ProPerData - A process model to support GDPR compliance

Dominik Huth, Florian Matthes
Chair for Informatics 19
Software Engineering for Business Information Systems (sebis)
Technical University of Munich
Boltzmannstr. 3, 85748 Garching bei München, Germany
{dominik.huth, matthes}@tum.de

May 2020
version 1.1

Abstract

The General Data Protection Regulation (GDPR) has changed the perception towards privacy and data protection worldwide. Passed in 2016 and in force since 2018, the regulation has been a steady part of the academic and practical discourse over the past years. However, companies still struggle with the task of becoming compliant, mainly because of the large interdisciplinary scope and the overall complexity of the regulation. Once established, maintaining GDPR compliance in an accelerating business environment remains a challenge.

With this report, we present ProPerData, a process model for the protection of personal data. It addresses software developers and enterprise architects of large organizations and aims to provide a structured overview of the GDPR and a clear definition of responsibilities.

ProPerData is organized along 11 tasks that are derived from the GDPR. 16 work units of ProPerData are assigned to the tasks and executed by ProPerData stakeholders. We account for 7 resources that support the work units and 12 work products that result from them. The work units take place at one or more of the 10 stages or events of ProPerData.

Acknowledgment

For compiling this report and our process model, we relied on the experiences and knowledge from a large number of professionals. We would like to express our gratitude for providing these invaluable insights into how their companies addressed the GDPR, what worked well and what did not. These experiences helped tremendously in designing an artifact that claims relevance in a practical context.

Further, we would like to thank the colleagues and students that contributed to these results with discussions, intelligent ideas, remarkable commitment and support in transcribing hours of interviews.

This research has been sponsored by the German Ministry of Education and Research (BMBF) through grant 01IS17049 / UMEDA. The responsibility for the content of this work lies with the authors.

München, 06.05.2020

Dominik Huth

Contents

CHAPTER 1

Motivation

1.1. Goal of this report

The GDPR has been discussed to a great extent in the media, both before and after the regulation entered into force. However, most of these contributions either elaborate on the importance of acting, or point out innovations for single challenges. What we did not observe were attempts to provide a complete understanding of the regulation and support interdisciplinary implementation projects.

With this report, we disseminate the consolidated findings of our research endeavor. We draw our findings from continuous monitoring of academic and non-academic literature, interviews with IT, legal and data protection experts, and many discussions with researchers in the IT and privacy domains and related fields.

Our process model *ProPerData* is especially aimed at IT professionals, whose primary field of expertise is not data protection management. The main IT professional roles we consider are software developers and enterprise architects. Thereby, ProPerData serves as a holistic entry point for GDPR efforts of large *data controllers*, i.e. organizations that determine the purpose of data processing (Huth et al., 2018). We provide a structured approach to understand the overall picture of GDPR compliance, support the identification of individual work units and present insights from research and practice on how to perform these work units. The common language of our visual summary supports communication among the stakeholders, including data protection management:

- **Software developers'** main work unit in ProPerData is the implementation of compliant processing activities (P-3). Our process model summarizes the most important responsibilities and approaches, and references sources for in-depth support for these methods.

In addition, the canvas fosters an understanding of the overall context and further tasks where developers are in a supporting role.

- **Enterprise architects** receive a tool for clearly communicating the benefits that enterprise architecture provides for data protection. The main work units are the record of processing activities (P-8) and coordinating the information needs in a shared model for mutual benefits among different stakeholders (P-16).

ProPerData covers both the initial setup phase and maintaining compliance in changing business environments. Despite the observation that there are still companies who struggle to implement the GDPR provisions, we see the more important contribution in the second purpose. Rapidly changing market environments and new development paradigms underline the need to proactively manage regulatory compliance for the years to come.

1.2. General Data Protection Regulation

The General Data Protection Regulation was passed in 2016, replacing the previous European Directive 95/46/EC from 1995 (European Union, 2016). Since then, considerable attention has been paid to the new legislation in the media and in academia, which continued even after the GDPR entered into force in 2018. Recent studies still report that companies have not fully implemented all provisions (TrustArc, 2018; CIPL and AvePoint, 2018; International Association of Privacy Professionals and TrustArc, 2019; Ernst & Young and International Association of Privacy Professionals, 2019). Consequently, GDPR implementation remains a continuing challenge (Mikkelsen and Strandell-jansson, 2018). Further, examples of alarming fines underline the need to take the regulation seriously: Google was fined €50 million in France for intransparent privacy statements, British Airways and Marriot were fined £183 million and £99 million, respectively, for insecure data processing, and a German real estate management company had to pay €14.5 million for noncompliance with general processing principles (CMS, 2020).

Drawing from a series of interviews with several organizations, Sirur et al. (2018) name the breadth of the regulation, the interpretation of qualitative recommendations and an understanding of complex data networks as the main challenges that companies report. In contrast to the industry studies, the respondents in larger organizations were quite confident that they are able to address the complete regulation.

1.3. Reference framework for classifying ProPerData work units

For the purposes of this work, we need a frame of reference to classify the work units and work products for GDPR compliance that we present with ProPerData. Tikkinen-Piri et al. (2017) conducted an in-depth comparison of the GDPR with the 1995 directive and present a set of twelve key implications of the GDPR on personal data intensive companies. Based on these implications, we developed a set of data protection officer (DPO) tasks and examined them in a survey with 38 data protection officers (Huth et al., 2020b). Combining these two approaches, we use the following classification of GDPR tasks for ProPerData:

- Awareness-raising and training
- Verifying compliance of existing processing activities
- Support creation of new processing activities
- Identify need & conduct DPIAs (data protection impact assessments)
- Cooperation with supervisory authority
- Maintaining the record of processing activities (RPA)
- Conducting audits
- Dealing with data subjects
- Report to management
- Superordinate activities
- Promote organizational benefits

1.4. Theoretical background and research approach

A number of privacy engineering methods have been published in the past years, which are aimed at developing privacy-compliant systems (Kalloniatis et al., 2008; Deng et al., 2011). As a way to conceptualize the common ground among these methods, Martin and Del Alamo (2017) develop a *metamodel for privacy engineering methods (MPEM)*, which is built as an extension to the ISO 24744 metamodel for software and systems development methodologies. Within these models, three layers of abstraction are recognized: (1) the metamodel layer, (2) the method layer and (3) the project layer. Figure 1.1 shows how our method ProPerData relates to the method layer.

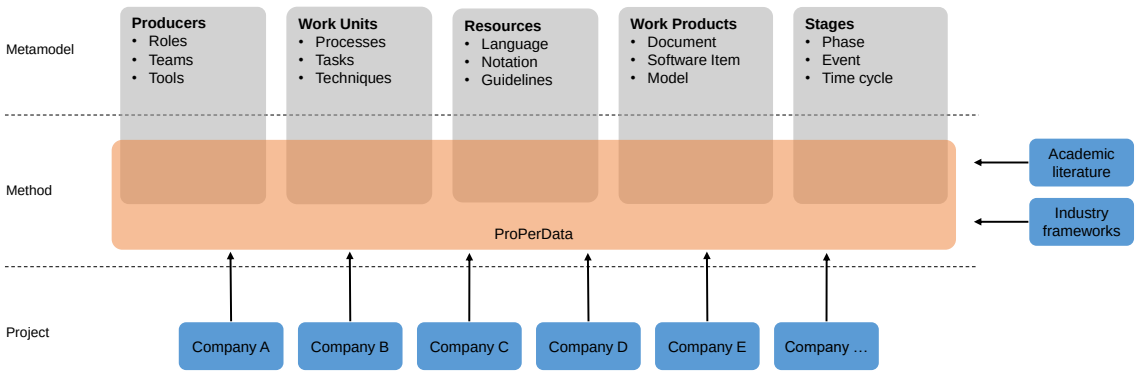

Figure 1.1.: ProPerData in the MPEM context

The metamodel is comprised of the following elements:

1. **Producers**: A role, team or tool that carries out one or more of the work units that are described in the method.
2. **Work units**: Modular activities that are part of the method, such as processes, tasks or techniques.
3. **Resources**: Preexisting elements that can be used 'as is' when executing work units.
4. **Work products**: Artifacts that are generated as results of the execution of work units within the method.
5. **Stages**: Time frames within the method that can either represent a duration (phases or time cycles) or instantaneous events.

ProPerData is derived from various input data and offers a holistic, unified view of all these perspectives:

- Continuous analysis of the academic literature over a timespan of more than 2 years, including overarching methods (such as the standard data protection model 2.0, (Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder, 2019)) and work addressing single aspects of the GDPR.
- A series of 24 interviews with 29 enterprise architecture experts, presented in Huth et al. (2020a) and Burmeister et al. (2020).
- Four interviews and a focus group interview with data protection experts that were centered on the RPA, but also yielded additional insights Huth et al. (2019b).
- Interviews with DPOs, as well as the results of a detailed survey on DPO tasks and challenges Huth et al. (2020b).
- Numerous topic discussions with data protection experts at practitioner conferences and privacy researchers at academic events and meetings.

CHAPTER 2

ProPerData - A process model for GDPR compliance

ProPerData is a method 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 an overview of the method in Figure 2.1 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 also shows the stakeholder roles and resources of the method, but does not incorporate relationships that involve these two groups. These relationships are explained in detail within the work units 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. An overview of the work unit titles is shown in Table A.2.

2.1. 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., 2020b).

R-2 Process owner: The person from the business department who is responsible for a business

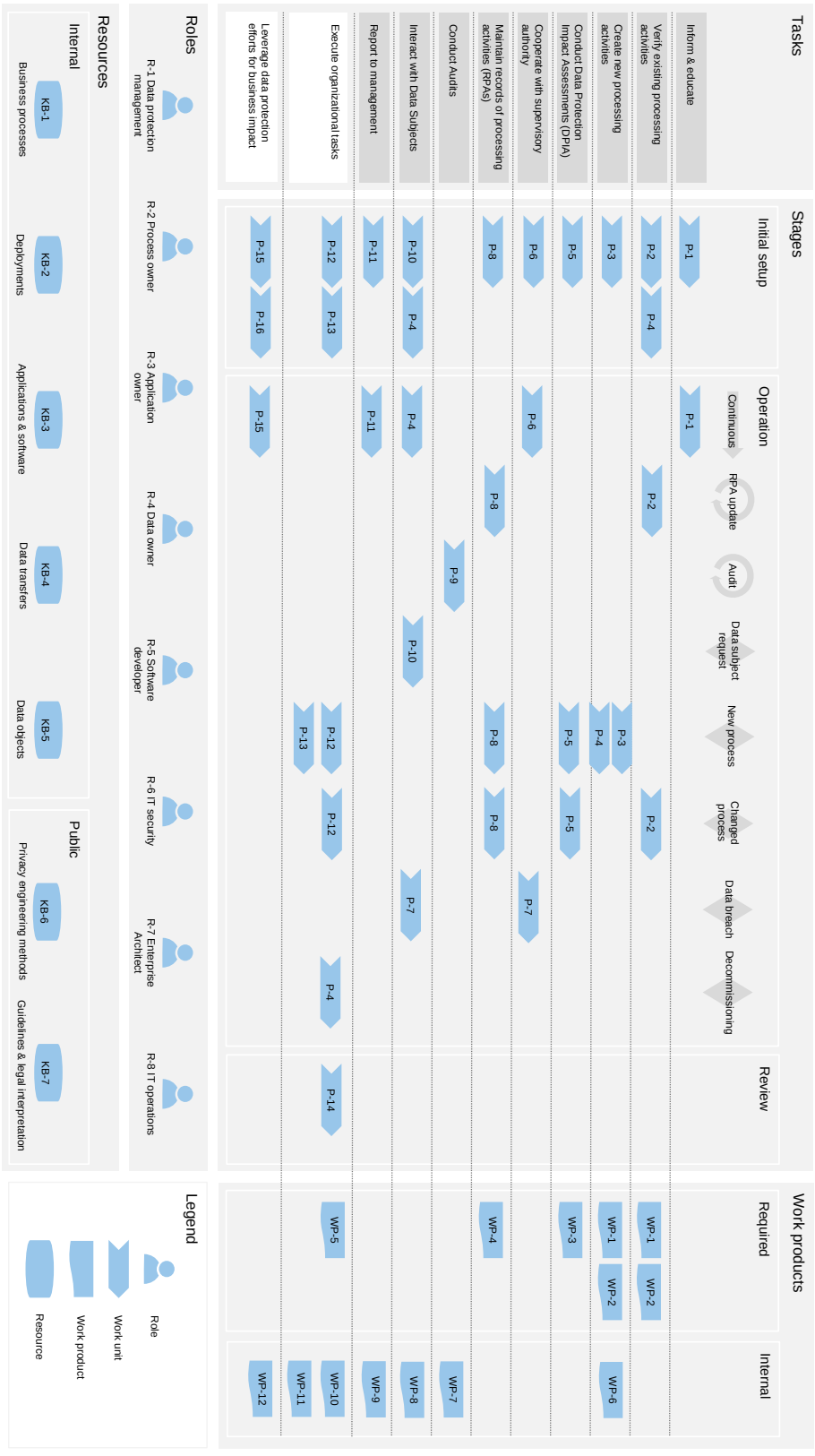

Figure 2.1.: The ProPerData method

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.

2.2. 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 (see section 1.3). 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
- RPA 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, 2017), 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 RPA 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.

2.3. 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 (2017) 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,

[1] SEMDM and MPEM define milestones rather than events. We adapted the notion to the organizational scope of ProPerData.

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 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 flows

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 2.2 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, 2017).
- 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

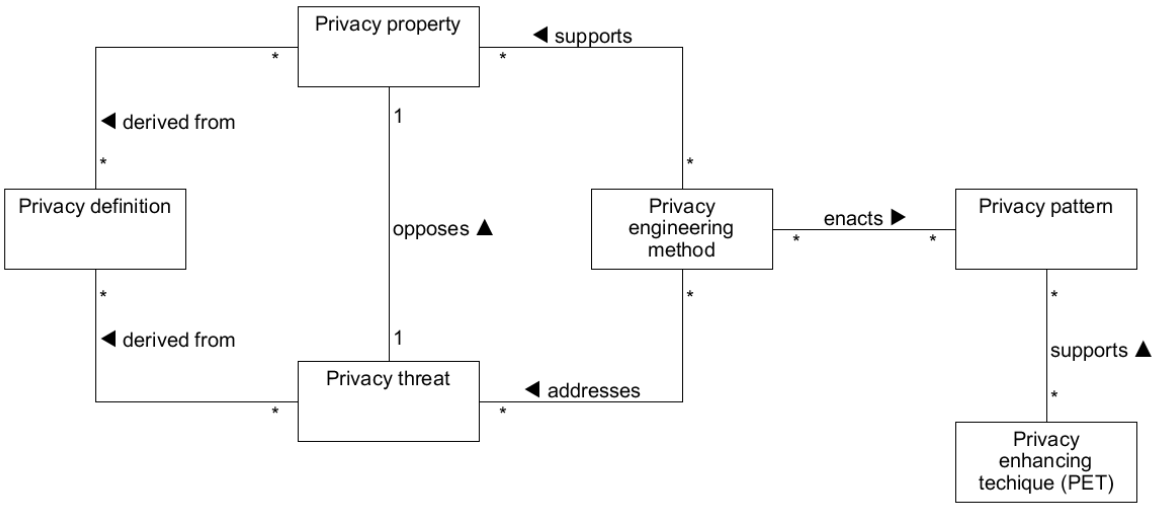

Figure 2.2.: Conceptual representation of privacy engineering methods

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 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.

2.4. 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.

2.4.1. 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.

2.4.2. 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 compli-

[2]e.g. via its Youtube channel, `https://www.youtube.com/user/icocomms` (accessed 01/24/2020).

ance 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
    - 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.

2.4.3. 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. 2.3). The current state of the practice for implementing privacy requirements in newly developed processes and software are often developer guidelines.

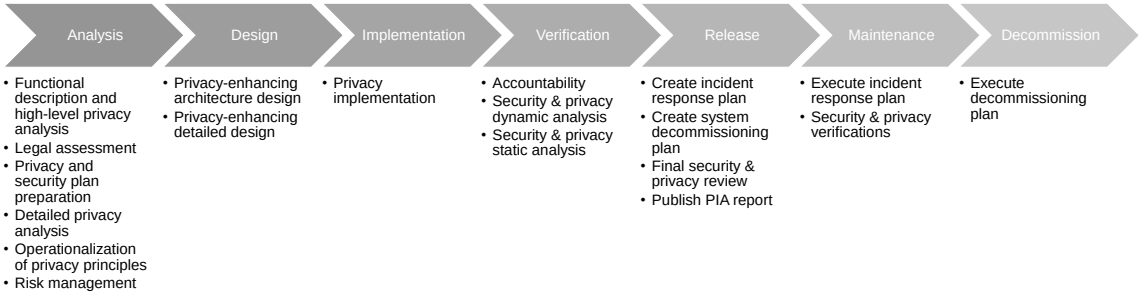

Figure 2.3.: The PRIPARE lifecycle for privacy-friendly system design (adapted from Crespo et al. (2015))

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 frameworks 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 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."

2.4.4. 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 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 2.4) for when to conduct a DPIA (cf. (Article 29 Data Protection Working Party, 2017a)).

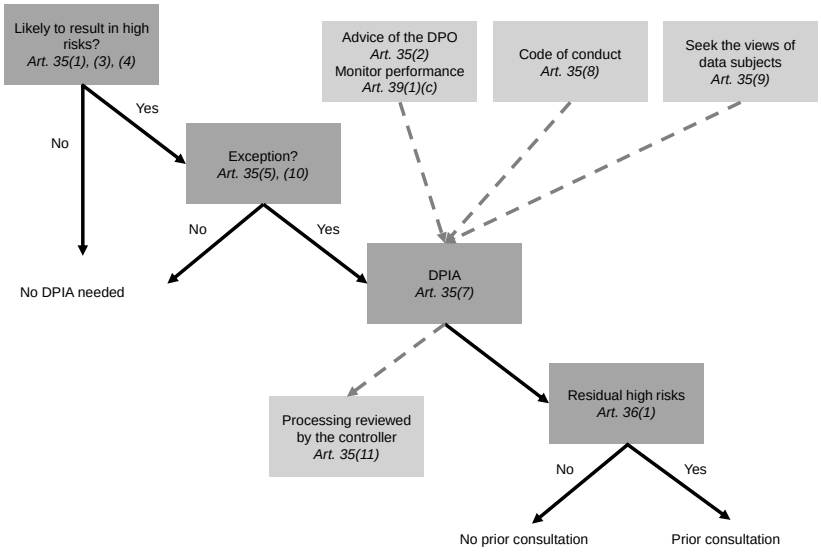

Figure 2.4.: DPIA decision diagram (Article 29 Data Protection Working Party, 2017a)

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)):

- 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:

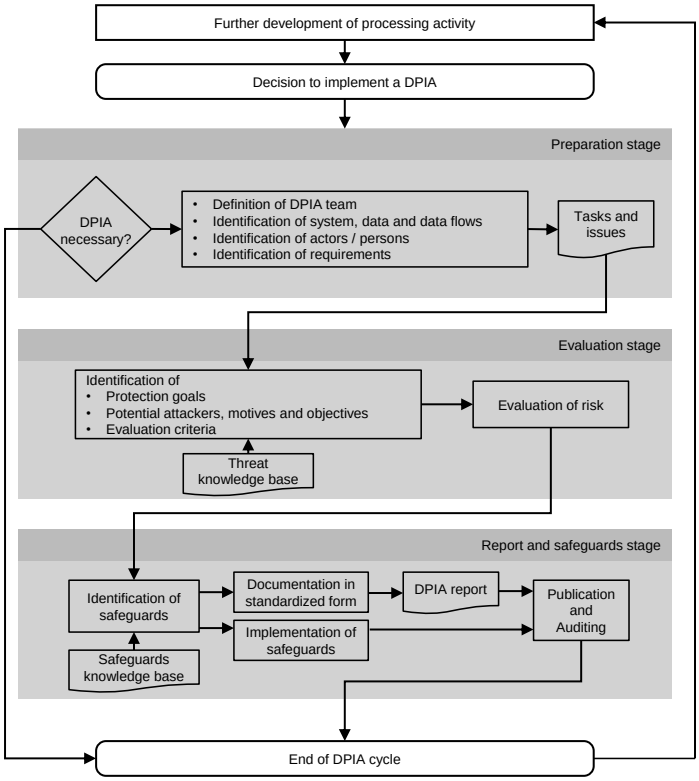

Figure 2.5.: The DPIA process (adapted from Bieker et al. (2016))

> "The question is: how critical is an application? [...] Risk always exists, but the probability of occurrence, the frequency of occurrence... they differ."

2.4.5. 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.

2.4.6. Maintain records of processing activities (RPA)

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 RPA 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 an RPA (cf. 2.6), the DPO identifies all departments that could be responsible for processing activities and contacts these stakeholders to collect the information, often via E-Mail. 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

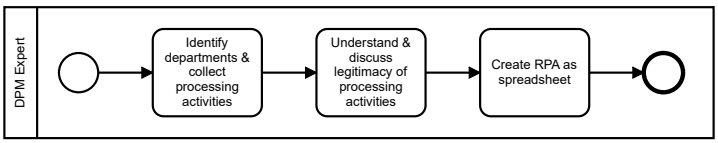

Figure 2.6.: A simple process of creating the RPA (Huth et al., 2019b)

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 RPA 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., 2020a).
- Some interviewees implement the entire RPA 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 RPA.

With stronger involvement of EA tools into the RPA 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 RPA, 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.

2.4.7. 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.

2.4.8. 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 subject 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.

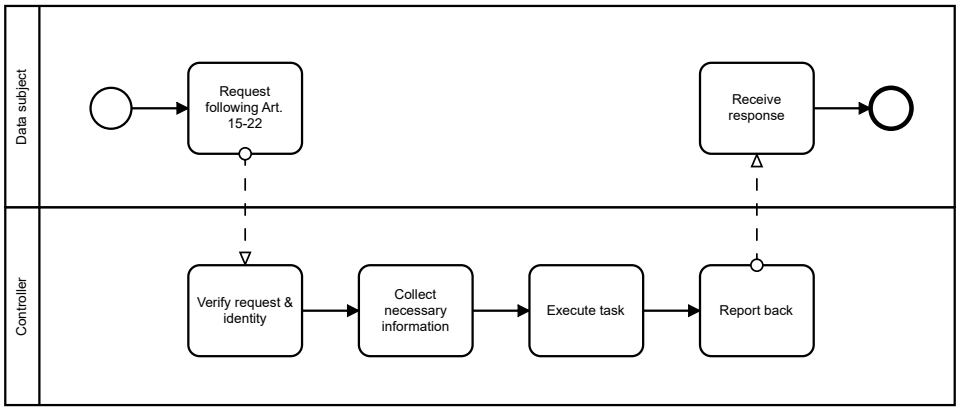

Figure 2.7.: Generic process for answering data subject requests

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."

2.4.9. 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. 2.8). 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.

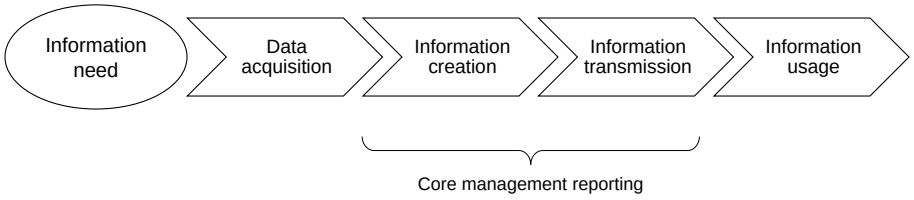

Figure 2.8.: A reporting process, adapted from Taschner (2015)

2.4.10. 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. 2.9), and propose that privacy notices should be integrated in a user-centered design process.

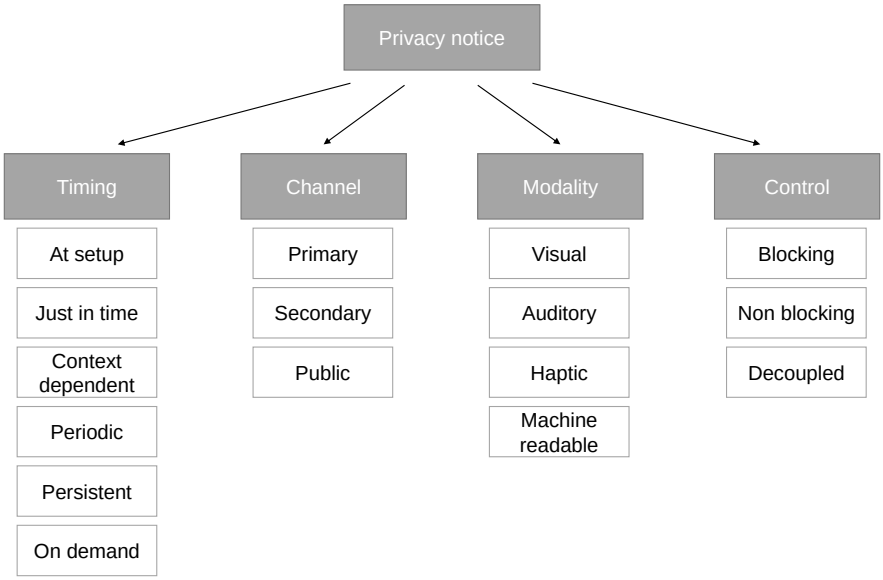

Figure 2.9.: Design space for effective privacy notices, cf. Schaub et al. (2017)

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.

2.4.11. 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. (**?**, p.7) draws 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 points and 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., 2020b).

2.5. 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 2.1 for each processing activity ("technical and organizational measures").

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 2.1.: Privacy properties that must be ensured with technical and organizational measures. Adapted from Huth and Matthes (2019)

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

- the risks to rights and freedoms of individuals
- methods to address the identified risks

WP-4 RPA

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 RPA templates, e.g. Deutsche Datenschutzkonferenz (2018). Functionalities to maintain the RPA are an important part of the offering of privacy tech companies, cf. International Association of Privacy Professionals (2019).

WP-5 Privacy statements

Privacy statements provide information about the data processor and the processing activities to the data subjects. While there is no specifiec format, Articles 13, 14 and 15 define which information the statement must include (**?**). Only data controllers, i.e. the organization that determines the conditions for the processing activity has to provide a privacy statement. Essential parts are (**?**, 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 2.2.: 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 2.3.

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 2.3.: Classification criteria for personal data

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.

> "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., 2020a).

CHAPTER 3

Conclusion

3.1. Summary

This report introduces ProPerData, a process model for the protection of personal data. ProPerData comprises 8 stakeholders, 10 stages, 16 work units, 6 resources and 13 work products.

We created ProPerData to support software developers and enterprise architects in gaining a holistic perspective on what the implementation of privacy legislation, in particular the GDPR, means for them. Necessarily, the vocabulary of one group differs from the other, creating the challenge of mutual understanding. ProPerData aims to bridge these differences by providing a scientifically developed frame of reference.

Further, ProPerData presents detailed concepts for fulfilling GDPR tasks: the record of processing activities (Article 30), the right to data portability (Article 20) and the creation of compliant processing activities (Article 5, among others). These concepts are based on practitioner input and literature analysis and synthesis as well.

This section summarizes our most important findings and suggestions for implementing data protection regulation in an effective manner.

Establish the communication channels and unlock the value of information that already exists within the organization.

A recurring theme in all our practitioner interviews were that the challenges of fulfilling the GDPR requirements were less technical than organizational:

- Enterprise architects reported isolated documentation efforts of the data protection teams and the usage of stale data, where established EA processes could have facilitated these efforts and created a sustainable framework for changes in the processing activities or future compliance efforts. Where a close collaboration existed, our interview partners underlined the mutual benefits for each collaborator.
- Data protection experts have a variety of educational backgrounds, including legal, business and IT backgrounds (Vilser, 2019). Depending on the background, knowledge of other departments may be focused on certain areas. This results in blind spots for organizational functions that have long been established. We encourage DPM experts to actively investigate the existence of processes that guide software development and IT documentation, as this can inspire fruitful collaboration across the organization.
- Software developers tend to describe data privacy with security terminology and neglect threats that are caused by improper usage of personal data or lack of control for the data subject (Hadar et al., 2018). Communicating the multifaceted challenges of privacy regulation may sensitize developers to other privacy aspects.

Integrate data protection considerations in every process

Data protection cannot be enforced as a standalone discipline and an organization's primary goal is advancing its business. Therefore, privacy considerations should be integrated seamlessly into development and management processes.

Leverage the necessary effort for data protection for business opportunities

Compliance with privacy regulation has the goal of avoiding penalties and preventing harm through negative publicity. However, the effort spent on privacy compliance may serve as a lever to the business and IT strategy. In particular, our interview partners articulated the following opportunities to do so:

- Most business activities involve personal data of customers, employees or business partners. A complete account of these business activities, organized as processing activities, illustrates potential for digitalization and digital transformation[1].
- The necessity to account for data protection in development processes is a trigger to rethink current tooling and processes and make them more efficient.
- Necessary documentation of processing agreements are a starting point to the assessment of cloud providers and hence an enabler for the cloud strategy.

[1] We adopt the notion that digitization refers to the digitization of documents, digitalization to the digitization of business processes and digital transformation to the digitization of business models.

3.2. Outlook

Our next step is to evaluate ProPerData together with software developers, enterprise architects and data protection experts. The assessment will cover ProPerData's comprehensiveness and fit to the stakeholders, its correctness, applicability and completeness, and suggested changes and enhancements. We will adapt ProPerData according to the feedback we receive.

The findings and structure of ProPerData serve as a general model for understanding the multifaceted tasks of data protection regulation. As one interview partner pointed out, the GDPR will be the point of reference for future data protection legislation all over the world. The recent California Consumer Privacy Act (CCPA) applies to a more narrow selection of companies, but also introduces enhanced data subject rights and is measured against the GDPR.

Bibliography

Article 29 Data Protection Working Party. Anonymisation Techniques. Technical Report April, Article 29 WP, 2014. URL http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

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. URL http://ec.europa.eu/justice/data-protection/index_en.htm.

Article 29 Data Protection Working Party. Guidelines on transparency under Regulation 2016/679. Technical report, Article 29 WP, 2017d. URL http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227.

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

Jef Ausloos, Rene Mahieu, and Michael Veale. Getting Data Subject Rights Right. 2019. URL https://osf.io/preprints/lawarxiv/e2thg.

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.

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.

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. doi: https://hdl.handle.net/10125/64430.

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.

CIPL and AvePoint. Organisational Readiness for the European Union General Data Protection Regulation. Technical Report March, CIPL, 2018. URL `https://www.informationpolicycentre.com/global-readiness-benchmarks-for-gdpr.html`.

CMS. GDPR Enforcement Tracker, 2020. URL `http://www.enforcementtracker.com/`. Last accessed: 02/22/2020.

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, 2015. URL `http://pripareproject.eu/`.

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. URL `www.enisa.europa.eu`.

Mina Deng, Kim Wuyts, Riccardo Scandariato, and Bart Preneel Wouter. A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering*, 16(1):3–32, 2011. doi: 10.1007/s00766-010-0115-7.

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.

Ernst & Young and International Association of Privacy Professionals. IAPP-EY annual privacy governance report 2019. Technical report, EY, 2019.

European Data Protection Supervisor. Guidelines on the Rights of Individuals with regard to the Processing of Personal Data. Technical report, EDPS, 2010. URL `https://edps.europa.eu/sites/edp/files/publication/14-02-25_gl_ds_rights_en.pdf`.

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`.

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.

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.

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 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, 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, 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, 2020a. doi: http://hdl.handle.net/10125/64457.

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, 2020b. to appear.

UK ICO. A guide to ICO audits. 2018. URL `https://ico.org.uk/media/for-organisations/documents/2787/guide-to-data-protection-audits.pdf`.

International Association of Privacy Professionals. 2019 Privacy Tech Vendor Report. Technical report, International Association of Privacy Professionals, 2019. URL `https://iapp.org/media/pdf/resource_center/2019TechVendorReport.pdf`.

International Association of Privacy Professionals and TrustArc. Measuring Privacy Operations. Technical report, IAPP and TrustArc, 2019.

ISO. ISO/IEC 29134:2017 Information technology — Security techniques — Guidelines for privacy impact assessment, 2017.

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.

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.

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, 2017.

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.

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.

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. URL `http://arxiv.org/abs/1808.07338`.

Daniel J Solove. A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154(3):477, 2006. ISSN 00419907. doi: 10.2307/40041279. URL `http://www.jstor.org/stable/10.2307/40041279?origin=crossref`.

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.

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*, 1(2017), 2017. ISSN 02673649. doi: 10.1016/j.clsr.2017.05.015.

TrustArc. GDPR Compliance Status. Technical report, TrustArc, 2018. URL https://download.trustarc.com/dload.php/?f=M9T2K99J-729.

UC Berkeley School of Information. privacypatterns.org, 2020. URL https://privacypatterns.org/.

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. URL http://ejlt.org/article/view/546.

Michael Vilser. Empirical Task Analysis of Data Protection Management, 2019. URL https://wwwmatthes.in.tum.de/pages/12wb2907mhi4k/Bachelor-s-Thesis-Michael-Vilser.

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.

APPENDIX A

Appendix

ID	Work unit
P-1	Data protection trainings
P-2	Analysis of existing processing activities for GDPR compliance
P-3	Developing GDPR-compliant processing activities
P-4	Data deletion process
P-5	Data protection impact assessment
P-6	Respond to supervisory authority requests
P-7	Communicate data breach
P-8	Maintain record of processing activities
P-9	Data protection audit
P-10	Define response process to data subject requests
P-11	Respond to data subject requests
P-12	Data protection reporting
P-13	Privacy statements
P-14	Assign ownership of data objects
P-15	Leverage documentation of processing activities to identify business potential
P-16	Create a model of shared information requirements with other departments and collaboratively collect data

Table A.1.: Overview of work units

ID	Work product
WP-1	Process description for data deletion
WP-2	Processing agreements
WP-3	Documentation of technical and organizational measures
WP-4	DPIA report
WP-5	RPA
WP-6	Audit results
WP-7	Processes for the execution of data subject rights
WP-8	High-level management report of data protection activities
WP-9	Classification criteria for personal data
WP-10	Guidelines for admissible processing vs. obligation to involve DPM
WP-11	List of data owners
WP-12	Data privacy coordinator
WP-13	Shared repository

Table A.2.: Overview of work units