A Pattern Catalog for GDPR Compliant Data Protection

Dominik Huth

Chair of Software Engineering for Business Information Systems
Fakultät für Informatik
Technische Universität München
Boltzmannstr. 3, 85748 Garching
dominik.huth@tum.de

Abstract. Today's digital Business Models challenge the concept of privacy of the previous century. As a legislative approach to keep up with the rapid technological change, the European Union has passed the General Data Protection Regulation (GDPR), which will be effective in May 2018. For companies, this implies extensive changes in established processes and new organizational duties. With this work, we aim to develop an understanding and concepts that support an organization - consisting of people, processes and IT systems - in the implementation of privacy regulation. The central result will be patterns, i.e. observed successful approaches of how to address specific requirements of the GDPR.

Keywords: Personal Data; Data Privacy; GDPR; Privacy Engineering; Dependency Model

1 Problem Definition

The rapid technological development has enabled the increasing collection, transmission and storage of user-generated data. By leveraging existing hardware, such as smartphones, the cost for this collection is decreasing as well. This vast amount of data has also led to substantial new developments in Business Models. It has enabled a shift from product development towards information aggregation: Facebook creates no content, Uber does not employ any drivers, Airbnb does not own any real estate [8], and the German long distance bus company Flixbus does not own any buses.

On the other side of the business model innovation, the collected data is often personal information. According to a famous definition by Alan Westin, privacy is *the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others* [22]. When considering the number of services and the amount of personal data that is being collected, with or without the user's explicit knowledge, it is evident that this claim is not completely fulfilled today.

There are multiple options to address the issue of data privacy, the first of them being reliance on the forces of economic markets. [20] argues that economic

J. Ralyté, B. Roelens, and S. Demeyer (Eds.):
PoEM 2017 Doctoral Consortium and Industry Track Papers, pp. 1-8, 2017.

markets without regulation can fail and produce non-optimal outcomes. In [14] we analyze the market for personal mobility data and alternative business models to handle personal data, but still many of today's business models rely on the exchange of free services for broad rights to data that is generated by the user. Another initial nucleus could be the personal motivation of users to protect privacy, but [9] and [15] suggest that individuals do not necessarily know how to value their privacy and are not willing to pay for privacy.

As an approach to address data privacy from a legislative perspective, the European Union (EU) adopted the General Data Protection Regulation (Regulation 2016/679, GDPR) in April 2016 [6]. It was passed after thorough consideration and includes 99 articles and 173 recitals that guide the articles. As a major change, failure to comply with the regulation can be fined with up to four percent of an organization's revenue. This underlines the importance of clearly understanding the requirements, analyzing the status quo of data protection measures, deriving an action plan and implementing the necessary technical and organizational measures to attain compliance with the GDPR. Thus, we set the research goal of this project:

> **Research Goal:** To support the adaptation of privacy legislation in organizations, consisting of people, processes and IT systems.

2 Related Work

For addressing the set research goal, we will first refer to our concept of an Enterprise Model and then analyze existing work using the established Enterprise Model in order to categorize the existing work and determine which of the aspects of the model it addresses. We rely on a simplified Enterprise Architecture model based on [4] and only consider the three layers *Business & Organization*, *Application & Information* and *Infrastructure*. Figure 1 locates the presented work within this Enterprise Architecture model.

Legal advice gives an abstract view of the challenges to be addressed. According to [12], these are the principle of accountability, the creation of records of processing activities, the execution of data protection impact assessments, the implementation of technical measures for protection, the extension of data subject rights and the installation of the role of data protection officer (DPO). [21] analyses the regulation and compares it in detail to Directive 95/46/EC, which it repeals when entering into force. The resulting practical implications are the obligation to specify data needs and usage, to consider conditions for data processing in an international context, to build privacy through data protection by design and default, to demonstrate compliance, to develop processes for dealing with data breaches, to consider possible sanctions when acting, to designate a DPO, to establish processes for providing information to and getting consent from data subjects, to create ways of deleting or transferring an individual's data and to document processing activities.

The paradigm of Privacy by Design was keyed by the Canadian Information Commissioner [5] and encompasses the notion of building privacy into IT systems

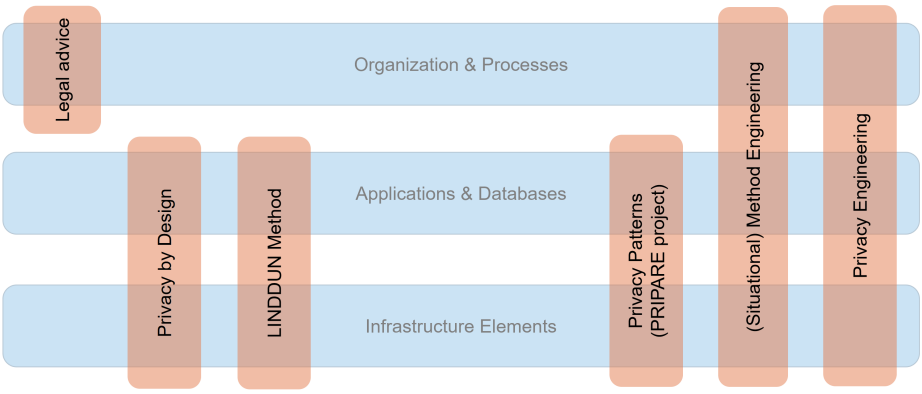

Fig. 1. Related work in Enterprise Architecture model

and processes, but also physical designs. However, the seven proposed principles are rather a goal definition than an instruction of how to implement privacy [19] and a translation of these principles into system requirements is necessary [11]. Nonetheless, there are also critical voices who claim that "privacy cannot be hardcoded" [16].

Method engineering, according to [1], is the discipline of creating methods that support the development of Information Systems and can be compared to Software Engineering for the development of software. When adapting a method for a specific situation and combining individual elements, so-called method chunks, it is called a situational method [13]. An application to the domain of governance, risk and compliance can be found in [7], where method fragments are categorized into five dimensions (conceptual, strategic, organizational, technical, cultural) and assigned to different roles within the enterprise. A recent development is the field of Privacy Engineering [10], which systematically addresses privacy issues during system development.

A notable method for modeling privacy threats is LINDDUN [23], [24]. It is intended to support software developers in identifying and addressing privacy threats early during software development. An extensive privacy threat tree [23] addresses all the security threats that are included in the acronym and gives guidance to software developers.

The operationalization of Privacy and Security by Design was the subject of study of the EU project PRIPARE (Preparing Industry to Privacy-by-design by supporting its Application in Research) [18], which we consider to cover all three layers of the model presented earlier. The main results include a process to address the accountability postulate of the GDPR and so-called patterns for privacy preserving development of software. They distinguish between risk-based (such as LINDDUN, [23]) and goal-oriented approaches, such as the GDPR, which have to be broken down into smaller requirements. These two approaches are

considered to be complimentary, with application of the goal-oriented principle first and detailed system analysis second [18].

Patterns are observed solutions for a given problem [3] that have emerged in three independent instances. Historically, they originate from the domain of software architecture, but have also been applied in Enterprise Architecture Management, a holistic approach to align business and IT [2]. They describe the solution with consequences, known uses, actors and related patterns [3]. The PRIPARE project also created a pattern catalog [17] that includes 26 privacy patterns at a technological level.

The gap we identify is the consideration of holistic patterns that are specifically targeted at compliance with the new GDPR regulation. As it has been passed in 2016 and becomes effective in 2018, there are currently industry efforts under way who aim at achieving compliance by the time the GDPR becomes effective. These efforts have not been investigated so far and have not been analyzed with regards to the existing body of knowledge.

3 Proposed Approach

The goal of this research project is to support the practical application of privacy patterns by developing theory and insight from the current industry efforts. An approach that was designed to ensure theoretical contributions of industry cooperation projects is pattern-based design research [3]. It combines the use of patterns, design theory and the concept of a design theory nexus to propose a four-step process (as depicted in figure 2) consisting of the steps *observe & conceptualize*, *pattern-based theory building*, *solution design and application* and *evaluation and learning*.

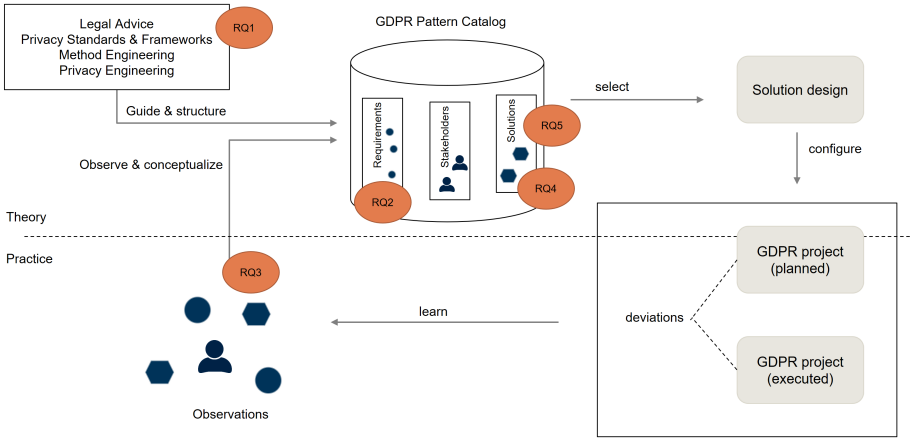

Fig. 2. Research Approach based on [3] with localization of the research questions within the overall process

The EU GDPR is not the first privacy regulation that organizations have to become compliant to. In fact, it repeals EU directive 95/46/EC, which was the basis for prior national privacy legislation. Researchers have addressed these concerns with threat frameworks and situational methods before. There are also similar legislation that requires action within the organization and has been researched before, such as financial or risk regulation. To develop a clear understanding of the existing work and to be able to concisely identify the differences that are invoked specifically by the GDPR, we derive the first research question:

> **RQ1:** Which conceptual frameworks exist that can be instrumented to describe regulatory requirements and the design of possible solutions?

The GDPR as a whole consists of 99 articles, some of which defining terms, some concerning administrative acts regarding the regulation, and some introducing rights for data subjects and obligations for the data controllers and processors. These have to be extracted from the legal text and formalized as elementary requirements. Each of the requirements can then be addressed by methods or techniques that are investigated at a later stage. Thus, we formulate the second research question:

> **RQ2:** What are the elementary requirements of the GDPR and how can they be modeled with the existing concepts?

Since the GDPR is applicable from 2018, there are currently ongoing efforts in organizations with the goal to prepare for the legislation and to implement the required measures, such as the establishment of a data protection officer. It is necessary to observe and describe these efforts and to classify the approaches. Who is responsible? Who is in charge of planning, who is in charge of execution? How are the requirements identified, how are project goals set, which processes are established? Which tools or methods are used to support this effort? In short, we state the third research question as:

> **RQ3:** What is the state of the practice in the implementation of GDPR requirements in Germany and how does the practice instrument theoretical knowledge?

From these practical investigations, we aim to identify common observations from various industry partners. These so-called pattern candidates have to be consolidated into patterns and described using the theoretical background that was identified before. It will be analyzed how successful the application of each pattern has been, since commonality does not necessarily imply quality. As the fourth research question, define:

> **RQ4:** Which successful patterns can be identified for fulfilling the elementary requirements of the GDPR?

There might be different patterns that address the same elementary requirement. At the same time, there might be patterns that require the application of another pattern in order to be executed successfully. In the last step, we aim to analyze dependencies among patterns. One such dependency could be the reduced need for on site data security when transferring personal data to cloud services, but a rigorous selection process for the cloud provider to comply with the data security requirements. This leads to the last research question:

> **RQ5:** How are solution options interrelated with each other? Which solutions are independent, which require other actions, and which replace other solution options?

Acknowledgment

This dissertation project is supervised by Prof. Dr. Florian Matthes at the chair for Software Engineering for Business Information Systems, TU München. It is part of the TUM Living Lab Connected Mobility (TUM LLCM) project and has been funded by the Bavarian Ministry of Economic Affairs and Media, Energy and Technology (StMWi) through the Center Digitisation.Bavaria, an initiative of the Bavarian State Government.

References

1. Brinkkemper, S.: Method engineering: Engineering of information systems development methods and tools. Information and Software Technology 38(4 SPEC. ISS.), 275–280 (1996)
2. Buckl, S., Ernst, A.M., Lankes, J., Matthes, F.: Enterprise Architecture Management Pattern Catalog. Sebis, TU München (February), 322 (2008)
3. Buckl, S., Matthes, F., Schneider, A.W., Schweda, C.M.: Pattern-Based Design Research An Iterative Research Method Balancing Rigor and Relevance. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 7939 LNCS, pp. 73–87 (2013)
4. Buckl, S., Matthes, F., Schweda, C.M.: Conceptual models for cross-cutting aspects in enterprise architecture modeling. Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC pp. 245–252 (2010)
5. Cavoukian, A.: Privacy by Design. IEEE Technology and Society (Winter 2012), 18–19 (2012)
6. European Union: Regulation 2016/679 of the European parliament and the Council of the European Union (2016)
7. Gericke, A., Fill, H.G., Karagiannis, D., Winter, R.: Situational method engineering for governance, risk and compliance information systems. Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology - DESRIST '09 p. 1 (2009)
8. Goodwin, T.: The Battle Is For The Customer Interface (2015), https://techcrunch.com/2015/03/03/in-the-age-of-disintermediation-the-battle-is-all-for-the-customer-interface/

9. Grossklags, J., Hall, S., Acquisti, A.: When 25 Cents is too much : An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information. Information Security pp. 7–8 (2007)
10. Gürses, S., Del Alamo, J.M.: Privacy Engineering: Shaping an Emerging Field of Research and Practice. IEEE Security and Privacy 14(2), 40–46 (2016)
11. Gürses, S., Troncoso, C., Diaz, C.: Engineering Privacy by Design. Computers, Privacy & Data Protection 317(5842), 1178–1179 (2011)
12. Hamann, C.: Europäische Datenschutz-Grundverordnung - neue Organisationspflichten für Unternehmen. Betriebs-Berater pp. 1090–1097 (2017)
13. Henderson-Sellers, B., Ralyté, J.: Situational Method Engineering: State-of-the-Art Review. Journal of Universal Computer Science 16(3), 424–478 (2010)
14. Huth, D., Faber, A., Matthes, F.: Personal Mobility Data: Threats and Opportunities. In: mobilTUM conference (2017)
15. Jentzsch, N.: State-of-the-art of the Economics of Cyber-security and Privacy. Tech. rep. (2016)
16. Koops, B.J., Leenes, R.: Privacy regulation cannot be hardcoded. A critical comment on the 'privacy by design' provision in data-protection law. International Review of Law, Computers and Technology 28(2), 159–171 (2014)
17. Le, D., Inria, M., Trilateral, I.K., María, J.: PRIPARE Privacy- and Security-by-Design Methodology Handbook (2015), http://pripareproject.eu/
18. Notario, N., Crespo, A., Martin, Y.S., Del Alamo, J.M., Metayer, D.L., Antignac, T., Kung, A., Kroener, I., Wright, D.: PRIPARE: Integrating privacy best practices into a privacy engineering methodology. Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015 pp. 151–158 (2015)
19. Rubinstein, I.S., Good, N.: Privacy by Design: A Counterfactual Analysis of Google and Facebook Privacy Incidents. Berkeley Technology Law Journal 28(2), 1333–1413 (2013)
20. Stiglitz, J.E.: The price of inequality: How today's divided society endangers our future. WW Norton & Company (2012)
21. Tikkinen-Piri, C., Rohunen, A., Markkula, J.: EU General Data Protection Regulation: Changes and implications for personal data collecting companies. Computer Law and Security Review (2017) (2017)
22. Westin, A.F.: Privacy and Freedom. Ig Publishing (2015)
23. Wuyts, K.: LIND(D)UN privacy threat tree catalog (September) (2014)
24. Wuyts, K., Joosen, W.: LINDDUN privacy threat modeling: a tutorial. Tech. Rep. July (2015)