. The email shortly explained the research topic and referred to a 2-minute introduction video on ProPerData. At the end of the video, we placed a prominent link to the technical report on ProPerdata (Huth and Matthes, 2020) and a less prominent link directly to the online survey. The links to the video, the technical report and the survey were also included in the contacting email. We encouraged the recipients to forward the request to knowledgeable colleagues as well.

The survey started with a self-classification of the participants to the ProPerData roles. We allowed multiple selections to reflect the possibility that multiple roles are held by the same person at once.

The questionnaire addressed the *engineering perspective* with the statement

> “ProPerData contributes to GDPR compliance.”

Further, we addressed the *deployment perspective* with the statements

> “ProPerData is understandable.”
> “ProPerData facilitates communication about data protection legislation.”

Lastly, we addressed the *economic perspective* with the statement

> "ProPerData would have a positive economic impact for my organization."

The participants were asked to rate their agreement on a 5 point Likert scale (disagree / rather disagree / neutral / rather agree / agree). The reason we chose to keep the survey as short as possible was to ensure a high completion rate of the survey, since long surveys tend to deter participants from answering all questions. Thereby, we intended to get a more meaningful result from the survey.

Lastly, we addressed the net promoter score (NPS), a simple measure of customer satisfaction first proposed by Reichheld (2003). To determine the NPS for ProPerData, survey participants were asked:

> "How likely is it that you would recommend ProPerData to a colleague?"

As per definition of the NPS, the scale ranged from 0 (very unlikely) to 10 (very likely). The NPS is designed to counter a possible 'grade inflation' that is associated with satisfaction surveys (Reichheld, 2003). Participants who choose values $x \leq 6$ are classified as *detractors*, participants who assign $7 \leq x \leq 8$ as *passively satisfied* and participants who assign values of $x \geq 9$ as *promoters*. The net promoter score is calculated as the percentage of promoters less the percentage of detractors.

6.3.2. Results

The survey was sent out to a total of 170 recipients via email and included the invitation to forward the request to other interested colleagues. One company reported distributing the survey in an internal communication channel. The video page was loaded 98 times and the video was played 41 times. Out of the 37 participants who started the survey, 29 completed the questionnaire, which was open for five weeks between May and June of 2020.

The distribution of self-assigned, non-exclusive roles (cf. Figure 6.5) shows a strong bias towards *enterprise architects* (21 participants), who represent the main stakeholder group and also the main source of information on implementation projects. Further well-represented groups with 8 participants each include *software developers*, *process owners* and *DPM experts*. The roles *application owner* (4 participants), *IT operations* (3), *IT security* (3) and *data owner* (2) were less represented in the sample. We consider the distribution of participants as relevant and suitable for assessing a reference process model that assumes an enterprise architecture perspective, although the sample is not representative.

Figure 6.6 shows how the participants responded to four statements that represent the deployment and economic perspectives of ProPerData. Nine respondents considered ProPerData as understandable, and 14 considered it as rather understandable. We trace this back to the fact that most respondents are familiar with the terminology that is used by ProPerData.

Regarding the statement that ProPerData facilitates communication, one respondent fully disagreed. Unfortunately, none of the voluntary comments relates to this assessment, so we can

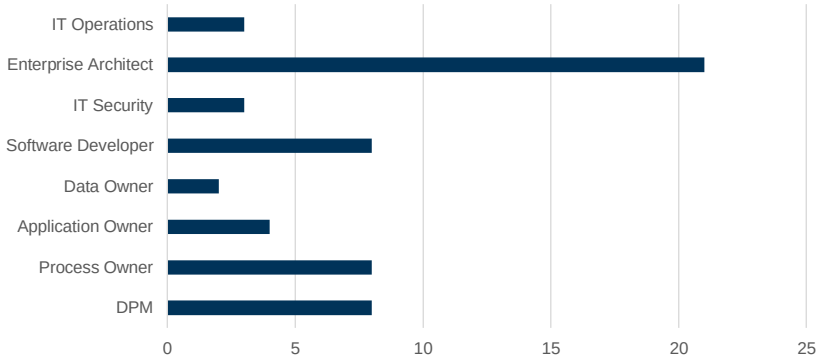

Figure 6.5.: Participant roles in quantitative evaluation (non-exclusive)

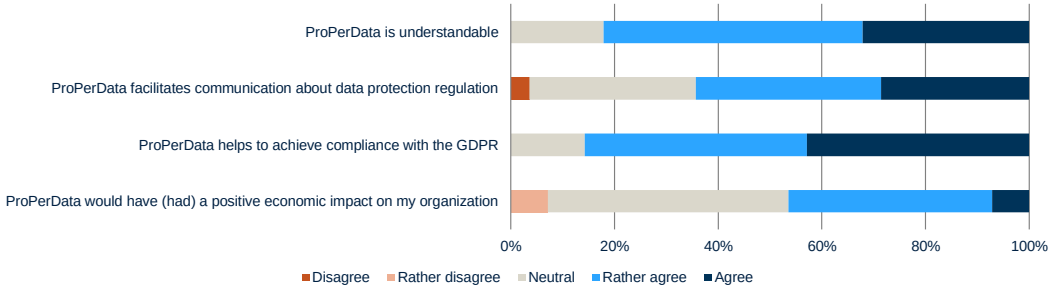

Figure 6.6.: Distribution of assessments of ProPerData

only speculate about the reason that the respondent had in mind. One possibility could be that the terminology is less helpful to other stakeholder groups.

The majority of participants saw a positive contribution of ProPerData in achieving compliance with the GDPR (12 agree / 12 rather agree). However, this strong agreement was not matched in the direct assessment of ProPerData's possible economic impact: less than half the participants responded positively to the statement (2 agree / 11 rather agree), while most respondents had a neutral position (13) and two respondents even rather disagreed with this statement.

A breakdown by selected participant roles is shown in Figure 6.7. On average, enterprise architects assign the highest rating in all four categories. This underlines the strong focus of ProPerData on the EA perspective.

While the ratings for ProPerData's contribution to GDPR compliance and its understandability are fairly similar (4 for DPM / 4.3 for EAM / 4 for developers and 3.8 / 4.2 / 3.6, respectively), some disagreement shows for the remaining two statements - software developers slightly disagree with the statement that ProPerData could have a positive economic impact (2.8), while DPM and EAM experts tend towards agreement (3.2 and 3.5). ProPerData's role in communication about data protection is assessed as rather positive by EAM experts and software developers (3.9 and 3.8), whereas the data protection experts seem unsure about this topic (3).

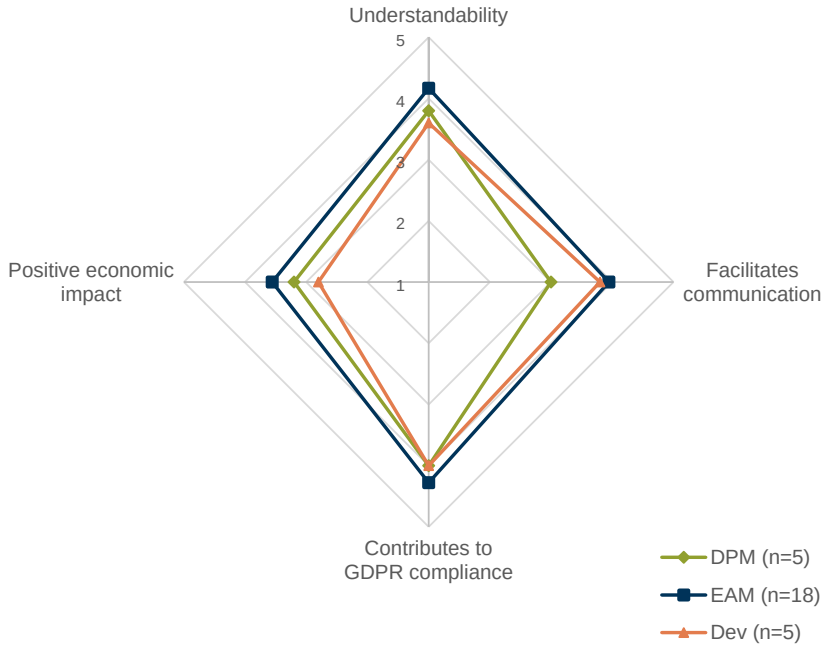

Figure 6.7.: Average assessments of ProPerData by participant role

Since DPM experts rated the understandability as rather good, a possible explanation is that DPM experts are already quite familiar with communicating data protection topics and do not feel like ProPerData would lead to a notable improvement in communication. Conversely, EA experts and software developers are less familiar with talking about data protection topics, so ProPerData might extend their communication toolset.

We also assessed the NPS for ProPerData with the question *"how likely is it that you would recommend ProPerData to a colleague?"*. As Figure 6.8 shows, the largest group of respondents (16 out of 29) are classified as *passively satisfied*, a group that the NPS does not include in the calculation. Nine respondents rated the likelihood with at most the value 6 and are therefore classified as *detractors*, while we observe four ratings of 9 or 10, which we classify as *promoters*. Overall, we can calculate the NPS as

$$\begin{aligned} NPS &:= \frac{\sum Promoters}{\sum Participants} - \frac{\sum Detractors}{\sum Participants} \\ &= \frac{4-9}{29} \\ &\approx -17\%. \end{aligned}$$

I.e., a net surplus of four respondents classify as detractors and the overall NPS is negative. The assessment of the NPS by role (cf. Figure 6.9) again shows differences between the stakeholder

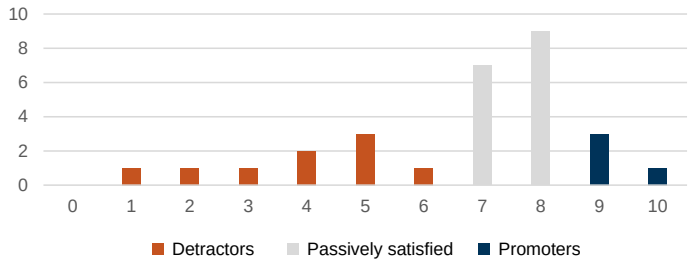

Figure 6.8.: Distribution of detractors, passively satisfied and promoters

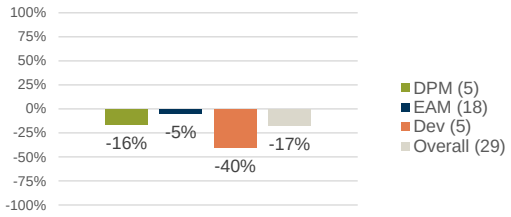

Figure 6.9.: The NPS by selected roles

groups EA experts, DPM experts and software developers. Even though EA experts rated the likelihood of recommendation higher than the participant average, their NPS is negative as well. Software developers showed a very low inclination towards recommending ProPerData. Borrowing from a statement in the qualitative evaluation, software developers do not consider overall compliance with data protection regulation as their responsibility. The value for DPM experts fits the overall result.

A possible explanation for the negative assessment may be found in the voluntary comments, which participants could leave at the end of the survey. Overall, ten participants used this opportunity. The comments included six messages of encouragement or praise, such as

> *"Great work! All relevant aspects covered and well-described."*

Two short remarks referred to the limited number of stakeholders that ProPerData applies to. Technically, management is accountable for establishing effective data protection measures, but as ProPerData shows, multiple roles must contribute towards this goal.

Another statement noted that ProPerData would have been helpful in 2018, when the GDPR entered into force. While this is a valid remark, it does not render ProPerData useless in the author's opinion. Only half of the companies in an industry study already claim a satisfactory level of GDPR compliance (IAPP, 2020), and changing business processes ensure relevance of ProPerData for the future.

One participant suggested placing more emphasis on the legal aspects, because even the prudent implementation of the measures described in ProPerData might leave some residual risk and should therefore be included in risk management.

Two participants offered to have an in-depth discussion of ProPerData. These two participants are included in the qualitative evaluation in Section 6.2.

6.4. Analytical discussion

6.4.1. Fulfillment of the GoM

As a conceptual model, a reference model must satisfy the general requirements for conceptual models. Schütte (1998) derives the *GoM*, which comprise the six aspects *adequacy of construction*, *adequacy of language*, *efficiency*, *systematic structure*, *clarity* and *comparability*, which we present in Table 6.3. We will discuss the considerations with respect to the GoM in this section.

Guideline	**Explanation**
Adequacy of construction	Making the construction process transparent; rationality of actions and decisions; adequacy of perspective.
Adequacy of language	Is the language capable of addressing the concerns that are inherent in the problem?
Efficiency	Considering a scarcity of resources, the model has to follow the efficiency postulate.
Systematic structure	Addresses the requirement to represent different perspectives on the modeling subject.
Clarity and understandability	The target group has to be able to understand the model, because they should be able to use it. Therefore, a clear visual representation is necessary.
Comparability	In an application context, multiple models might exist to solve the same problem. If no comparable model exists, transferability can also be assessed.

Table 6.3.: The GoM (Schütte, 1998)

Schütte considers *adequacy of construction* as the most important guideline, as it supports decision making about a reference model. Our research process is made transparent in Section 1.6 and throughout this thesis. Further, we refer to the construction process throughout this thesis and thus provide the reader with a sound base for judging the validity of the obtained results.

The second guideline, *adequacy of language*, is accounted for by the discussion that we present in Section 4.2.2. The metamodel we instrument for the representation of ProPerData captures all relevant aspects that the model describes and is therefore a suitable choice. ProPerData is intended to be an information model without any claim to be executable, so the chosen language with simple elements is sufficient for attaining this goal.

For the guideline *efficiency*, Schütte highlights the need to present a model that is adaptable to various factors. The trigger that forces such a change process could be motivated by the system itself, or by exogenous factors. ProPerData accounts for internal changes with the *review* phase, where the stakeholders have the opportunity to reflect and adapt the current data protection processes. An exogenous factor could be changes in the privacy regulation, as the reference model is based on the GDPR. We believe that ProPerData is open to such legislative changes and provides a sound structure for future changes to come.

In the overview canvas of ProPerData, the DPM tasks provide a clear structure for the work units and the work products. The stages further organize the work units, and resources and work products are subdivided into internal and external elements. This segmentation contributes towards fulfilling the guideline *systematic structure*.

To support *clarity*, we employed two hierarchical levels of abstraction - overview level and work unit level - that pursue the goals of communication and implementation, respectively. The layout of the overview canvas includes as few elements as possible, but is nonetheless extensive. This is owed to the complexity of the GDPR itself.

6.4.2. Discussion of specific requirements

In addition to the general GoM, Section 4.1 derived eight requirements for a reference process model that we used as further guidance in the development of ProPerData.

R1: A clear conceptual visualization of GDPR implementation approaches.

Since GDPR implementation projects differ significantly among companies and industries, we concentrated on the common elements. The ProPerData overview canvas links tasks, work units and work products, and specifies to which task a work unit belongs and which work products should be expected within this task. Further, the representation of stakeholders and resources completes the set of elements that we defined within the MPEM.

The visualization on the overview level is accompanied by the visual concepts on the work unit abstraction level. Since the work unit level is aimed at specific stakeholders, the visualization approaches depend on each work unit.

R2: Ability to capture temporal units and dependencies.

We encountered particular difficulties in specifying temporal relationships, since multiple work units can take place simultaneously. Additionally, a GDPR implementation approach can take place one step at a time, and although some work units lend themselves naturally be executed before others, there is no imperative sequence. Therefore, we restricted the model to specify the time frames and events at which work units take place. In this sense, we adopt the notion of a process model as a set of responsibilities, actions and time frames.

R3: Adaptability to the context of the organization as a socio-technical system.

It is clear from the regulatory text of the GDPR, which makes ample use of the term *'technical and organizational measures'*, that the DPM process is cross-cutting and involves both

people and technology. ProPerData supports the implementation of the work units without any prescription on technologies. The defined resources are intentionally named after the concepts they represent, e.g. *applications & software*, to allow the interpretation within the context of an organization - the concept could be represented by an EA application list, or it could be the applications that are managed through another approach.

R4: Incremental and iterative applicability.

As we already discussed in the context of Requirement 2, there is no rigid sequence in which the work units must be applied. However, there are dependencies between single work units, which are depicted in Figure 5.6. Despite these dependencies, an incremental implementation is possible, adding one work unit after another. The review phase is a checkpoint to reflect and adapt the implementation approach.

R5: Correspond with regulatory requirements and be representative for GDPR implementation projects from the perspective of EA.

We ensured that the set of tasks that describe the structure of ProPerData are validated by DPM experts, such that they can serve as a solid basis with a claim of completeness at an overview level. Nonetheless, the process model cannot replace the detailed specifications of the legal text. Each implementation challenge is different, but a solid conceptual frame supports in the specification of the details.

R6: Provide practical insights for implementation of single GDPR work units.

ProPerData is rooted in empirical investigations with enterprise architects, DPM experts and software developers, and therefore represents a practice-based perspective on GDPR implementation projects. The claim of generality of a reference process model, however, implies that certain company-specific details are omitted. We provide anecdotal practical references throughout the construction section of this thesis (Section 5.1) and within the work unit descriptions in the appendix.

R7: Foster reuse and value of the established artifacts and processes.

EA management addresses multiple concerns with a set of methods, concepts, models and tools that revolve about *the* architecture of the enterprise, i.e. its nature is reuse of established artifacts. By focusing on the perspective of enterprise architects, we foster the notion of reusability of the artifacts that EA makes available, e.g. relational knowledge about processes, applications, data and people.

R8: Account for different stakeholders and emphasize the value of collaboration between departments.

We define two levels of abstraction of ProPerData - the first one is the overview level, which addresses all stakeholders equally and should foster communication about GDPR implementation projects, or analyzing and planning them. The second level is targeted at specific stakeholders, and therefore fulfills part of this requirement. To emphasize the value of collaboration, we

discuss practical insights in the work unit descriptions, which point at opportunities that arise from collaboration.

6.5. Summary of evaluation

In this section, we evaluated the design artifact of this thesis, ProPerData. We conducted eleven in-depth interviews with experts from the stakeholder groups enterprise architects, software developers and data protection experts. 29 industry experts participated in a quantitative survey, whereby the possible roles in ProPerData were selected between 2 and 18 times. An analytical evaluation with respect to modeling guidelines and the identified requirements for a process model to support GDPR compliance closes the evaluation.

The qualitative evaluation yielded rich insights about ProPerData from the perspectives of three stakeholder groups, which we described in detail. Enterprise architects largely identified with the way ProPerData presents DPM. Software developers welcomed the overview canvas, but would mostly use concrete instructions for their own tasks instead of the whole model. Data protection experts confirmed the correctness of ProPerData, but stated that it would provide the most value for non-experts in the field.

From the expert survey, we obtained mixed results. The support for the each of the four statements was quite strong, with averages of 4.1, 3.9, 4.3 and 3.5, yet this strong support did not translate to equally high ratings for the likelihood of recommending ProPerData to a colleague. In fact, more participants classified as detractors, resulting in a negative NPS. Therefore, there is no indication that ProPerData will become a widely adopted industry standard. Nevertheless, according to the generally positive tendency and the positive textual feedback, ProPerData is a valuable contribution to foster understanding and planning among the experts we surveyed.

Lastly, our discussion of the requirements that we defined in Chapter 4 lays out how the construction of ProPerData followed these requirements.

CHAPTER 7

Conclusion

In this final Chapter, we summarize the thesis chapters and recapitulate the contributions to the five research questions that we defined in Chapter 1 (Section 7.1). We present threats to the validity of our results, which include limitations of considered material (7.2.1), the perspective (7.2.2) the research method (7.2.3), and the evaluation approach 7.2.4). Finally, we reflect on this research project and discuss future research opportunities that this thesis motivates (7.3).

7.1. Summary

Chapter 1 outlines the reasoning for updating privacy legislation in Europe and the challenges that companies face when implementing the GDPR. We derived five research questions from the problem description and elaborated our research approach for developing and evaluating a reference process model for GDPR implementation projects.

To establish a common terminology, we discuss the most important GDPR definitions, the stakeholders in the regulation and in GDPR implementation projects, and the tasks that DPM encompasses in Chapter 2. Subsequently, we present empirical results on the challenges of DPM and on the suitability of EAM to support their execution.

In Chapter 3, we present approaches that have been published by the academic community in the IS field, as well as industry frameworks that support privacy, security and general management of IT assets. We discuss their suitability to support GDPR compliance and identify the research gap the we address with this thesis.

The reference model frame for our main artifact is constructed in Chapter 4. We first state and discuss the guiding requirements for the development of the reference process model. Then, we define the metamodel, key elements and modeling approach for our reference process model.

We discuss the construction of the reference process model ProPerData itself in Chapter 5. The construction approach is presented and exemplified. We characterize the roles based on interview results and define responsibilities for work units and the relationships between work units. Further, we explain the design decisions we made in the construction of ProPerData. The chapter closes with an analysis of the interrelationships of the DPM process that ProPerData embodies and other internal processes in an organization, with a particular focus on the EAM process.

Chapter 6 covers evaluation methods for reference models and presents our detailed evaluation results from a qualitative interview series with the three stakeholder groups *enterprise architects*, *software developers* and *data protection experts*. Further, we review and discuss the results from a quantitative survey and discuss the fulfillment of the requirements that we presented in the construction of the model frame.

Within this thesis, we addressed five research questions. In the following, we discuss our findings with respect to these research questions.

> ***RQ1:*** What are the tasks and stakeholders that have to be considered for GDPR compliance?

While the GDPR itself focuses on the description of the external roles data subject, data controller, data processor and supervisory authority (with the exception of the description of the DPO), we identify eight internal roles from expert interviews: DPM experts, process owners, application owners, data owners, software developers, enterprise architects, IT security and IT operations.

We identified eleven tasks that are encompassed by DPM: Inform & educate, verify existing processing activities, create new processing activities, conduct DPIA, cooperate with supervisory authority, maintain RoPA, conduct audits, interact with data subjects, report to management, execute organizational tasks, and leverage data protection efforts for business impact. From these tasks, DPM experts evaluated the tasks that address single processing activities as the most complex and the most time consuming.

Our results indicate that most DPM tasks suffer from a lack of clear guidelines and practical knowledge. Regarding single processing activities, identification of the right contact persons, the missing holistic view and the missing insight into single activities are the most severe problems.

To evaluate the suitability of EAM for supporting DPM tasks, we asked DPM experts to assess the usefulness of EAM for addressing DPM tasks. The group that had collaborated with EAM before indicated strong support for this assumption.

> ***RQ2:*** Which methods exist in literature to address GDPR compliance?

Our analysis of existing work to support compliance with the GDPR revealed only few holistic approaches, with the notable exception of the PRIPARE method by Crespo et al. (2015), the capability-based approach by Labadie and Legner (2019), and the method by Koç et al. (2018). Further work addressed single aspects of the GDPR, where the majority of contributions focuses on designing privacy-compliant information systems. Multiple publications point out the relationship between EAM and DPM.

While many consultancies advertise full methods for GDPR compliance, the only published industry approach that we identified was the Standard Data Protection Model (SDM). The SDM uses the seven protection goals of the GDPR to structure the protection activities and lists possible solutions for each of the activities. Additionally, we studied the industry frameworks ISO27001, COBIT and IT4IT for their support for data protection efforts.

Even though Koç et al. (2018) and Labadie and Legner (2019) derive their approaches from practitioner interviews, we identified a lack of practice-based insights into GDPR compliance approaches. Further, the SDM as a contribution from the German data protections authorities does not address the interaction of multiple stakeholders, which is an essential part of PRIPARE or COBIT. Lastly, the contributions do not clearly work out the temporal aspects of the work units that are necessary for GDPR compliance.

> ***RQ3:*** What are the requirements and concepts of a reference process model to address GDPR compliance?

We develop a list of eight requirements that our reference process model should fulfill in order to support the research goal. Four requirements address general aspects of reference models, two requirements address the specific goal of supporting GDPR compliance, and another two requirements originate from the opportunities and barriers that we identified during expert interviews. These last two requirements underline the fact that DPM is not an isolated effort, but affects the core processes and other supporting processes in an organization and must therefore take these relationships into account as well.

An established metamodel for privacy engineering provides the basis for the metamodel of our reference process model. This metamodel abstracts the elements of ProPerData, which in turn abstracts the observed company instances. We select a visual modeling approach at the overview level and rely on a combination of textual description and other modeling techniques in the detailed work unit descriptions.

> ***RQ4:*** How can a reference process model for GDPR compliance be defined?

The reference process model ProPerData is the result of a structured construction process that draws from a series of expert interviews with enterprise architects, input from data protection experts and software developers, and extensive literature sources.

ProPerData, which is presented in full in the Appendix B, includes eight roles, seven resources, sixteen work units, eleven temporal components, and twelve work products. Further elements, such as DPM tasks or the separation into internal and external work products, provide additional structure. The work unit descriptions shortly motivate each work unit and refer to possible solutions and additional sources.

> ***RQ5:*** How do practitioners assess the economic, deployment and engineering aspects of a reference process model for GDPR compliance?

To assess the implications of ProPerData, we conducted eleven qualitative interviews with three stakeholder groups: enterprise architects, software developers and DPM experts. We incorporated the feedback from the qualitative interviews into ProPerData. Subsequently, we conducted

a quantitative survey with 29 participants, which included all stakeholder roles. We present and discuss the evaluations of enterprise architects, software developers and DPM experts in detail.

From the perspective of enterprise architects, ProPerData could be used as a frame of reference in any type of organization, but lacks the specific instructions on how to initiate a GDPR compliance approach. Enterprise architects put a particular focus on the relationships between the elements within the organizational context. The stakeholder group assigned the highest ratings to ProPerData in the quantitative evaluation. This seems logical, because due to the selection of enterprise architects as the main source we expect the reference model to convey the thought processes of that group. Multiple comments confirmed this positive observation. Nonetheless, there was also some criticism from enterprise architects, which included the following points:

Level of detail The level of detail was viewed as insufficient to be transferred directly into practice. In an application scenario, all of the described work units would require multiple meetings and iterations.

Small stakeholder group An expert pointed to the very limited number of affected employees, even in a very large organization. We do not share this opinion. Even though different responsibilities exist, many stakeholders are involved in developing and running processing activities.

No method to initiate a company-specific data protection approach based on ProPerData This point of criticism referred to the wish of stakeholders to have a step by step procedure to follow. Indeed, creating a method to implement a data protection approach could be a next step, as we will discuss in Section 7.3. The identification of elements and relationships, which we consolidated in ProPerData, was a necessary step towards that goal.

Software developers generally welcomed the visual overview, as it fosters understanding of the own tasks in GDPR compliance approaches. They tended to see data protection requirements as a necessary burden, and therefore welcomed any approach that would provide quick insight into GDPR tasks and responsibilities. Software developers assigned high values for ProPerData's understandability, its ability to facilitate communication and its contribution to GDPR compliance, but indicated no conviction that it could provide economic benefits. Criticism by software developers included:

Too much information Software developers would like to have specific, short checklists that ensure complete fulfillment of data protection requirements. The overview perspective of ProPerData includes too much information for this stakeholder group.

Challenge to adhere to data protection guidelines in agile development Up-front planning of data protection measures clashes with an agile approach in software development. Interview partners proposed training and assigning responsibilities to software developers, which would require a more holistic perspective from developers (in contrast to the first point of criticism).

DPM experts also appreciated the holistic overview canvas of ProPerData, which could allow non-experts to quickly grasp the core concepts and responsibilities. The level of detail is adequate for this purpose, but does not cover the specifics of a data protection manager's duties. According

to the expert's observation, GDPR implementation knowledge is typically proprietary and offered by consulting companies. Apart from the SDM, there is no openly available approach to put the GDPR to work. This underlines the relevance of ProPerData. Criticism of DPM experts addressed:

Lack of instructions To make ProPerData easier to use, a step-by-step instruction should be added. As it is, stakeholders would not know where to start with a GDPR compliance approach.

Unfamiliar terminology Especially the technical and EA terminology deterred DPM experts from following up with ProPerData.

7.2. Limitations

The validity of this work depends on the validity of the considered material, the selection of the perspective, the construction and evaluation methods and the group of experts who participated in the evaluation. Therefore, we discuss each of the known limitations of this work in the following section.

7.2.1. Limitations of the considered material

The highly interdisciplinary nature of data protection - situated at the overlap among legal science, ISR, computer science and management science, as well as the multiplicity of topics that the GDPR addresses - from organizational responsibilities to technical implementation - make it difficult to ensure completeness for the sources that were included in this work. We executed partial searches, e.g. for data portability (Huth et al., 2019a), technical and organizational measures (Huth and Matthes, 2019), or the RoPA (Huth et al., 2019b). Additionally, we used forward and backward searches from these sources to identify relevant academic work in the field. Identification of relevant industry material occurred either through references in other publications or 'encouraged coincidences', i.e. newsletters and media articles covering the GDPR implementation. Another source were interview partners, e.g. for DIN 66398 about deletion concepts.

Our investigation and the resulting artifact are only based on the GDPR, which was released in 2016 and will continue to be in force in the coming years. In the latest IAPP annual privacy governance report 2020, less than half the respondents assessed the compliance status of their organization as 'compliant' or 'fully compliant'. The expandable implementation of the GDPR is also reflected in the large fines (cf. Section 1.3), so the topic of GDPR implementation is still as relevant as in 2018. Another reason the GDPR is considered such a landmark is that other legislation tends to follow existing rules, and organizations are safe if they aim for the strictest regulation, especially in an increasingly globalized world and increased international data transfers. Further, legislators in other countries are modeling their updated data protection legislation on the GDPR, e.g. in China, Canada, India, Singapore and the U.S. (IAPP, 2020, p.4).

7.2.2. Limitations in the perspective of ProPerData

The perspective that is presented in ProPerData is defined by the selection of the organizations and interview partners in the motivation, construction and evaluation phases of this research project.

The motivational survey maps out the perception of 38 DPOs. Ten participants worked for organizations with 1.000 to 10.000 employees, but there were no respondents from organizations with more than 10.000 employees. This threatens the validity of the hypothesis that EAM played an important role in the implementation of the GDPR. Since (a) the proportion of companies in our sample who did collaborate with EAM increases with the organization size and (b) almost all enterprise architects from large organizations in our interview series confirmed this collaboration, we nevertheless regard this assumption as valid.

29 enterprise architects represent the main source of information for the construction of ProPerData. These experts were distributed over all organization sizes from less than 5.000 employees to more than 50.000 employees. The focus emphasizes the internal perspective of enterprise architects, while the external perspective on the enterprise architects' work is less represented. To counter this threat to validity, we incorporated insights from more informal exchanges with data protection experts.

Another consequence of this focus is that the description of collaboration between other stakeholders is observational. We addressed this limitation by incorporating sources that do not make use of EA concepts, such as the SDM (Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder, 2019) or recommendations by the A29WP, and interviewing experts from software development and the data protection field.

All of the experts who contributed their expertise work in the German-speaking area. Even though the GDPR as a European regulation is the same for all countries, citizens of different countries have different perceptions towards privacy (European Commission, 2015). Certain opening clauses allow national legislators to detail on the requirements of the GDPR (although none of our interview partners mentioned the German legislation, the *Bundesdatenschutzgesetz*). Therefore, the implementation practices in other countries might differ. However, our close study of research publications across international conferences and journals did not indicate importance of the national adaptations of the GDPR in the member states. Since Germany has a long-standing tradition in data protection, we are confident that the perspective represented in ProPerData is adequate for other countries as well.

In the evaluation, we focused on enterprise architects, data protection managers and software developers. Eight out of 29 participants stated that they held more than one role, but we cannot exclude the possibility that other stakeholders (or another selection of experts with the same roles) would evaluate ProPerData differently.

7.2.3. Limitations in the method

The two central claims or characteristics of reference models are that they accurately describe an application domain on the one hand (*universality*), and that they provide a blueprint for de-

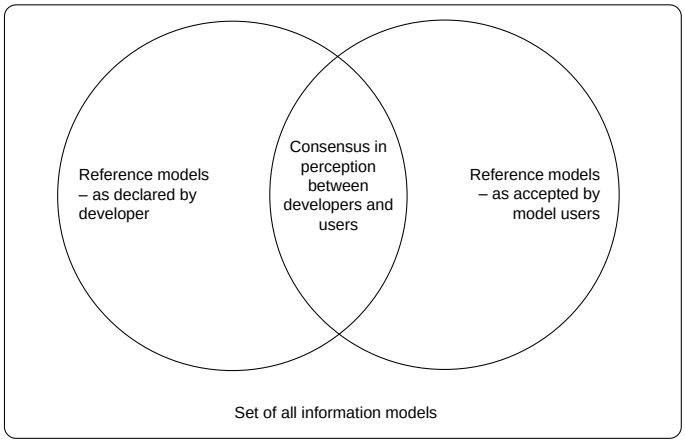

Figure 7.1.: Venn diagram of the perception of reference models. Adapted from Thomas (2005)

signing high-quality information systems (*recommendation character*) on the other hand (Frank, 2006, p.119; Vom Brocke, 2003, p.31). Both of these claims have been discussed critically in literature:

Universality Thomas (2005) notes that the universality attribute should not be mistaken as a claim for universal validity, but rather that the reference model may be regarded as universal if certain conditions are met. Examples for such conditions are a category of enterprises or a category of projects. Since we derived the reference process model from ex-post accounts given by enterprise architects, we can only reasonably make this claim for GDPR implementation approaches in large organizations with EA departments. Given some relaxations on terminology, we still argue that the model could be helpful for GDPR implementation approaches in other organizations as well.

Recommendation character The recommendation character of a reference model is inseparable from the question whether it is a good model for the chosen problem domain, because only then can it fulfil the promise of contributing to higher quality information systems. This leads to the challenge of evaluating the reference model, which we will discuss in Section 7.2.4.

Whether or not a model can be called a reference model can also not be determined objectively. The claim of an author that a model is in fact a reference model does not necessarily mean that others will equally accept it as a reference model (Thomas, 2005). As Figure 7.1 shows, an element within the set of all information models can only be called a reference model if it lies within the overlap of what both the developer and the user of a model define as a reference model. Thus, we can only make this claim and leave it to the model user to accept it as a reference model or not.

7.2.4. Limitations of the evaluation

Evaluating reference models is challenging due to methodological, philosophical and practical reasons (Schermann et al., 2007). As Frank (2006) puts forward, the claim for universality requires considering the variety of possible applications and possible objectives. In line with other researchers, Frank proposes a conceptual framework that structures the evaluation problem. The framework includes the four perspectives *engineering*, *deployment*, *economic* and *epistemological*, which the researcher may select for the evaluation. In this thesis, we do not investigate the *epistemological* perspective because of the practical orientation of ProPerData.

The evaluation took place in qualitative interviews and a quantitative survey. Although we addressed a (subjectively) meaningful selection of professionals, there can be no certainty that we achieved a representative perspective. Overall, 43 experts participated in the evaluation (14 in the qualitative part, 29 in the quantitative survey). The majority of the participants identified as enterprise architects, while only seven (2 qualitative / 5 quantitative) were data protection experts.

Another threat to the validity of this evaluation is that ProPerData was not evaluated in practice, but only through expert opinions. These opinions were formed from examination of the ProPerData material, which was made available before the evaluation, and the discussion with the researcher, which allowed clarification of doubts.

7.3. Reflection and outlook

This thesis motivates and describes the construction of a reference process model to support implementation approaches for the GDPR or, more generally, DPM. The resulting reference model ProPerData can serve as a blueprint for enterprise architects and data protection managers to support such approaches. Other stakeholders can equally use the ProPerData as a tool for understanding the big picture, but are more likely to limit the usage of ProPerData to the single work unit descriptions that are crucial for their role.

From the research perspective, this thesis constitutes an effort to address challenges from the discipline of data protection with the methods and perspectives of ISR. Thereby, we give an extensive account of existing work that aims to fill this gap in a similar way. In light of the number of 43 experts that participated in the evaluation, ProPerData represents a validated reference process model.

The introduction section shortly touched the motivation behind establishing new privacy laws - the clear distinction between the public space and the private space that balances societal and personal development. Whether the GDPR achieves this goal remains to be seen and is certainly in the eye of the beholder. In the author's view, one of the main merits of the GDPR is the management attention and the public awareness that it created. The first two and a half years have shown that supervisory authorities do not shy away from applying the unprecedented range of penalties, which makes it economically rational to follow the GDPR. The omnipresent cookie banners constantly remind data subjects of their rights. Considering that societal developments

rarely take place overnight, the GDPR may well have been a valuable stimulus in the right direction.

Reflecting on ProPerData, the stakeholder perspective appears to be a crucial factor for the assessment of such a reference process model. Enterprise architects approved of ProPerData and highlighted the value of models as structuring and communication tools. Software developers viewed the overall model as an interesting addition to the actual privacy-aware implementation work. Surprisingly, data protection experts seemed most hesitant towards ProPerData.

One key learning of this research endeavor is that the information, knowledge and tools for proper data protection already exist within the organizations. The challenge is therefore mainly to locate and exchange the information and collaborate across departments in large organizations.

As the previous section suggests, the limitations themselves provide opportunity to extend this work by adding more sources, including more perspectives, extending the approach to other types of regulation or conducting a practical evaluation of ProPerData. However, especially during the evaluation phase of this research project, three topics for possible future work emerged:

7.3.1. Design a method based on ProPerData

ProPerData is a reference process model that comprises the roles, resources, temporal units, work units and work products in a GDPR compliance effort. As a reference model, it serves two purposes (Frank, 2006): on the one hand, it is a description of its application domain, and on the other hand, it is a blueprint for a GDPR compliance project. However, it does not give step by step instructions on which work units to implement first. In the evaluation, multiple experts (e.g. E5, E8, E11) suggested creating a method based on ProPerData as a next logical step.

Koç et al. developed such a method in practice and published parts of it on creating the RoPA. We present interrelations between work units in Figure 5.6, which could serve as starting point in designing a method to instantiate or maintain GDPR compliance across the organization. Especially the aspect of maintaining compliance as business processes or business models evolve will gain more relevance in the future.

7.3.2. DPM as EA stakeholder

In various exchanges, we discovered that EA tool providers supported data protection concerns in their products, mainly functionalities to map IT assets to business processes and create the RoPA. Future academic work could conceptualize and generalize the existing methods that are implicitly implemented in current EA tools. As cross-organizational collaboration is one of the key challenges in many business domains, future work could focus on investigating how organizations can leverage the work for data protection for other business objectives.

Timm and Sandkuhl (2018) present EA viewpoints, i.e. perspectives on EA models, that address different (financial) compliance concerns, e.g. anti money-laundering aspects. This concept could be transferred to monitoring GDPR compliance, e.g. documentation status, last audit dates, security measures, or other aspects as proposed by Burmeister et al. (2019).

7.3.3. Ensuring data protection in agile organizations

Especially the developers in the evaluation referred to the challenge of planning data privacy and security measures before an agile development project and following through with this plan. In an emerging research contribution, we described the challenge to apply extensive development guidelines for development items in agile projects (Huth et al., 2020a). In a prototypical implementation, we proposed to attach granular data protection items, such as *ensure right to be forgotten*, to individual development projects. Developers could document the fulfillment or rejection of data protection items, and thereby gradually build an internal knowledge base of past solutions. Future work may focus on developing lightweight approaches to ensure data protection (or other non-functional requirements) in agile software development.

Bibliography

Syed Abdullah, Syed Norris Hikmi, Marta Indulska, and Shazia Sadiq. A study of compliance management in information systems research. In *17th European Conference on Information Systems, ECIS 2009*, 2009. ISBN 9788861293915.

Frederik Ahlemann and Heike Gastl. Process model for an empiracally grounded reference model construction. In *Reference modeling for business systems analysis*, pages 77–97. IGI Global, 2007.

Pouya Aleatrati Khosroshahi, Matheus Hauder, Alexander W Schneider, and Florian Matthes. Enterprise Architecture Management Pattern Catalog v2.0. Technical report, TU Munich, 2015.

Gonçalo Almeida Teixeira, Miguel Mira da Silva, and Ruben Pereira. The critical success factors of GDPR implementation: a systematic literature review. *Digital Policy, Regulation and Governance*, 2019. ISSN 2398-5038. doi: 10.1108/DPRG-01-2019-0007.

Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007.

Article 29 Data Protection Working Party. Anonymisation Techniques. Technical Report April, Article 29 WP, 2014.

Article 29 Data Protection Working Party. Guidelines on Data Protection Officers (DPOs), 2016.

Article 29 Data Protection Working Party. Guidelines on data protection impact assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of regulation 2016/679 (WP29), 2017a. ISSN 1556-5068.

Article 29 Data Protection Working Party. Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. Technical report, Article 29 WP, 2017b.

Article 29 Data Protection Working Party. Guidelines on the right to data portability. Technical Report April, Article 29 WP, 2017c.

Article 29 Data Protection Working Party. Guidelines on transparency under Regulation 2016/679. Technical report, Article 29 WP, 2017d.

Jef Ausloos and Pierre Dewitte. Shattering One-Way Mirrors. 2018.

Jef Ausloos, Rene Mahieu, and Michael Veale. Getting Data Subject Rights Right. 2019.

Vanessa Ayala-Rivera and Liliana Pasquale. The grace period has ended: An approach to operationalize GDPR requirements. In *Proceedings - 2018 IEEE 26th International Requirements Engineering Conference, RE 2018*, pages 136–146, 2018. ISBN 9781538674185. doi: 10.1109/RE.2018.00023.

J. Becker, R. Knackstedt, D. Kuropka, and P. Delfmann. Subjektivitätsmanagement für die Referenzmodellierung: Vorgehensmodell und Werkzeugkonzept [Subjectivity Management for Reference Modeling: Procedure Model and Tool Concept]. *KnowTech*, 49(0):1–19, 2001.

Kristian Beckers and Sebastian Pape. A Serious Game for Eliciting Social Engineering Security Requirements. In *IEEE 24th International Requirements Engineering Conference, RE*, pages 16–25, 2016. ISBN 9781509041213. doi: 10.1109/RE.2016.39.

France Bélanger and Robert E. Crossler. Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly: Management Information Systems*, 35(4):1017–1041, 2011. ISSN 02767783. doi: 10.2307/41409971.

Victoria Bellotti and Abigail Sellen. Design for Privacy in Ubiquitous Computing Environments. *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW '93*, pages 77–92, 1993. ISSN 20901232. doi: 10.1007/978-94-011-2094-4_6.

Martin Bichler, Ulrich Frank, David Avison, Julien Malaurent, Peter Fettke, Dirk Hovorka, Jan Krämer, Daniel Schnurr, Benjamin Müller, Leena Suhl, and Bernhard Thalheim. Theories in Business and Information Systems Engineering. *Business and Information Systems Engineering*, 58(4):291–319, 2016. ISSN 18670202. doi: 10.1007/s12599-016-0439-z.

Felix Bieker, Michael Friedewald, Marit Hansen, Hannah Obersteller, and Martin Rost. A Process for Data Protection Impact Assessment Under the European General Data Protection Regulation. In *Proceedings - 4th Annual Privacy Forum 2016*, pages 21–37, Cham, 2016. Springer. ISBN 978-3-319-44760-5. doi: 10.1007/978-3-319-44760-5_2.

Michael Brenner, Nils Gentschen Felde, Wolfgang Hommel, Stefan Metzger, Helmut Reiser, and Thomas Schaaf. *Praxisbuch ISO/IEC 27001*. Hanser, Munich, 2011. ISBN 978-3-446-43026-6.

Sabine Buckl. *Developing organization-specific enterprise architecture management functions using a method base*. PhD thesis, Technical University of Munich, 2011. URL `https://mediatum.ub.tum.de/doc/1069959/1069959.pdf`.

Fabian Burmeister, Paul Drews, and Ingrid Schirmer. A Privacy-driven Enterprise Architecture Meta-Model for Supporting Compliance with the General Data Protection Regulation. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, pages 6052–6061, 2019. ISBN 9780998133126.

Fabian Burmeister, Dominik Huth, Ingrid Schirmer, Paul Drews, and Florian Matthes. Enhancing Information Governance with Enterprise Architecture Management : Design Principles Derived from Benefits and Barriers in the GDPR Implementation. In *53rd Hawaii International Conference on Systems Sciences*, pages 5593–5602, 2020.

Julio C. Caiza, Yod Samuel Martín, Danny S. Guamán, Jose M. Del Alamo, and Juan C. Yelmo. Reusable Elements for the Systematic Design of Privacy-Friendly Information Systems: A Mapping Study. *IEEE Access*, 7:66512–66535, 2019. ISSN 21693536. doi: 10.1109/ACCESS.2019.2918003.

Ann Cavoukian and Mark Dixon. Privacy and Security by Design: An Enterprise Architecture Approach. Technical Report September, 2013.

CIPL. Organisational Readiness for the European Union General Data Protection Regulation. Technical Report March, CIPL, 2018.

Anne Cleven and Robert Winter. Regulatory compliance in information systems research - Literature analysis and research agenda. *Lecture Notes in Business Information Processing*, 29 LNBIP:174–186, 2009. ISSN 18651348. doi: 10.1007/978-3-642-01862-6_15.

CMS. GDPR Enforcement Tracker, 2021. URL `http://www.enforcementtracker.com/`. Last accessed: 2021-01-01.

Michael Colesky, Jaap-Henk Hoepman, and Christiaan Hillen. A Critical Analysis of Privacy Design Strategies. In *Proceedings - 2016 IEEE Symposium on Security and Privacy Workshops, SPW 2016*, pages 33–40, 2016. ISBN 9781509008247. doi: 10.1109/SPW.2016.23.

Alberto Crespo, Nicolas Notario, Carmela Troncoso, Daniel Le Métayer, Inga Kroener, David Wright, Jose M Del Alamo, and Yod Samuel Martin. PRIPARE Privacy- and Security-by-Design Methodology Handbook. Technical report, Pripare Project, 2015.

George Danezis, Josep Domingo-Ferrer, Marit Hansen, Jaap-Henk Hoepman, Daniel Le Métayer, Rodica Tirtea, and Stefan Schiffner. Privacy and Data Protection by Design. Technical Report December, ENISA, 2014.

Steven De Haes and Wim Van Grembergen. IT governance and its mechanisms. *Information Systems Control Journal*, 1:27–33, 2004.

Deutsche Datenschutzkonferenz. Template for the Record of Processing Activities pursuant to Art. 30, 2018. URL `https://www.datenschutzkonferenz-online.de/media/ah/201802_ah_muster_verantwortliche.pdf`. Last accessed: 02/22/2020.

Deutscher Bundestag. Internetnutzung Globale Entwicklung und Darstellung empirischer Daten. 2007.

Economist. The world's most valuable resource is no longer oil, but data. *The Economist*, 2017. URL `https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data`. last accessed: 02-13-2021.

European Commission. Eurobarometer 431 data protection. Technical report, 2015.

European Commission. What is personal data?, 2020. URL https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en.

European Data Protection Supervisor. Guidelines on the Rights of Individuals with regard to the Processing of Personal Data. Technical report, EDPS, 2010.

European Parliament. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995. ISSN 1098-6596.

European Union. Regulation 2016/679 of the European parliament and the Council of the European Union, 2016. ISSN 1977-0677. URL https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02016R0679-20160504.

Peter Fettke and Peter Loos. Classification of reference models: a methodology and its application. *Information Systems and e-Business Management*, 1(1):35–53, 2003. ISSN 1617-9846. doi: 10.1007/BF02683509.

Peter Fettke and Peter Loos. Referenzmodellierungsforschung. *Wirtschaftsinformatik*, 46(5): 331–340, 2004. ISSN 09376429. doi: 10.1007/BF03250947.

Peter Fettke, Peter Loos, and Jörg Zwicker. Business process reference models: Survey and classification. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3812 LNCS:469–483, 2005. ISSN 03029743. doi: 10.1007/11678564_44.

Ralph Foorthuis, Frank Hofman, Sjaak Brinkkemper, and Rik Bos. Assessing business and IT projects on compliance with Enterprise Architecture. *CEUR Workshop Proceedings*, 459:1–15, 2009. ISSN 16130073.

Ulrich Frank. Evaluation of reference models. *Reference Modeling for Business Systems Analysis*, pages 118–140, 2006. doi: 10.4018/978-1-59904-054-7.ch006.

Ulrich Frank, Stefan Strecker, and Stefan Koch. Open Model: Ein Vorschlag für ein Forschungsprogramm der Wirtschaftsinformatik. *eOrganisation: Service-, Prozess-, Market-Engineering, 8. Internationale Tagung Wirtschaftsinformatik*, 2:217–234, 2007.

Maria da Conceição Freitas and Miguel Mira da Silva. GDPR Compliance in SMEs: There is much to be done. *Journal of Information Systems Engineering & Management*, 3(4):1–7, 2018. ISSN 2468-4376. doi: 10.20897/jisem/3941.

Andreas Gadatsch. *Grundkurs Geschäftsprozess-Management*. Springer Fachmedien Wiesbaden, 2017. ISBN 9783658171780. doi: 10.1007/978-3-658-17179-7.

Markus Gaulke. *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.

Seda Gürses and Jose M. Del Alamo. Privacy Engineering: Shaping an Emerging Field of Research and Practice. *IEEE Security and Privacy*, 14(2):40–46, 2016. ISSN 15584046. doi: 10.1109/MSP.2016.37.

Irit Hadar, Tomer Hasson, Oshrat Ayalon, Eran Toch, Michael Birnhack, Sofia Sherman, and Arod Balissa. Privacy by designers: software developers' privacy mindset. *Empirical Software Engineering*, 23(1):259–289, 2018. ISSN 15737616. doi: 10.1007/s10664-017-9517-1.

Ben Halpert. *Auditing Cloud Computing*. Wiley Online Library, 2011.

Volker Hammer. DIN 66398: Die Leitlinie Löschkonzept als Norm. *Datenschutz und Datensicherheit - DuD*, 40(8):528–533, 2016. ISSN 1614-0702. doi: 10.1007/s11623-016-0651-5.

Matheus Hauder, Sascha Roth, Florian Matthes, and Christopher Schulz. Organizational factors influencing enterprise architecture management challenges. *ECIS 2013 - Proceedings of the 21st European Conference on Information Systems*, 2013.

Matheus Hauder, Sascha Roth, Christopher Schulz, and Florian Matthes. Agile enterprise architecture management an analysis on the application of agile principles. In *BMSD 2014 - Proceedings of the 4th International Symposium on Business Modeling and Software Design*, pages 38–46, 2014. ISBN 9789897580321. doi: 10.5220/0005424100380046.

Maurice Hendrix, Ali Al-Sherbaz, and Victoria Bloom. Game Based Cyber Security Training: are Serious Games suitable for cyber security training? *International Journal of Serious Games*, 3(1):53–61, 2016. ISSN 2384-8766. doi: 10.17083/ijsg.v3i1.107.

Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design science in information systems research. *MIS Quarterly*, 28(1):75–105, 2004.

Mike Hintze. Privacy Statements Under the GDPR. *Seattle U. L. Rev*, 42:1129–1154, 2018.

Jaap-Henk Hoepman. Privacy Design Strategies. In *IFIP International Information Security Conference*, pages 446–459, Berlin, Heidelberg, 2014. Springer. ISBN 978-3-642-55414-8. doi: 10.1007/978-3-642-55415-5. URL `http://arxiv.org/abs/1210.6621`.

Dominik Huth. A Pattern Catalog for GDPR Compliant Data Protection. In *IFIP Working Conference on The Practice, Doctoral Consortium*, Leuven, 2017.

Dominik Huth and Florian Matthes. "Appropriate Technical and Organizational Measures": Identifying Privacy Engineering Approaches to Meet GDPR Requirements. In *25th Americas Conference on Information Systems*, Cancún, 2019.

Dominik Huth and Florian Matthes. ProPerData - A process model to support GDPR compliance. Technical Report March, Technical University of Munich, Munich, 2020.

Dominik Huth, Anne Faber, and Florian Matthes. Towards an Understanding of Stakeholders and Dependencies in the EU GDPR. In Paul Drews, Burkhardt Funk, Peter Niemeyer, and Lin Xie, editors, *Multikonferenz Wirtschaftsinformatik*, pages 338–344, Lüneburg, 2018.

Dominik Huth, Laura Stojko, and Florian Matthes. A Service Definition for Data Portability. In *21st International Conference on Enterprise Information Systems*, pages 169–176, 2019a.

Dominik Huth, Ahmet Tanakol, and Florian Matthes. Using Enterprise Architecture Models for Creating the Record of Processing Activities (Art . 30 GDPR). In *23rd IEEE International Distributed Object Computing Conference (EDOC)*, pages 98–104, Paris, 2019b. doi: DOI10.1109/EDOC.2019.00021.

Dominik Huth, Andreas Both, Jeffrey Ahmad, Gerhard Sauer, Fatih Yilmaz, and Florian Matthes. Process and Tool Support for Integration of Privacy Aspects in Agile Software Engineering. In *Americas Conference on Information Systems 2020*, pages 1–5, Virtual, 2020a.

Dominik Huth, Fabian Burmeister, Florian Matthes, and Ingrid Schirmer. Empirical Results on the Collaboration Between Enterprise Architecture and Data Protection Management during the Implementation of the GDPR. In *53rd Hawaii International Conference on Systems Sciences*, pages 5839–5848, 2020b.

Dominik Huth, Michael Vilser, Gloria Bondel, and Florian Matthes. Empirical Task Analysis of Data Protection Management and its Collaboration with Enterprise Architecture Management. In *22nd International Conference on Enterprise Information Systems*, Prague, 2020c. to appear.

IAPP. IAPP-EY Annual Privacy Governance Report 2019. Technical report, EY, 2019a.

IAPP. 2019 Privacy Tech Vendor Report. Technical report, International Association of Privacy Professionals, 2019b.

IAPP. 2020 Privacy Tech Vendor Report 4.1. Technical report, 2020.

IAPP. IAPP-FTI Consulting Privacy Governance Report 2020. Technical report, 2020.

IAPP and TrustArc. Measuring Privacy Operations. Technical report, 2019.

UK ICO. A guide to ICO audits. 2018. URL `https://ico.org.uk/media/for-organisations/documents/2787/guide-to-data-protection-audits.pdf`.

ISO. International Standard ISO/IEC/IEEE 42010:2011 Systems and software engineering — Architecture Description, 2011. ISSN 2167082X.

ISO. ISO/IEC 24744:2007. Software engineering — metamodel for development methodologies., 2014.

ISO. ISO/IEC 29134:2017 Information technology — Security techniques — Guidelines for privacy impact assessment, 2017.

ISO. ISO27001:2018 Information technology — Security techniques — Information security management systems, 2018.

ISO. ISO27701:2019 Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines, 2019.

Dierk Jugel, Christian M. Schweda, Christina Bauer, Jawed Zamani, and Alfred Zimmermann. A metamodel to integrate control objectives into viewpoints for EA management. *CEUR Workshop Proceedings*, 2218:110–119, 2018. ISSN 16130073.

Ilya Kabanov. Effective frameworks for delivering compliance with personal data privacy regulatory requirements. In *2016 14th Annual Conference on Privacy, Security and Trust, PST 2016*, pages 551–554, 2016. ISBN 9781509043798. doi: 10.1109/PST.2016.7907015.

Christos Kalloniatis, Evangelia Kavakli, and Stefanos Gritzalis. Addressing privacy requirements in system design: The PriS method. *Requirements Engineering*, 13(3):241–255, 2008. ISSN 09473602. doi: 10.1007/s00766-008-0067-3.

Sebastian Klipper. *Information security risk management*. Springer Fachmedien, Wiesbaden, 2015. ISBN 978-3-658-08773-9. doi: 10.1007/978-3-658-08774-6.

Hasan Koç, Kai Eckert, and Daniel Flaig. Datenschutzgrundverordnung (DSGVO): Bewältigung der Herausforderungen mit Unternehmensarchitekturmanagement (EAM) - Challenging the General Data Protection Regulation (GDPR) with Enterprise Architecture Management (EAM). *HMD Praxis der Wirtschaftsinformatik*, 55(5):942–963, 2018. ISSN 1436-3011. doi: 10.1365/s40702-018-00449-7.

Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder. Das Standard-Datenschutzmodell 2.0 (The standard data protection model) 2.0. Technical report, DSK, 2019.

Christian Kurtz, Martin Semmann, and Tilo Böhmann. Privacy by design to comply with GDPR: A review on third-party data processors. *Americas Conference on Information Systems 2018: Digital Disruption, AMCIS 2018*, pages 1–10, 2018.

Clément Labadie and Christine Legner. Understanding Data Protection Regulations from a Data Management Perspective : A Capability-Based Approach to EU-GDPR. *14th International Conference on Wirtschaftsinformatik, February 24-27*, 2019.

Jorg Lenhard, Lothar Fritsch, and Sebastian Herold. A literature study on privacy patterns research. In *Proceedings - 43rd Euromicro Conference on Software Engineering and Advanced Applications, SEAA 2017*, pages 194–201, 2017. ISBN 9781538621400. doi: 10.1109/SEAA.2017.28.

Yod Samuel Martin and Jose M Del Alamo. A metamodel for privacy engineering methods. In *CEUR Workshop Proceedings*, pages 41–48, 2017a.

Yod Samuel Martin and Jose M Del Alamo. A metamodel for privacy engineering methods. *CEUR Workshop Proceedings*, 1873(731711):41–48, 2017b. ISSN 16130073.

Philipp Mayring. Qualitative Content Analysis. *Forum: Qualitative Social Research*, 1(2):105–114, 2000. ISSN 1438-5627. doi: 10.17169/fqs-1.2.1089.

A. McDonald and Lorrie Faith Cranor. The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society*, 4(3):543 – 568, 2008.

Daniel Mikkelsen and Malin Strandell-jansson. GDPR compliance after May 2018 : A continuing challenge. Technical Report April, McKinsey & Company, 2018.

Jacques-alain Miller. Jeremy Bentham' s Panoptic Device. *MIT Press*, 41(Summer):3–29, 1987.

Miniwatts Marketing. Internet World Stats, 2020. URL `https://internetworldstats.com/stats.htm`. Last accessed: 2020-10-01.

Nicolas Notario, Alberto Crespo, Yod Samuel Martin, Jose M Del Alamo, Daniel Le Metayer, Thibaud Antignac, Antonio Kung, Inga Kroener, and David Wright. PRIPARE: Integrating privacy best practices into a privacy engineering methodology. In *Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015*, pages 151–158, 2015. ISBN 9781479999330. doi: 10.1109/SPW.2015.22.

Hubert Österle, Jörg Becker, Ulrich Frank, Thomas Hess, Dimitris Karagiannis, Helmut Krcmar, Peter Loos, Peter Mertens, Andreas Oberweis, and Elmar J Sinz. Memorandum on design-oriented information systems research. *European Journal of Information Systems*, 20(1):7–10, jan 2011. ISSN 0960-085X. doi: 10.1057/ejis.2010.55.

Ken Peffers, Tuure Tuunanen, Marcus A Rothenberger, and Samir Chatterjee. A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3):45–77, dec 2007. ISSN 0742-1222. doi: 10.2753/MIS0742-1222240302.

Frederick F Reichheld. The One Number You Need to Grow. *Harvard Business Review*, 81(12): 46–54+124, 2003. ISSN 00178012.

Hannah Ritchie and Max Roser. Technology Adoption in US households, 2019. URL `https://ourworldindata.org/`. last accessed: 02-13-2021.

Daniel Rösch, Thomas Schuster, Lukas Waidelich, and Sascha Alpers. Privacy control patterns for compliant application of GDPR. *25th Americas Conference on Information Systems, AMCIS 2019*, pages 1–10, 2019.

Petr Rozehnal and Vitezslav Novak. The Core of Enterprise Architecture as a Management Tool: GDPR Implementation Case Study. In *26th Interdisciplinary Information Management Talks*, pages 359–366, Kutná Hora, Czech Republic, 2018. Trauner Verlag.

Johnny Saldaña. *The Coding Manual for Qualitative Researchers.* Sage, 2013. ISBN 9781446247365. doi: 10.1017/CBO9781107415324.004.

Florian Schaub, Rebecca Balebako, and Lorrie Faith Cranor. Designing Effective Privacy Notices and Controls. *IEEE Internet Computing*, 21(3):70–77, 2017. ISSN 10897801. doi: 10.1109/MIC.2017.75.

Michael Schermann, Tilo Böhmann, and Helmut Krcmar. Fostering the Evaluation of Reference Models : Application and Extension of the Concept of IS Design Theories. *eOrganisation: Service-, Prozess-, Market-Engineering : 8. Internationale Tagung Wirtschaftsinformatik 2007*, pages 181–198, 2007.

Bruce Schneier. *Data and Goliath: The hidden battles to collect your data and control your world.* WW Norton & Company, 2015. ISBN 0393244822.

Reinhard Schütte. *Grundsätze ordnungsmäßiger Referenzmodellierung.* Springer Fachmedien Wiesbaden, 1998. ISBN 9783409128438.

Sean Sirur, Jason R. C. Nurse, and Helena Webb. 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*, pages 88–95. ACM, 2018.

Daniel J Solove. A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154(3):477, 2006. ISSN 00419907. doi: 10.2307/40041279.

Daniel J Solove. I've got nothing to hide and other misunderstandings of privacy. *San Diego l. Rev.*, 44:745, 2007.

Dirk Steuperaert. COBIT 2019: a Significant Update. *Edpacs*, 59(1):14–18, 2019. ISSN 19361009. doi: 10.1080/07366981.2019.1578474.

Latanya Sweeney. k-Anonymity: A model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5):557–570, 2002.

Andreas Taschner. *Management Reporting und Behavioral Accounting.* Springer Gabler, Wiesbaden, 2015. ISBN 978-3-658-23492-8. doi: 10.1007/978-3-658-23492-8.

The Open Group. The Open Group IT4IT ™ Reference Architecture , Version 2.1, 2017.

R. Oliver Thomas. Understanding the term reference model in information systems research: History, literature analysis and explanation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3812 LNCS:484–496, 2005. ISSN 03029743. doi: 10.1007/11678564_45.

Christina Tikkinen-Piri, Anna Rohunen, and Jouni Markkula. EU General Data Protection Regulation: Changes and implications for personal data collecting companies. *Computer Law and Security Review*, 2017. ISSN 02673649. doi: 10.1016/j.clsr.2017.05.015.

Felix Timm and Kurt Sandkuhl. A reference enterprise architecture for holistic compliance management in the financial sector. In *International Conference on Information Systems 2018, ICIS 2018*, pages 1–17, 2018. ISBN 9780996683173.

TrustArc. GDPR Compliance Status. Technical report, 2018.

UC Berkeley School of Information. privacypatterns.org, 2020. URL `https://privacypatterns.org/`. Last accessed: 2020-07-31.

M. W. Van Roosmalen and S. J.B.A. Hoppenbrouwers. Supporting corporate governance with enterprise architecture and business rule management: A synthesis of stability and agility. In *CEUR Workshop Proceedings*, volume 342, pages 13–24, 2008.

Aysem Diker Vanberg and Mehmet Bilal Ünver. The right to data portability in the GDPR and EU competition law: odd couple or dynamic duo? *European Journal of Law and Technology*, 8(1):1–22, 2017. doi: 10.2139/ssrn.2216088.

Michael Vilser. Empirical Task Analysis of Data Protection Management, 2019. URL `https://wwwmatthes.in.tum.de/pages/12wb2907mhi4k/Bachelor-s-Thesis-Michael-Vilser`. Bachelor's Thesis.

Paul Voigt and Axel von dem Bussche. *EU-Datenschutz-Grundverordnung (DSGVO)*. 2018. ISBN 9783662561867. doi: 10.1007/978-3-662-56187-4.

Jan Vom Brocke. *Referenzmodellierung: Gestaltung und Verteilung von Konstruktionsprozessen*. Logos Verlag, Berlin, 2003. ISBN 3-8325-0179-7.

Willis Ware. Records, Computers, and the Rights of Citizens: Report. Technical report, Department of Health, Education, and Welfare. Secretary's Advisory Committee on Automated Personal Data Systems, 1973.

Samuel Warren and Louis Brandeis. The Right to Privacy. *Harvard Law Review*, 4(5):193–220, 1890.

Maike Weiß and Kathrin Strauß. DSGVO-konformer Unternehmens-Datenschutz durch ISO 27001 und ISO 27701?, 2019. URL `https://www.datenschutzexperte.de/blog/datenschutz-im-unternehmen/iso-27001-und-iso-27701/`. Last accessed: 2020-07-31.

Alan F. Westin. *Privacy and Freedom*. Atheneum Press, New York, 1967. ISBN 0370013255. doi: 10.2307/2092293.

Johannes Wichmann, Kurt Sandkuhl, Nikolay Shilov, Alexander Smirnov, Felix Timm, and Matthias Wißotzki. Enterprise Architecture Frameworks as Support for Implementation of Regulations: Approach and Experiences from GDPR. *Complex Systems Informatics and Modeling Quarterly*, (24):31–48, 2020. ISSN 2255-9922. doi: 10.7250/csimq.2020-24.03.

Katharina Winter, Sabine Buckl, Florian Matthes, and Christian M Schweda. Investigating the State-of-the-Art in Enterprise Architecture Management Methods in Literature and Practice. In *Proceedings of the Mediterranean Conference on Information Systems*, 2010.

Robert Winter and Ronny Fischer. Essential layers, artifacts, and dependencies of enterprise architecture. *Journal of Enterprise Architecture*, (May):1–12, 2007. doi: 10.1109/EDOCW.2006.33.

Shoshana Zuboff. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books, 2019.

Abbreviations

A29WP Article 29 Working Party

BI business intelligence

CMDB configuration management database

DPIA data protection impact assessment

DPM data protection management

DPMS data protection management system

DPO data protection officer

GoM guidelines of modeling

EA enterprise architecture

EAM enterprise architecture management

EDPB European Data Protection Board

EU European Union

GDPR General Data Protection Regulation

IAPP International Association of Privacy Professionals

IS information system

ISACA Information Systems Audit and Control Association

ISR information systems research

ISMS information security management system

MPEM Metamodel for Privacy Engineering Methods

NPS net promoter score

PbD Privacy by Design

PDCA plan-do-check-act

RoPA record of processing activities

SDM Standard Data Protection Model

SOX Sarbanes Oxley Act

TOM technical and organizational measures

UML Unified Modeling Language

APPENDIX A

Interview partners

ID	Position	Industry	Company size
A01	Enterprise Architect	Logistics	5000 - 15000
A02	Business Architect	Insurance	<5000
A03	Lead IT Strategy & Architecture	Government	15001 - 50000
A04	Lead Enterprise Architect	Automotive	>50000
A05	Lead Enterprise Architect	Professional Services	5000 - 15000
A06	Enterprise Architect	Insurance	5000 - 15000
A07	Lead Enterprise Architect	Manufacturing	15001 - 50000
A08	Enterprise Architect	Insurance	15001 - 50000
A09	Lead Enterprise Architect	Industrial Services	5000 - 15000
A10	Enterprise Architect	Insurance	5000 - 15000
A11	Enterprise Architect	IT Services	<5000
A12	Enterprise Architect	Consumer Goods	15001 - 50000
A13	Lead Enterprise Architect	IT Services	15001 - 50000
A14	Enterprise Architect, EA Lead	Banking	15001 - 50000
A15	Chief IT Architect	Insurance	<5000
A16	Enterprise Architect (2)	Automotive	>50000
A17	Enterprise Architect	Banking	<5000
A18	Enterprise Architect	Logistics	15001 - 50000
A19	IT Architect	Banking (CH)	5000 - 15000
A20	Lead IT Strategy & Architecture	Sports	<5000
A21	IT Solution Architect	IT Services	>50000
A22	Enterprise Architect	Automotive	>50000
A23	Enterprise Architect (4)	Insurance	5000 - 15000
A24	IT Architect	IT Services	<5000

Table A.1.: Interview partners from interview series in (Huth et al., 2020b) and (Burmeister et al., 2020)

ID	Position	Industry	Company size
B1	Cyber Security Portfolio Manager	Industrial Manufacturing	large enterprise
B2	Head of IT Strategy	Industrial Manufacturing	large enterprise
B3	Co-Founder Compliance Tool	IT Service & Consulting	Start-Up
B4	Corporate Data Privacy Officer	Industrial Manufacturing	large enterprise
B5	Cyber Security Architect	Industrial Manufacturing	large enterprise
B6	Head of Sales for Privacy Tech	IT Service & Consulting	SME
B7	Business Intelligence (2)	Finance	large enterprise

Table A.2.: Expert sources with focus on data portability (Huth et al., 2019a)

ID	Position	Industry	Company size
C1	Lawyer	undisclosed	undisclosed
C2	Data protection expert	undisclosed	undisclosed
C3	Data protection officer	undisclosed	undisclosed
C4	Data protection officer	undisclosed	undisclosed
C5	Focus group interview with 4 DPM experts	undisclosed	undisclosed

Table A.3.: Expert interviews with focus on the record of processing activities (Huth et al., 2019b)

APPENDIX B

ProPerData - A process model for GDPR compliance

In Chapter 5, we described our search process and the design decisions that led to ProPerData. We explained the overview canvas and shortly referred to the ProPerData elements, i.e. the roles, resources, work units, stages, and work products. This appendix is intended as a complete and consistent presentation of ProPerData that is separated from the research details in the thesis. It should serve as a reference for exploring and using the process model.

B.1. Roles

A key reason why the GDPR is so complex is its interdisciplinary nature. In this section, we present the organizational roles that are described in ProPerData.

R-1 Data protection management: The team or role that is responsible for conducting and coordinating the overall data protection efforts of the organization. The team is headed by the data protection officer (DPO), a stakeholder explicitly mentioned in the GDPR (Huth et al., 2020c).

R-2 Process owner: The person from the business department who is responsible for a business process and the processing activity. The process owner defines why a business process / a processing activity is conducted. Note that this relates to the definition of the controller as the entity who *"determines the purposes and means of the processing of personal data"* (Huth et al., 2018).

R-3 Application owner: The person who is responsible for a single application, i.e. who coordinates the operation and maintenance of an application. In some cases, application owner and product owner can be the same person.

R-4 Data owner: The responsible person for a data object. This is important for master data that is accessed by multiple applications and used in multiple processing activities within the organization. The data owner knows which processing activities use a set of personally identifiable information (PII, e.g. address) and gives permission to use or change that data. Consequently, the data owner is also the contact person if PII should be deleted.

R-5 Software developer: The person who translates existing business requirements into executable code. We do not distinguish between software architects and programmers.

R-6 IT security: The IT security department has the objective of ensuring the attributes confidentiality, integrity and availability of the applications in an organization (the "security triad").

R-7 Enterprise architect: EA management has the goal of strategically developing the enterprise architecture, consisting of people, processes, applications, and their interrelationships. To this end, various elements of the architecture are documented, with applications as the most common element.

R-8 IT operations: The IT department that is responsible for the technical operation of an application, i.e. hosting and virtualization.

B.2. Stages

There are three distinct stages in ProPerData: The initial setup for introducing a new regulatory framework, the ongoing operation under the existing regulatory framework, and a review phase for improvement and adaptation of the applied process framework.

In the *initial setup* stage, the driving forces for action are the external pressures that originate from the new or changed set of rules. Hence, the goal in the *initial setup* stage is to *achieve regulatory compliance*. In the case of the GDPR, the rules comprise the new documentation obligations, the need for processing agreements, and the obligation to execute the new data subject rights. Following the GDPR timeline, the time frame for the initial setup phase was the period from when the regulation was passed in 2016 to when it became effective in 2018. However, industry reports suggest that some companies are still in the process of adapting to the new regulation (TrustArc, 2018). It is therefore an essential part of ProPerData despite the release after the GDPR deadline.

Once the initial GDPR compliance measures are in place, the driving forces are not changes in the regulatory framework, but in the underlying organization. Consequently, the goal in the *operation* stage is *maintaining regulatory compliance* despite these changes. Within the operation stage, there are various other time frame descriptors:

- Ongoing/continuous operation
- RoPA update cycles
- Audit cycles

Following the *software engineering metamodel for development methods (SEMDM)* and the *metamodel for privacy engineering methods (MPEM)* (Martin and Del Alamo, 2017a), we define cycles and events[1] in the *operation* stage. The *operation* stage itself groups processes that take place on an ongoing basis, without a particular cyclical or event-based trigger. The two cycles we identified are (1) the RoPA update cycles (typically one year, cf. (Huth et al., 2019b)) that are determined internally, and (2) the audit cycles for external auditors. Regarding events, we define the five events:

- Data subject request: A request that is based on GDPR Articles 13-22.
- New process: Establishment of a new processing activity that originates, transfers or processes personal data.
- Changed process: Changes to a processing activity that affect personal data, e.g. collecting data for analytical purposes.
- Data breach: Gaining knowledge that personal data has been accessible by unauthorized individuals, either internally or publicly.
- Decommissioning: Discontinuing a processing activity.

During the *review* stage, the driving force for acting is to improve the ongoing regulatory compliance measures by reflecting and adapting.

B.3. Resources

Resources, according to Caiza et al. (2019), are reusable elements that are assumed to exist and can be used "as is" for attaining a set goal. Among them are general concepts, such as language or notation, which we do not describe here. Caiza et al. (2019) and Martin and Del Alamo (2017a) also include privacy conceptual models ("what is privacy") and privacy normative frameworks ("how should the concept of privacy be enforced?") in the MPEM. Of course, our privacy normative framework is the GDPR itself. We choose to restrict this practice-driven publication of ProPerData to the resources that are specific to enacting the tasks of ProPerData.

The resources we list in this section do not necessarily exist in all organizations explicitly, but we believe they apply in any kind of organization that processes personal data: Business processes, applications, deployments, data objects and data flows might not always be documented, but are useful mental concepts for fulfilling the tasks that we specify in ProPerData.

KB-1 Business processes

A business process that processes personal data matches the concept of a processing activity in the sense of the GDPR. Business processes are either documented or exist as implicit knowledge of the stakeholders. Examples for explicit documentation could be dedicated process repositories

[1] SEMDM and MPEM define milestones rather than events. We adapted the notion to the organizational scope of ProPerData.

or the EA business process documentation. Yet, the business process documentation has not been used extensively in GDPR endeavors. Our interview partners reported that these repositories are often incomplete and that only selected processes are modeled. Thus, the information in a business process documentation should be handled carefully.

KB-2 Deployments

To identify processing activities that are supported by IT systems, a Configuration Management Database (CMDB) is a detailed technical documentation of the application hosts from an operational perspective. Since a record in the CMDB is mandatory in many organizations in order to have an application hosted centrally, it has been used in many cases as the entry point to identify relevant applications. However, the technical documentation lacks the meta-information about applications that is necessary to understand the nature of data processing and if processing of personal data is involved.

KB-3 Applications and software

Application repositories or application lists represent the information requirements for applications from an enterprise architecture perspective. The information includes the business domain or business capability that the application supports, the type of processing, the used technologies, and the application owner. Even though applications do not match directly to processing activities in the sense of the GDPR, they implicitly hint at the business process they support. All (non-consulting) enterprise architects in our study reported having a satisfactory level of completeness in the application repository. Thus, enterprise architects and DPM experts alike agreed on the usefulness of this resource.

KB-4 Data transfers

Modeled data flows between services illustrate which services exchange which type of information. Enterprise architects oblige service or application owners to register in a service repository in order to gain access to central (data) services, such as customer master data. This central repository then serves as a gatekeeper for compliance with the prerequisites for registration, for instance adherence to the GDPR processing principles. A central service repository also allows identifying the data that is exchanged via the interface and creating logs of that exchange. Beside the benefit for identification, the logs can serve as a forensic tool in case of a data breach.

KB-5 Data objects

Similar to data flows, data objects are an explicit representation of metadata in the enterprise architecture model. This representation allows marking data objects as personal data and tracing its flow across the organization. Various EA tools support this task with advanced analysis capabilities. Defining a data owner to each data object assigns a clear organizational responsibility for all processes on a particular set of personal data.

KB-6 Privacy engineering methods

The field of privacy engineering is concerned with the design and implementation of privacy-aware systems. It has produced a wealth of well-founded theories and methods that support this purpose. In previous work, we have presented a selection of privacy engineering methods and concluded that they are capable of addressing technical measures that are required by the GDPR (Huth and Matthes, 2019). Figure B.1 presents a general concept of privacy engineering methods. The elements in this figure are:

- Privacy definition: Solove (2006) characterizes privacy as an umbrella term for a set of related problems that concern personal information. His taxonomy distinguishes between problems of (1) information collection, (2) information processing, (3) information dissemination and (4) invasion.
- Privacy properties are positive statements of privacy goals. Conversely, privacy threats represent the opposite of the same properties.
- A privacy engineering method is designed to either support privacy properties or identify and prevent possible threats to privacy. A privacy engineering method combines this conceptual perspective with a framework of roles, stages, tasks, resources and outcomes (Martin and Del Alamo, 2017a).
- Privacy patterns are common solutions to recurring problems that are related to information privacy. They describe the context, the problem they address, the solution, and known implementations and effects in a structured manner. The website privacypatterns.org (UC Berkeley School of Information, 2020) originates from a cross-institutional research collaboration and provides a large collection of these patterns.
- Privacy enhancing techniques (PET) are technical mechanisms that support the concepts of privacy patterns.

KB-7 Guidelines & legal interpretation

The Article 29 Working Party was an independent supervisory body to the European Union that was established with Article 29 of the 1995 directive. It was made up of members of the national supervisory authorities of the member states and consulted the legislative bodies in data protection matters. Leading up to the GDPR, the Article 29 WP published a series of guidelines that discuss the implementation of single provisions of the GDPR, such as the right to data portability or transparency. The Article 29 WP ceased to exist when the GDPR entered into force and is now replaced by the European Data Protection Board with similar duties. The EDPB adopted the Article 29 WP recommendations and continues to publish advisory material on the GDPR.

In addition to advice on single aspects of the GDPR, there are a few holistic approaches that address the GDPR in its entirety. Most notably, the independent German supervisory authorities 2019 published their "Standard data protection model version 2.0", which discusses the GDPR based on protection goals and provides a method for implementing and maintaining compliance

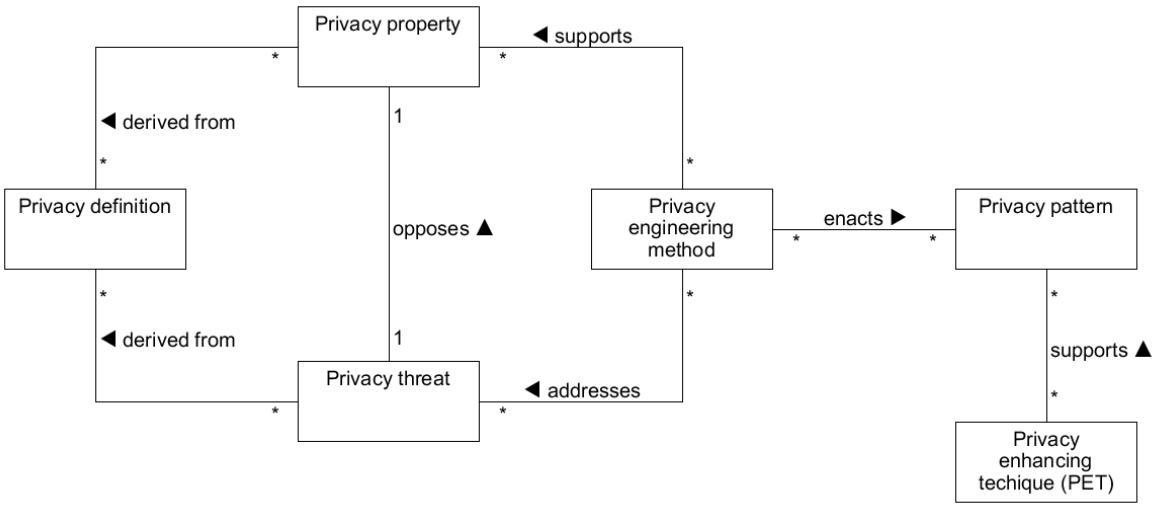

Figure B.1.: Conceptual representation of privacy engineering methods

with the GDPR. Besides the SDM, proprietary knowledge of how to operationalize GDPR requirements is often offered by consulting companies.

We see ProPerData as a practice-based reference model for achieving and maintaining GDPR compliance and hence, as a resource element for ProPerData itself.

B.4. Work units

Tasks, as used by ProPerData, are categories of activities that serve as classification scheme for the *work units* and *work products* of the method. This section presents the in-depth descriptions of each work unit that is part of ProPerData. Due to the diverse nature of the work units, a rigid, unified description is not capable of properly describing each of them. We rely on textual descriptions as a general rule, but enhance them with models and visualizations where appropriate.

B.4.1. Inform & educate

P-1 Data protection trainings

Rationale: Article 39 specifies the responsibilities of the data protection officer. Among them, the DPO is responsible for raising awareness for the regulation and training staff that is involved in processing activities. Further, data protection trainings are mandatory elements of the binding corporate rules (Art. 47).

Data protection trainings are aimed at presenting an overview of data protection within a relatively short amount of time, typically one or two days. Their broad, high-level objective limits

the amount of academic work from an engineering perspective, but gives rise to work that is focused on the behavioral aspects of data protection. In the domain of information security, serious games have been proposed for raising employee awareness in information security (Hendrix et al., 2016; Beckers and Pape, 2016). Serious games are games that combine the entertaining nature of games with educational aspects. To date, we are not aware of academic work that focuses on data protection training or serious games for data protection.

Supervisory authorities, e.g. the UK's Information Commissioner's Office (ICO), offer educational material on the GDPR[2]. Other private institutions contribute by offering on-site or online seminars.

Process:

- Identify affected employee groups: The employee groups with a general need to understand the regulation are all the roles in ProPerData (process owners, application owners, data owners, developers, IT operations employees, IT security employees and enterprise architects).
- Set time interval for repetition / refreshment of trainings. An interview partner referred to yearly trainings of all employees with customer contact.
- Create training material or employ suitable external trainer.

Discussion: Increasing awareness for data protection regulation has assisted the emergence of a new market for data protection companies. The fear of fines creates business opportunities and establishes data protection as an important topic, and spending money on data protection increases the overall perception of its value.

B.4.2. Verify existing processing activities

P-2 Analysis of existing processing activities for GDPR compliance

Rationale: Existing processing activities might have been established under different regulatory conditions, i.e. before the GDPR came into effect. Thus, they have to be checked for compliance with the general processing principles (Article 5) and the requirements on the security of processing (Article 32).

Process:

- Identify relevant processing activities, e.g. with the help of business process documentation (KB-1), the EA application repository (KB-2) or a CMDB export (KB-3).
- For each processing activity, verify the following properties:
  - Lawful basis of the processing: is the processing activity based on at least one of the following (Article 6):
    - The data subject has consented to the processing
    - Processing is necessary to fulfill a contract with the data subject
    - The controller processes data to comply with a legal obligation

[2]e.g. via its Youtube channel, `https://www.youtube.com/user/icocomms` (accessed 01/24/2020).

        - Data is processed to protect the vital interests of the data subject or another person
        - An official authority requires the processing in the public interest
        - The controller has a legitimate interest for processing the personal data
    - Is the processing tied to a clearly defined purpose, and does the purpose justify all the stored data? Will the data be deleted or anonymized if it is no longer processed?
    - Is the data accurate, and protected through organizational and technical measures?
- If any shortcomings are identified, the processing activity has to be adapted to meet the requirements of the regulation. If that is not the case, it is either re-engineered (P-3) or retired (P-4).

Discussion: The legal basis for processing is a very strict requirement at first sight. Consent cannot be faked (although some interface designs lure data subjects into consenting), there is not always a contract to be fulfilled or the data that is necessary for fulfillment of a contract is usually limited, and the legal reasons of legal obligation, vital interest or public interest typically do not hold. What companies have been using increasingly is *legitimate interest*, because there are no clear delimitations on what is legitimate or not. As a result, it has been interpreted (and stretched) to fit purposes from data analysis to improve services to sending out newsletters (as *"legitimate interest to maintain our business activity"*). Ultimately, future fines by supervisory authorities and subsequent court rulings will determine what may be considered as a legitimate interest.

B.4.3. Create new processing activities

P-3 Developing GDPR-compliant processing activities

Rationale: The principles for data processing are defined in Article 5. Data must be processed in a lawful, fair and transparent manner. The processing has to be limited to only the data that is necessary for the specified purpose and for as long as it is necessary. The data controller is fully accountable for these provisions and has to ensure the security of the data.

Engineering privacy-aware systems is the most widely researched topic of the ProPerData work units. Starting from the *Fair Information Practice Principles (FIPP)* (Ware, 1973), researchers have shaped the field of *Privacy Engineering* to "systematically address privacy issued while engineering information systems" (Gürses and Del Alamo, 2016, p.40).

Process: A generalized process for engineering privacy aware systems is a general software development process that incorporates privacy aspects. According to Crespo et al. (2015), this includes high-level functional analysis early on, the design of a privacy-friendly architecture, the incorporation of privacy patterns and privacy-enhancing techniques, planned responses to incidents and a plan for decommissioning (cf. B.2). The current state of the practice for implementing privacy requirements in newly developed processes and software are often developer guidelines.

Discussion: The scientific frameworks cover the necessary privacy properties that the GDPR requires, but there is no publication that evaluates the effectiveness or adoption of the frame-

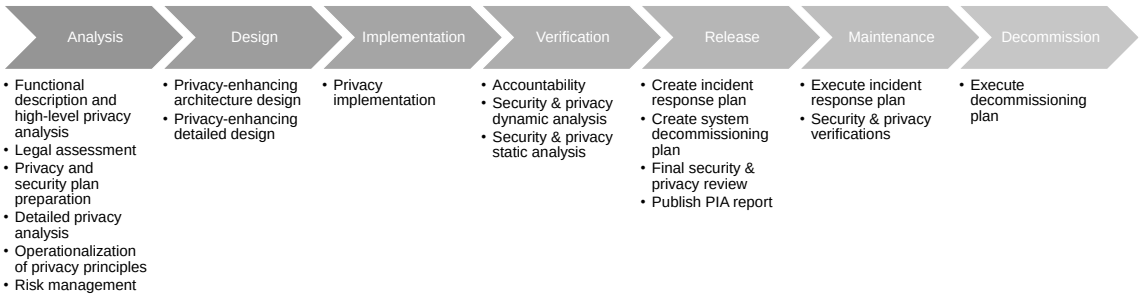

Figure B.2.: The PRIPARE lifecycle for privacy-friendly system design (adapted from Crespo et al. (2015))

works in practice (Huth and Matthes, 2019). Practitioners confirmed being unaware of such comprehensive frameworks and questioned their applicability.

Privacy patterns are solutions to recurring privacy problems that emerged from practical application (Colesky et al., 2016). There is little scientific work on the effectiveness of these patterns (Lenhard et al., 2017), but their origin in practical application implicitly validates their effectiveness. Privacy design strategies by Hoepman (2014) conceptualize privacy patterns. In our analysis of privacy engineering approaches, we found that privacy design strategies are able to provide "technical and organizational measures" to support the privacy properties in the GDPR.

A field that has not been studied adequately yet is how these properties can be ensured in agile development processes. We suggest that developing lightweight tools that support the practical application could support the development of privacy-aware systems more than complex frameworks.

P-4 Data deletion process

Rationale: Article 5 (1) (e) postulates that personal data may only be stored in an identifiable way for as long as the specified purposes require such storage. After the storage period, the data has to be anonymized or deleted. In addition to the planned deletion of all data that has been processed in a particular processing activity, Article 17 forces the data controller to delete personal data of single individuals upon request, given that there is no conflicting obligation.

There are multiple concerns when deleting data: (1) all affected data has to be deleted and (2) functionality of the processing application must remain intact, i.e. a deleted data point may not lead to inconsistencies.

Process: Data deletion should be considered early on in the establishment of a new processing activity. The GDPR itself does not specify how to implement this provision, but standards describe such processes and refer to the respective GDPR articles. As described by Hammer (2016), deletion concepts following DIN 66398 must have the following elements:

1. Deletion rules: “Deletion classes” are defined for combinations of holding periods and starting times. These holding periods could either follow directly from the legal provisions or they are defined by the company. For each deletion class, a deletion rule is specified.
2. Implementation instructions: The technology-agnostic standard deletion rules are detailed in implementation guidelines.
3. Exceptions: To allow for necessary flexibility, e.g. in case of lawsuits, exception rules can be defined.
4. Documentation: Deletion rules, implementation instructions and exceptions should be stored separately.
5. Responsibilities: The different stakeholders of the deletion concept must be assigned to the tasks that are specified by the norm.

Discussion: Data subjects interpreted the new provision (especially Article 17) as a general right to have all data deleted, and data controllers referred to the challenges in determining whether data can be deleted. If an enactable plan is established early on, uncertainties may not even arise. The same process described above holds for existing processing activities, though with less flexibility for early-on changes.

Deletion is only the most obvious action that is required by Article 5 (1) (e), but the original text states "personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary [...]" (European Union, 2016). In other words, anonymization is a permissible way of implementing this provision. However, the regulation deliberately does not give a clear definition of *anonymity*. Researchers have shown that by linking a publicly available voter registration list to seemingly anonymized health records, it is possible to re-identify individuals (Sweeney, 2002). Subsequently, Sweeney introduced the concept of k-anonymity:

Definition: *k*-anonymity
Sharing a combination of traits with at least k individuals in a sample.

The concept of l-diversity extends the measure of k-anonymity:

Definition: *l*-diversity
The property of having at least l well-represented values for each confidential attribute in a k-anonymous dataset (Danezis et al., 2014).

Practical application of anonymization methods largely depends on the type and interconnectedness of data. Relatively flat data structures can be anonymized quite easily, but with an increasing amount of touchpoints with the real world this is increasingly harder to do. Names are simple to replace with other valid names, but anonymizing addresses in a way that the result are valid addresses with the same distribution as before is hardly possible. Article 29 Data Protection Working Party (2014) issued guidelines on anonymization that are based on directive 95/46/EC, but should serve as a good reference for anonymizing personal data.

Enterprise architects supported deletion projects by supplying exports from the EA application repository (KB-3). A holistic account of the dependencies between applications facilitates the analysis of possible consequences if data is deleted in one system. The EA application repository may also be used to collect meta-information, such as the storage period. However, this process involves a large amount of manual work.

The data owner (R-4) is responsible for reviewing deletion requests and possible conflicts with other legislation:

> "We established a process to inform the data owner of a request to object processing or to delete data, and where the data owner has to report 'yes, I can block processing or delete' or 'no, I can't'. [...] Unless you have the feedback from all the involved data owners, you cannot execute the deletion process."

B.4.4. Conduct Data Protection Impact Assessments (DPIA)

P-5 Data protection impact assessment

Rationale: Article 35 of the GDPR states that a data protection impact assessment (DPIA) has to be carried out if a processing activity is likely to result in a high risk to the freedom of natural persons (European Union, 2016). According to Recital 89, the DPIA should replace the general obligation to notify the supervisory authority from directive 95/46/EC, which has shown to be costly and ineffective (European Union, 2016). The supervisory authority should be consulted if the DPIA indicates severe risks for the data subject, in particular if:

- the processing involves automated decisions with legal effects for the data subject;
- special categories of data according to Article 9 or criminal records according to Article 10 are processed; or
- a publicly accessible area is systematically monitored.

In addition to these general cases, the supervisory authority shall publish a list of the kinds of processing operations that require a DPIA.

The Article 29 Working Party, an independent advisory body to the European Union that was established with directive 95/46/EC, presents a simple decision diagram (see Figure B.3) for when to conduct a DPIA (cf. Article 29 Data Protection Working Party (2017a)).

Process: Bieker et al. (2016) derive a process for conducting a DPIA from recommendation guidelines by the supervisory authorities from France and the UK. The authors describe a three-stage process that involves (1) the identification of tasks and issues, (2) the evaluation of risks and (3) the identification, implementation and documentation of appropriate safeguards. Ideally, the person responsible for implementing the processing activity should also conduct the DPIA, with support from the DPO.

Alternative methods for a DPIA (such as ISO (2017)) should meet the following criteria to satisfy the requirements of the GDPR (sub-criteria and details can be found in Article 29 Data Protection Working Party (2017a)):

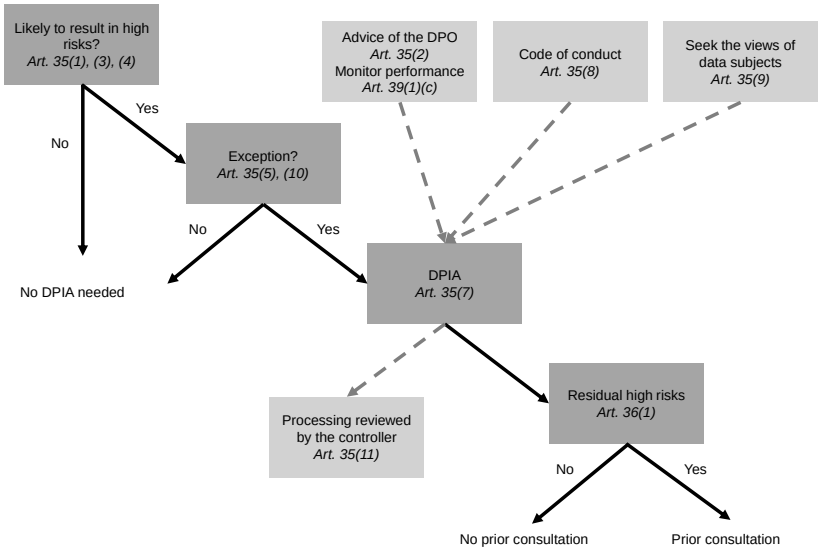

Figure B.3.: DPIA decision diagram (Article 29 Data Protection Working Party, 2017a)

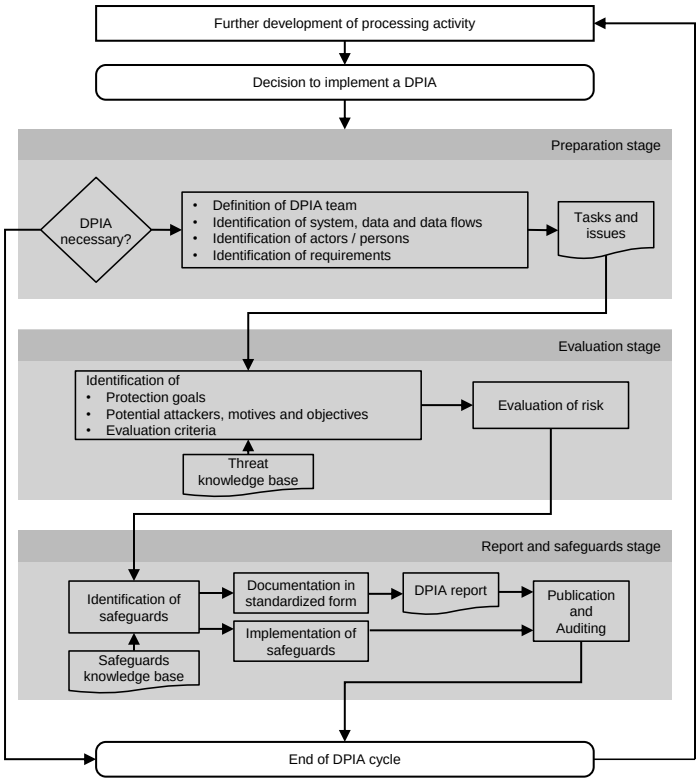

Figure B.4.: The DPIA process (adapted from Bieker et al. (2016))

- a systematic description of the processing is provided
- necessity and proportionality are assessed
- risks to the rights and freedoms of data subjects are managed
- interested parties are involved (i.e. DPO and the data subjects)

Discussion: Enterprise architects (R-7) reported supporting the DPIA through the organizational frame that EA provides: Established tools are able to send out and track surveys to application owners (R-3). This proved especially helpful in cases where the EA repository is used for documenting data protection information. An important element of the DPIA is the criticality of the processing. As one enterprise architect remarked:

> "The question is: how critical is an application? [...] Risk always exists, but the probability of occurrence, the frequency of occurrence... they differ."

B.4.5. Cooperate with supervisory authority

P-6 Respond to supervisory authority requests

Rationale: Article 31 shortly mentions the obligation of the controller to cooperate with the supervisory authority. This includes:

- Making the record of processing activities available to the supervisory authority (Article 30 (4))
- Collaboration regarding the DPIA (P-5)
- Communicate the binding corporate rules to the supervisory authority upon request

P-7 Communicate data breach

Rationale: Article 33 states that the controller has to notify the competent supervisory authority within 72 hours of becoming aware of a data breach. A data breach in terms of the GDPR is defined as *"the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed"* (Article 4(12)).

Process: The notification has to (cf. Article 33 (3))

- describe the nature of the breach,
- state the contact detail of the DPO,
- outline likely consequences, and
- describe measures taken or proposed in response to the data breach.

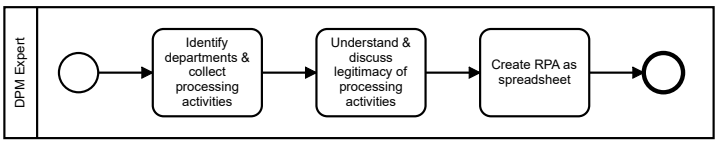

Figure B.5.: A simple process of creating the RoPA (Huth et al., 2019b)

B.4.6. Maintain records of processing activities

P-8 Maintain record of processing activities

Rationale: The record of processing activities serves the purpose of demonstrating compliance with the GDPR to the supervisory authorities (Recital 82). It has to be made available upon request only, but should always be readily available.

Process: Mandatory information for the RoPA includes the following:

- The name and contact details of the controller
- The name of the data processing activity
- The purposes and lawful basis of the processing activity
- The categories of data subjects and personal data
- The categories, names and contact details of recipients to whom the personal data have been or will be shared (both internal and external)
- The identification of third countries or international organizations in the case of transfers of personal data
- Retention period of different categories of data
- A description of the technical and organizational security measures

In a simple process for a RoPA (cf. B.5), the DPO identifies all departments that could be responsible for processing activities and contacts these stakeholders to collect the information, often via email. For further understanding, a direct discussion of the processing activity can take place.

Identifying the relevant processing activity is by far the largest challenge. Implicitly, many experts use organizational charts to find the right people (Huth et al., 2019b). Rather than starting from scratch, EAM experts have reported using the existing IT landscape documentation to identify applications that process personal data, because the applications point to the processing activities that are supported by these applications (Burmeister et al., 2020). Unpublished findings from our interview series indicate that the useful databases are configuration management databases (CMDB) and enterprise architecture application lists. While a CMDB holds only operational information, EA application lists typically contain metadata, such as the business domain or the application owner. Less frequently our interview partners reported documenting processes in their EA repositories.

Overall, we observed the following approaches for addressing the RoPA with EA support:

- Handing over the IT documentation to the DPM experts without further involvement of the enterprise architects. DPM experts used additional tools for survey in some cases.
- Enterprise architects used existing tools and their data collection functionalities to support and track responses from the application owners (Huth et al., 2020b).
- Some interviewees implement the entire RoPA in their EA tool, an approach that Huth et al. (2019b) also put forward.

Discussion: Some DPM experts were unaware of the extent of documentation that exists and emphasized the usefulness of having a starting point for the data protection documentation. For the first approach, what the interviewed enterprise architects criticized was not being consulted with respect to which list or which repository to use. These one-time exports could be outdated, leading to missing (or unnecessary) entries in the RoPA.

With stronger involvement of EA tools into the RoPA creation process, our experts referred to the established data collection process that helped tremendously in gathering the additional information from the application owners (and, in some cases, process owners). Where the EA tool served as the RoPA, the most common pitfall was too fine-grained information. An interviewee reported an effort with too many categories for personal data, which was hard to maintain and ultimately failed.

The EA tool industry already captures the synergy potentials between enterprise architecture management and data protection management. Multiple tools, among them ADO, LeanIX and BiZZdesign, incorporate modeling capabilities for this rather new field for EAM. Huth et al. (2019b) add custom properties to standard ArchiMate elements to model data protection documentation capabilities.

B.4.7. Conduct Audits

P-9 Data protection audit

Rationale: Audits are "an assurance function that some standard, method or practice is followed" (Halpert, 2011, p.16). The DPO as representative of the data controller has the responsibility to monitor compliance with the GDPR by executing audits (Article 39 (1)(b)). While data protection audits are mostly conducted in a collaborative manner (ICO, 2018, p.3), Article 58 (1)(b) grants the supervisory authority the right to assess an organization's compliance with the GDPR. A data protection audit ensures, verifies and tests policies and procedures to protect personal data, as well as detects gaps and yields change recommendations (ICO, 2018, p.4). The controller benefits from this procedure through independent expert opinions and resources (ICO, 2018, p.3).

Process: The UK ICO (ICO, 2018) describes three steps for a data protection audit by a supervisory authority:

1. Audit program development: In the planning phase, the supervisory authority identifies

high-risk controllers by considering past data breaches, data subject complaints and media reports of questionable data practices.

2. Audit approach: The supervisory authority and the organization agree on the scope of the audit, depending on generic known risks and specific concerns of the organization. Based on the agreement, the DPO sends requested documents to the supervisory authority, such as data protection documentation, training material or employee guidelines for handling personal data. In the subsequent on-site visit the auditors look for gaps and possibly undiscovered data breaches. They conclude with a final report that includes an assurance rating and suggestions to mitigate risks that arise in personal data processing. High-level results of the report are published.
3. Audit follow up: 6 to 12 months after the audit the organization demonstrates how the suggestions from the audit were implemented. The supervisory authority either approves the actions or decides on further steps.

Discussion: Since data protection audits are initiated and conducted by the supervisory authority, and the DPO takes a supportive role, our experts did not report on personal experiences with this task.

B.4.8. Interact with data subjects

P-10 Respond to data subject requests

Rationale: Enhanced data subject rights are a significant new addition in the GDPR. They include:

- The right to transparency (Article 12) and the right to information (Articles 13 and 14) grant the data subject to be informed of the processing before the processing takes place.
- The right of access (Article 15) represents a pivotal element of the data subject rights, because without knowing which data is processed and how, the rights to changes in the processing could not be exercised correctly (Ausloos and Dewitte, 2018, p.3).
- The right to rectification (Article 16) is meant to prevent adverse consequences of a controller processing incorrect personal data.
- The right to erasure (Article 17).
- The right to restriction of processing (Article 18).
- The right to data portability (Article 20) should "empower data subjects [...] to move, copy or transmit personal data easily from one IT environment to another" (Article 29 Data Protection Working Party, 2017c, p.4).
- The right to object to processing (Article 21).
- The right to object to automated individual decision making (Article 22).

Process: It is important to distinguish between different categories of data subjects: data

subjects can be clients, business partners or employees. For clients, the volume of data subjects is substantially higher than for other categories of data subjects, which makes the definition of processes more feasible.

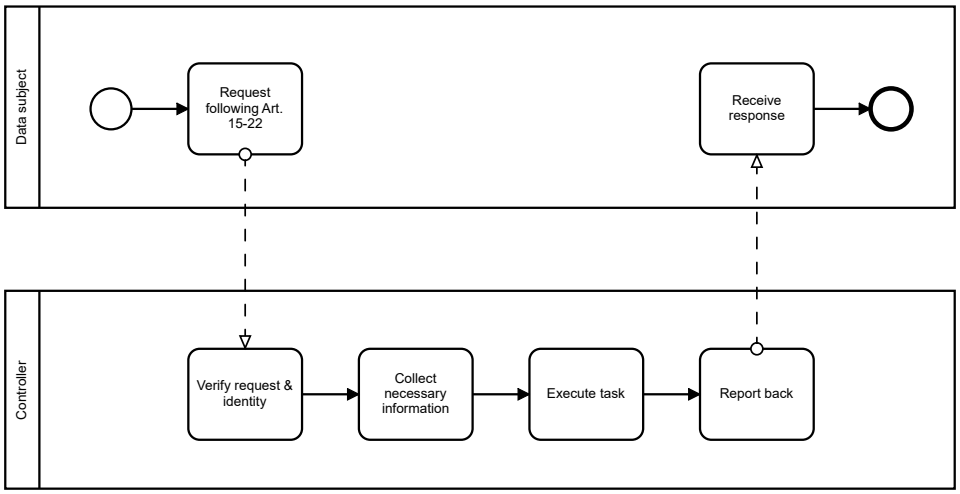

Figure B.6.: Generic process for answering data subject requests

The European Data Protection Supervisor (2010) issued guidelines on the implementation of data subject rights for *Regulation 45/2001 on the protection of personal data by European Union institutions and bodies*. We combine the information from this publication and on the *Guidelines on Automated individual decision-making and profiling for the purposes of Regulation 2016/679* (Article 29 Data Protection Working Party, 2017b). The latter gives a convenient summary of all the data subject rights.

- The right of access should be executable without constraints (i.e. not require to specify a reason for the request), free of charge, and the results should be returned within a reasonable time frame. However, it should not lead to disproportionate efforts for the controller. The format of the response depends on the nature of the data, but be understandable in a way that would allow the data subject to influence the processing.
- The right to rectification applies only to factual data, not to subjective statements. Ausloos et al. (2019) oppose that view: "The right to rectification applies to opinions and inferences of the data controller" (p.2).
- The recommendations on the right to erasure and the right to object are specific to European Union institutions (the processing institutions that Regulation 45/2001 addresses). Ausloos et al. (2019) argue that it is not enough to anonymize personal data and that a request for erasure should be taken as a request to immediately stop any processing of data from that individual.
- Regarding automated individual decision making, which is defined as a decision without meaningful assessment by a human (Article 29 Data Protection Working Party, 2017b, p.9), the principle of lawfulness, fairness and transparency (Article 5 (1) (a)) and the information requirements by Article 12 must be followed.

The documents do not give advice on the right to data portability, which is a new provision to the GDPR and intersects with competition law (Vanberg and Ünver, 2017). Article 29 Data Protection Working Party (2017c) and Huth et al. (2019a) discuss which data is affected by data portability requests and how it can be transferred. However, interviews with practicioners from non-information society enterprises (i.e. companies that mostly sell physical products) suggest that these requests are rare (Huth et al., 2019a).

Discussion: One enterprise architect referred to the importance of collaboration in defining processes for data subject requests:

> "We were involved in a project to ensure that we can answer data subject requests from clients. We were in a consulting role in that project, because what you don't want is another uncoordinated list."

However, the ability to use existing EA repositories hinges on the completeness of the documentation. Another interview partner remarked:

> "This is the great potential, to know at the click of a button which application processes which business objects and whether they contain personal data. And that is where the efficiency will be later on."

B.4.9. Report to management

P-11 Data protection reporting

Rationale: Since Management is accountable for GDPR compliance within the organization, data protection managers asserted that reporting is an essential task. Article 38 affirms that the DPO "shall report to the highest management level of the controller".

Process: Data protection reporting is not fundamentally different from other reporting activities (cf. B.7). Arising from an information need, intelligible information is created from raw data and presented to the accountable stakeholders. In the case of the GDPR, the accountable stakeholders are from top management. We believe that the overall structure of ProPerData provides a blueprint of preparing such reports.

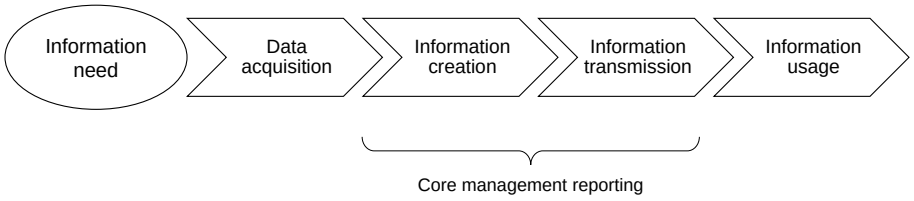

Figure B.7.: A reporting process, adapted from Taschner (2015)

B.4.10. Execute organizational tasks

P-12 Update privacy statements

Rationale: Articles 12, 13 and 14 lay out the requirements for making data processing transparent to the data subject. Recital 39 requires such information to be "easily accessible and easy to understand". Article 29 Data Protection Working Party (2017d) recommends making a privacy statement accessible with at most two taps/clicks in an online interaction.

Privacy statements should include all information that is necessary for making an informed decision to engage with a data controller: details about the data controller and DPO; the purposes and legal basis of processing, the categories of personal data and the (types of) recipients of that data; safeguards and storage periods; a statement of the data subject rights and, if applicable, the existence of automated decision making (Article 29 Data Protection Working Party, 2017d, p.38-40). According to Schaub et al. (2017), information requirements for other privacy legislation add to the overall length of privacy statements. McDonald and Cranor (2008) estimate the overall time effort to skim short privacy policies at 81 hours per year.

Process While privacy statements are generally created by legal experts and are therefore catering to legal obligations, Schaub et al. (2017) propose to distinguish between privacy statements and privacy notices. Privacy notices, in this context, are easily understandable complements to the privacy statements that are tailored to the transactional context and shall support the principles of notice and choice for the user. The authors present a design space for delivering such privacy notices (cf. B.8), and propose that privacy notices should be integrated in a user-centered design process.

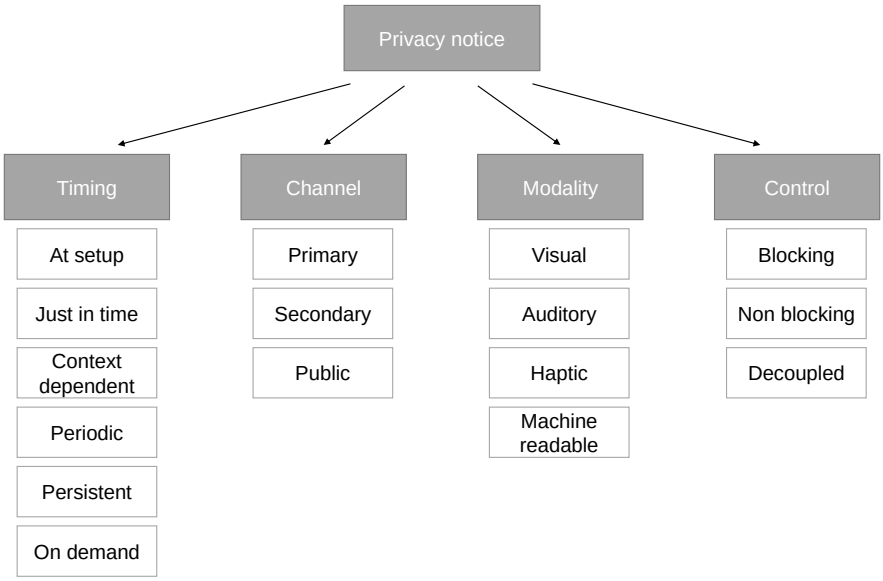

Figure B.8.: Design space for effective privacy notices, cf. Schaub et al. (2017)

P-13 Harmonize processing activities for data objects

Rationale: It is rarely the case that personal data is processed in only one processing activity. This might lead to conflicts regarding that data, for example:

- An online shop uses the personal address of a data subject for consent-based advertising and for delivery (fulfillment of the contract). The data subject revokes consent for advertising and requests immediate deletion of her data. Retention requirements might force the online shop to still keep the data for a fixed time period.
- A telephone carrier collects communication data for billing purposes only. The marketing department wants to make personalized, usage-based suggestions.

These examples illustrate the conflict that has to be resolved between different processing activities.

Process: Multiple companies reported establishing the role of *data owner* for data objects that are considered personal data. In an integrated creation or update process for processing activities (cf. P-3), the process owner must contact the data owner and negotiate the terms of processing. Likewise, for the update or deletion of the data object itself, the data owner has to be aware of possibly conflicting legislation and possible effects on data consistency in the application landscape.

P-14 Reflect and adapt GDPR implementation practices

Rationale: As multiple interview partners remarked, it is important to consider how the regulatory compliance efforts evolve in order to assess and improve the effectiveness of the processes that are already in place. Often, the GDPR processes were established bottom-up and evolved over time.

B.4.11. Leverage data protection efforts for business impact

While not immediately a topic of data protection management, we suggest that the implementation of data protection regulation should be associated with benefits as well. (Cavoukian and Dixon, 2013, p.7) draw the analogy that race cars have brakes to make them stop (defensive approach), but they also allow them to go faster around difficult tracks (enablement posture). This section highlights possible benefits of collaboration that our interview partners pointed out in a rather anecdotal way.

P-15 Leverage documentation of processing activities to identify business potential

When asked about the benefits of the GDPR implementation, an IT leader replied:

> "If I know how to organize data based on processes, then I have the capability to discover what I can digitize. [...] The right approach to digitalization is to look at the processes and organize the information objects."

Thus, the obligation to analyze and document the processing of personal data should not only be seen as an unproductive task, but as a chance to question established processing activities and understand the organization better.

P-16 Align information requirements and collection processes with other departments

An enterprise architect reported a particularly fruitful collaboration between data protection management, IT security management and enterprise architecture management:

> "From an [enterprise] architecture perspective, you always have the problem that models become obsolete. And the more people use it, the more it remains up to date. That is a huge benefit for the [enterprise] architecture model in itself. And the users, among them the data protection experts, can save a lot of work because of the up-to-date model."

A single shared model might not always be feasible, but the general importance of cross-departmental collaboration must be emphasized. DPM experts generally rated the value contribution of EAM as positive, but 26 out of 38 respondents did not collaborate with EAM. The main reasons for this were that the function does not exist in the organization (14 respondents), unawareness (4), no contact persons (4), doubts about the objectives and the necessary level of detail (3), or time limitations (5) (Vilser, 2019; Huth et al., 2020c).

B.5. Work products

WP-1 Processing agreements

A legally binding agreement between the controller and the processor that defines "the subject-matter of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects and the obligations and rights of the controller" (Article 28).

WP-2 Documentation of technical and organizational measures

A textual description of the measures taken to ensure the privacy properties in Table B.1 for each processing activity (“technical and organizational measures”).

WP-3 DPIA report

The DPIA report should follow a standardized form for readability (Bieker et al., 2016) and, along with a description of the processing activity and the purpose of processing, should include statements on (Article 35):

- the necessity and proportionality of the processing operation

GDPR Article	Privacy property
Pseudonymity	Art. 4 (5), Art. 25 (1), Art. 32 (1) (a)
Non-identifiability	Art. 5 (1) (e)
Unlinkability	Art. 34 (3) (a)
Confidentiality	Art. 5 (1) (f), Art. 32 (1) (b)
Integrity	Art. 5 (1) (f), Art. 32 (1) (b)
Availability	Art. 32 (1) (b)
Storage limitation	Art. 5 (1) (e)
Purpose limitation	Art. 24 (2)
Data minimization	Art. 25 (2)
Encryption	Art. 32 (1) (a), Art. 34 (3) (a)
Resilience	Art. 32 (1) (b)
Access	Art. 32 (1) (c)
Demonstrate compliance	Art. 24 (1)

Table B.1.: Privacy properties that must be ensured with technical and organizational measures. Adapted from Huth and Matthes (2019)

- the risks to rights and freedoms of individuals
- methods to address the identified risks

WP-4 RoPA

The record of processing activities should contain the following information (Huth et al., 2019b):

- The name and contact details of the controller
- The name of the data processing activity
- The purposes and lawful basis of the processing activity
- The categories of data subjects and personal data
- The categories, names and contact details of recipients to whom the personal data have been or will be shared (both internal and external)
- The identification of third countries or international organization in the case of transfers of personal data
- Retention period of different categories of data

- A description of the technical and organizational security measures

Local supervisory authorities provide RoPA templates, e.g. the Deutsche Datenschutzkonferenz (2018). Functionalities to maintain the RoPA are an important part of the offering of privacy tech companies, cf. the IAPP (2019b).

WP-5 Privacy statements

Privacy statements provide information about the data processor and the processing activities to the data subjects. While there is no specific format, Articles 13, 14 and 15 define which information the statement must include (Hintze, 2018). Only data controllers, i.e. the organization that determines the conditions for the processing activity has to provide a privacy statement. Essential parts are (Hintze, 2018, p.1131):

- the identity of the data controller
- the categories of data processed, if they are not obtained directly from the data subject
- whether providing personal data is mandatory, if the data is obtained directly from the data subject
- the recipients of the data
- the purposes of processing
- the existence of the data subject rights (access, correction, erasure, object, portability)

WP-6 Process description for data deletion

A documented process that describes preconditions, responsibilities and tasks to be executed for deletion of bulk data. This can be the case if the defined storage period is over or the processing activity is discontinued.

WP-7 Audit results

The auditor will issue a report with the audit results, including (ICO, 2018, p.9):

- an assurance rating for each scope area
- details on non-conformities and associated risks
- prioritized recommendations to mitigate the identified risks

WP-8 Processes for the execution of data subject rights

The processes define:

- The initial point of contact for data subjects and a verification procedure

- Dissemination of the request to the responsible person of the processing activity
- Instructions for identifying relevant/affected data
- Guidelines or templates for responses to data subjects
- A time constraint for answering the request

WP-9 High-level management report of data protection activities

A management report of data protection activities should be integrated in the regular reporting process and may include:

- Overall assessment of compliance with the regulation (Article 39 (1)(d))
- Results of DPIAs
- Status of workforce data protection trainings

WP-10 Guidelines for admissible processing vs. obligation to involve DPM

Guidance material for product owners and developers regarding which type of processing is admissible without involving data protection management and when they must consult data protection experts. This can include:

Guidelines on	**Example**
General statements on types of personal data	e.g. obligation to ask for consultation when location data is involved
General permissions and necessary conditions	e.g. capturing usage statistics for product optimization if the data subject has consented
Admissible technologies	e.g. certain third-party libraries or encryption algorithms

Table B.2.: Examples for guidelines to process owners and developers

For classification of personal data, and easy-to-follow set should be defined, e.g. as presented in Table B.3.

WP-11 Data privacy coordinator

The role of data privacy coordinator serves as a facilitator for addressing possible conflicts between business requirements, data protection requirements and the overall IT strategy. While the DPM experts are frequently assigned to top management, the data privacy coordinator is an employee of the business or IT departments.

Criterion	Example
Type of data subject	Prospective client; client/customer; client (child); employee; business partner
Type of personal data	Address; location; financial; medical; political/ethnic/religious; interests/preferences

Table B.3.: Classification criteria for personal data

> "The privacy coordinator must be in very close contact with the central data privacy department. There is a privacy coordinator within the IT department, who is responsible for data protection topics in IT. The HR data privacy coordinator has other topics, of course."

The data privacy coordinator serves as an extended arm of the central DPM experts in the organization.

WP-12 Shared repository

A shared documentation of IT applications and business processes that captures the information requirements of multiple stakeholders, e.g. IT security, data protection and enterprise architecture management. Each stakeholder consumes and contributes information.

Note that not each information requirement of each stakeholder can and should be captured. As multiple interview partners remarked, such a shared model cannot "represent the whole world", and should therefore be seen as a consolidated entry point to further investigation.

If such a shared, collaborative repository cannot be established, central documentation should be made available. DPM experts in many organizations used EA application lists or CMDB exports as a starting point for their compliance endeavor. However, descriptive information about these documents should clarify the information base and the timeliness of the data, since some enterprise architects reported that DPM experts used outdated versions of these lists (Huth et al., 2020b).