Assessing and Improving Compliance to Privacy Regulations in Business Processes

Jake Tom
Supervisors: Raimundas Matulevičius and Peeter Laud

Institute of Computer Science, J. Liivi 2
University of Tartu - 50409, Estonia
jaketom@ut.ee

Abstract. Personal data privacy is emerging as an area of significant interest in recent years as more regions adopt data protection regulations in an effort to protect the rights of consumers. Organizations are in need of novel approaches to upgrade their business processes to make them compliant with these regulations. However, current research towards privacy management is tailored towards the development of new processes from the ground up and little is done to address the assessment and improvement of existing processes. This PhD thesis proposes two artifact-based techniques to evaluate compliance to privacy regulations and capture privacy properties in business processes within the context of a new privacy management approach.

Keywords: Business process management, Privacy management, BPMN, Business process compliance

1 Introduction

Personal data privacy has garnered increasing attention over recent years. Frequent and large scale breaches of sensitive data around the world [13] indicate that organizations have not yet fully prioritized private data protection. The last decade has seen the introduction and strengthening of data protection laws in several countries [19]. In Europe, the General Data Protection Regulation (GDPR) [1] will replace the Data Protection Directive (Directive 95/46/EC) [11] from from May, 2018. Such regulations levy heavy penalties upon organizations that fail to adequately ensure security and offer transparency into their data processing activities. In addition, the EU-US Privacy Shield [12] enforces privacy guidelines upon companies in the USA that process the data of European citizens. Organizations that intend to implement privacy controls through either business process modifications or the implementation of privacy-enhancing technologies (PETs) require techniques to assess and improve the current state of their processes. A study of research in this area reveals several approaches and methods to assist the development of new processes that support desired privacy guarantees (e.g. [5][2][20][16]). However, further work is required to develop approaches aimed at the improvement of *existing* processes.

The objective of the thesis is to develop the foundations of an approach for organizational privacy management with a focus on artifact-based Business Process Model and Notation (BPMN) analysis techniques. For the sake of practical relevance, this approach will be developed and validated with reference to the GDPR. The primary research objective can be formulated as follows: *How can we assess and improve business process compliance to privacy regulations?* The results of this thesis - privacy extensions to BPMN and a meta-model of the GDPR along with their usage within the intended approach, are aimed towards the assessment and improvement of processes and information systems within existing organizations. It should also be noted that the specific domain of this thesis is *privacy* management and not *information security.* While there is an overlap between the two domains, we view security as defined under the *Confidentiality-Integrity-Availability triad* for expressing security objectives. On the other hand, privacy deals with how sensitive information is obtained and processed. Implementing security can mitigate certain privacy risks but may not be the solution for them all.

This paper describes the research goals of this PhD thesis and highlights its contributions so far. It is structured as follows: Section 2 describes the intended approach and the primary research objective is broken down into five research questions. Section 3 goes into the contributions made so far within the context of the research methodology employed. Section 4 presents the focus areas for the literature review and an overview of its results so far.

2 Research Objectives

A proposed view of the approach is illustrated in Fig. 1. The figure is composed of generic steps with their instantiations within this thesis described in smaller boxes. For example, GDPR is an instantiation of Privacy Policy. The approach is broadly composed of two phases:

1. **Assessment:** The goal of this phase is to identify areas of non-compliance to the GDPR to arrive at a set of requirements to achieve compliance. This identification is based on the comparison of two kinds of UML diagrams - (*i*) a meta-model extracted from the contents of the GDPR that illustrate compliance rules and (*ii*) existing business processes of the organization under evaluation documented in BPMN that are then converted into class diagrams to provide a policy-oriented view of the processes. The GDPR meta-model and its comparison techniques form one side of the thesis contribution.
2. **Improvement:** The goal of this phase is to produce process models that are verifiably compliant to the GDPR. To express this, the output of the assessment phase is converted into a set of requirements. These requirements will fall under two categories - (*i*) technological improvements and (*ii*) process improvements. While process improvements are simply revisions to the existing business processes in standard BPMN, technological improvements that can be addressed using PETs cannot yet be adequately captured. This

provides the motivation for the development of extensions to the BPMN language, labeled Privacy-Enhanced BPMN (PE-BPMN) which is another contribution of this thesis.

To develop such an approach and its components, the following research questions (RQ) have to be addressed:

- RQ1: What is the state of the art in business process assessment and improvement with regards to privacy regulation compliance?
- RQ2: What is a conceptual representation (meta-model) of the GDPR?
- RQ3: How can the conceptual representation be validated?
- RQ4: How can the BPMN language be extended to capture PETs?
- RQ5: How can the PE-BPMN language extensions be validated?

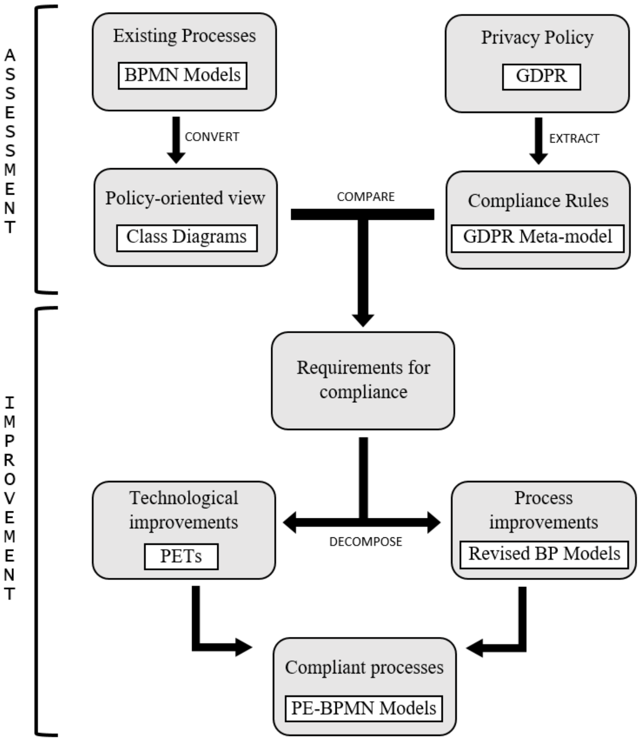

Fig. 1. Proposed approach for assessment and improvement of an existing information system towards privacy regulations

3 Research Methodology and Current Status

The work accomplished during this first year of study has been primarily focused on the development and validation of privacy extensions to BPMN (RQ4, RQ5) and secondarily, on evaluating the state of the art (RQ1). Preliminary development of the GDPR meta-model is currently underway at the University of Tartu which will soon be handed over for completion as part of this thesis (RQ2, RQ3).

3.1 Design Science

PE-BPMN [4] has been developed in line with the design science research guidelines proposed by Hevner et al. [7]. The following list describes the state of the thesis and future plans within the context of each guideline:

1. **Design as an artifact**: The techniques being developed in this thesis are based on the production of specific artifacts in the form of process models with an extended BPMN syntax and meta-models drawn in UML. The PE-BPMN language extension provides an approach (method) and syntax (model) to capture PETs at multiple levels of abstraction in process models. The work on PE-BPMN is summarized in Section 3.2.
2. **Problem relevance**: The literature review in Section 4 has identified a relative shortage of techniques to assess and improve existing business processes. Additionally, while BPMN is particularly suited to capture the state of processes around an existing information system, it is not suitable for capturing technological privacy properties of the same. To be relevant in the privacy domain, the BPMN standard requires extension.
3. **Design evaluation**: A prototype modeler for PE-BPMN extensions and subsequent analysis has been implemented and is currently being validated on case studies from the DARPA Brandeis program [22].
4. **Research contributions**: The contribution of the thesis so far is the PE-BPMN language extension. It is intended that the GDPR meta-model and its comparison methods will be completed and consolidated along with PE-BPMN into the proposed approach in Fig. 1.
5. **Research rigor**: PE-BPMN is based on BPMN and the GDPR meta-model is based on UML which are accepted standards. PE-BPMN extends the abstract and concrete syntax and semantics of BPMN 2.0. The addition of semantics is based on existing PET taxonomies, the abstract syntax is an extension of the existing model and the concrete syntax is based on best practices of UML.
6. **Design as a search process**: The development of PE-BPMN is founded on a review of the gaps identified in the current state of privacy modeling in business processes (see Section 4.4). After its initial version was published, the language was extended and refined in a later iteration that added support for modeling additional PETs as well as refinements to its foundations.
7. **Communication of research**: The paper on PE-BPMN [4] is presented from a technical viewpoint directed at system analysts. While higher level

motivations for managerial audiences are provided to some degree, they are not the focus. However, the broader scope of the thesis (i.e. the approach) will be aimed at addressing concerns of managerial audiences as well.

3.2 Privacy-Enhanced BPMN

PE-BPMN, originally developed for another project aims to address the gap of expressing technological privacy safeguards in business process models. It is also directly applicable to solving related steps of the approach in Fig. 1. The paper on PE-BPMN [4] proposes a multi-leveled model of PET abstraction that, on one level, views PETs in terms of their general targets, for example, data confidentiality or user anonymity. This abstraction is expressed in process models using language extensions called *generic stereotypes*. A process model defined in terms of generic stereotypes can then be used to advance to the next level of abstraction where these generic stereotypes are instantiated with specific PETs using *concrete stereotypes*. In Fig. 2(a) we see how individual tasks related to a mechanism that ensures data confidentiality are expressed with generic stereotypes. We assume application of a protection mechanism (via ProtectConfidentiality), computation on the protected data with a public input (via PETComputation and finally, removal of the protection mechanism (via OpenConfidentiality. At this stage, our process under consideration has achieved the goal of ensuring data confidentiality. It is possible to take this further by instantiating the generic stereotypes with concrete stereotypes as shown in Fig. 2(b) which describes the application of public key encryption through equivalent tasks.

Finally, it proposes a method to qualitatively analyze disclosures of information that can potentially occur along a process described in PE-BPMN through a set of *disclosure matrices*. A prototype demonstrating modeling with privacy extensions, syntax verification and automated information disclosure analysis is also presented. Pullonen et al. [10] introduced the first version of PE-BPMN and described the PE-BPMN syntax and PET selection method. The contribution of this thesis to PE-BPMN includes: (*i*) The multi-leveled model of PET abstraction, (*ii*) the inclusion of additional PETs into the PE-BPMN syntax and improvement of existing syntax and (*iii*) the information disclosure analysis method.

4 Background and Related Work

Based on the goals and steps of the approach described in Fig. 1, the following research areas have been identified for the literature review:

4.1 Privacy Regulations and their Implementation

DLA Piper [19] has developed a web application that provides an overview of the strength of data protection regulations around the world. A review of works

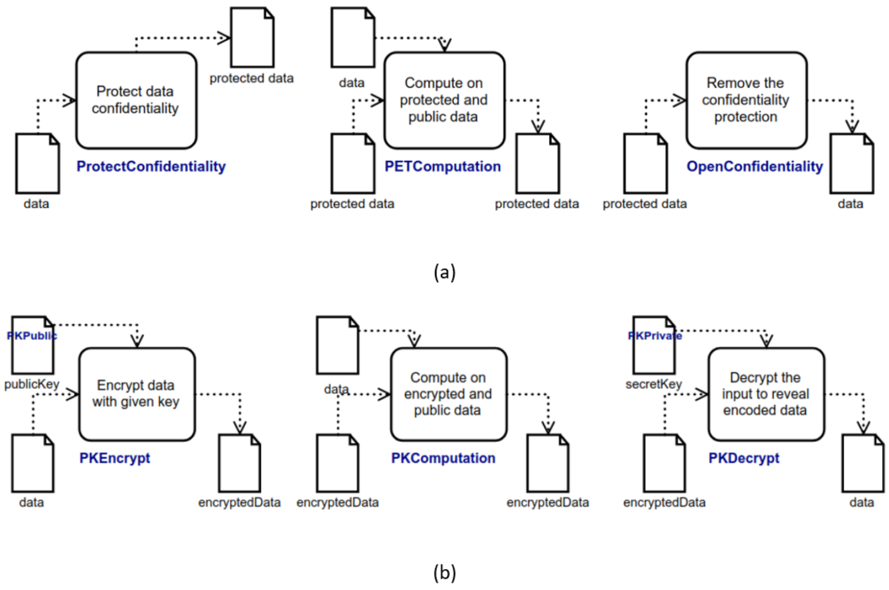

Fig. 2. Generic and concrete stereotypes in PE-BPMN

related to evaluation and implementation of these regulations is currently underway. It is focused particularly on those regions recognized as having *robust* or *heavy* data protection laws as described in the application.

Outside EU: The Personal Data Protection Act (PDPA) introduced in Malaysia in 2010 is a collection of categorized principles [3] that regulate commercial usage of personal data. The authors in [18] present the results of a case study aimed at identifying the impact of the PET adoption after PDPA enforcement on employee work performance. The results suggest that a holistic evaluation of business process compliance to privacy regulations would include techniques to measure PET adoption among employees and compliance to restructured business processes. However, the discussion on techniques to manage the same is limited.

GDPR: In [5], the authors introduce an extended version of Socio-Technical Security modeling language (STS-ml) to capture and verify the social aspects of the GDPR such as the relationship between employers and employees and consent. However, STS-ml does not yet capture technical aspects of the GDPR such as data security measures and is most suitable for systems being developed. A meta-model to illustrate key relationships between entities defined in the GDPR is proposed by the authors of [2]. The meta-model is intended to be used by the designers of e-services for the development of adequate Privacy Level Agreements (PLAs) that demonstrate compliance to the GDPR.

4.2 General Approaches for Regulation-Compliant System Design

It is relevant to also consider work done with regards to IS design and compliance to regulations that do not relate to privacy. At a more abstract level, any regulation can be viewed as a set of constraints imposed upon a process. An analysis of related work into regulation compliance outside privacy may yield useful insights when studied from this perspective.

In [20], the authors introduce a method that provides a high-level set of guidelines to select appropriate measures to ensure regulatory compliance. A framework to elicit security requirements from laws and regulations is proposed by the authors of [16]. While [20] is not specifically applied to GDPR and speaks in terms of implementing *any* regulation, its high-level perspective provides some guidance to IS designers conforming to legal requirements. The framework in [16] could be adapted for some privacy requirements such as the concept of consent in GDPR. The approach uses goal modeling to extract security requirements which are then translated to a secure system design with Model Based Security Engineering.

4.3 Business Process Compliance Checking

Business process compliance (BPC) checking has received substantial research attention and there are several approaches towards it based on graph pattern matching, computation tree logic and other computer science concepts. Pattern based approaches to compliance checking are presented by the authors of [15] and [6]. In [15], the authors illustrate a method to extract a catalogue of compliance patterns from a regulation that can then be compared against business processes using a graph pattern-based compliance checking approach. In [6] the authors use security risk-oriented patterns applied to the ISO27001:2013 security standard to check as well as improve process compliance.

4.4 Privacy Modeling in Business Processes

There has been significant research into modeling security and privacy in BPMN. BPMN has been adapted to the domain of security risk management and security modeling extensions have been proposed by the authors of [14], [17] and [21]. While these are applicable to security modeling which covers some aspects of privacy management, they are not designed specifically for privacy. Privacy-aware BPMN and syntax extensions to capture specific aspects of privacy concerns like consent and access control are presented in [9] and [8]. However, neither of them address how to capture PETs in BPMN and identify privacy losses along the process chain.

5 Concluding Remarks

In this paper, we provided the motivation, scope and approach for the PhD thesis. While the focus of this thesis will be in the context of the GDPR, the goal is to generalize the approach for achieving compliance with other emerging privacy regulations as well. The main research question is broken down into sub-questions that address the development and validation of the privacy management techniques proposed in the thesis. The applied research methodology is introduced and explained along with current progress. Then, the state of the art is reviewed with respect to four identified areas of focus to motivate the thesis and position it with regards to existing research.

As for future work, while a preliminary evaluation of the state of the art (RQ1) has been done, the review needs to be completed, potentially using a systematic literature review method. The GDPR meta-model (RQ2, RQ3) remains to be developed and then validated on case studies within the DARPA Brandeis [22] program - a collective research effort aimed at the development of techniques to preserve privacy while analyzing large amounts of sensitive information across varied industries. We plan to elicit feedback on the artifacts and the general approach from program managers through direct interviews and qualitative surveys from other program members.

References

1. EU General Data Protection Regulation, http://data.consilium.europa.eu/doc/document/ST-5419-2016-INIT/en/pdf
2. Angelopoulos, K., Diamantopoulou, V., Mouratidis, H., Pavlidis, M.: A Metamodel for GDPR-based Privacy Level Agreements. ER Forum/Demos (2017)
3. Personal Data Protection Act 2010, Malaysia, http://www.agc.gov.my/agcportal/uploads/files/Publications/LOM/EN/Act%20709%2014%206%202016.pdf
4. Pullonen, P., Tom, J., Matulevičius, R., Toots, A.: Privacy-Enhanced BPMN: A Multi-Level Approach to Information Disclosure Analysis, Submitted for publication (2018)
5. Giorgini, P., Robol, M., Salnitri, M.: Toward GDPR-Compliant Socio-Technical Systems: Modeling Language and Reasoning Framework. PoEM (2017)
6. Alaküla, M. L., Matulevičius, R.: An Experience Report of Improving Business Process Compliance Using Security Risk-Oriented Patterns. In IFIP Working Conference on The Practice of Enterprise Modeling (pp. 271-285). Springer (2015)
7. Von Alan, R. H., March, S. T., Park, J., Ram, S.: Design science in information systems research. MIS quarterly, 28(1), 75-105. (2004)
8. Brucker, A. D., Hang, I., Lückemeyer, G., Ruparel, R.: SecureBPMN: Modeling and enforcing access control requirements in business processes. In Proceedings of the 17th ACM symposium on Access Control Models and Technologies (pp. 123-126). ACM (2012).
9. Labda, W., Mehandjiev, N., Sampaio, P.: Modeling of privacy-aware business processes in BPMN to protect personal data. In Proceedings of the 29th Annual ACM Symposium on Applied Computing (pp. 1399-1405). ACM (2014).

10. Pullonen, P., Matulevičius, R., Bogdanov, D.: PE-BPMN: Privacy-Enhanced Business Process Model and Notation. In International Conference on Business Process Management (pp. 40-56). Springer (2017)
11. Directive 95/46/EC of the European Parliament, https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A31995L0046
12. Weiss, M. A., Archick, K.: US-EU data privacy: from safe harbor to privacy shield. (2016)
13. Visualization of data breaches around the world, http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/
14. Rodríguez, A., Fernández-Medina, E., Piattini, M.: A BPMN extension for the modeling of security requirements in business processes. IEICE transactions on information and systems, 90(4), 745-752 (2007).
15. Delfmann, P., Hübers, M.: Towards Supporting Business Process Compliance Checking with Compliance Pattern Catalogues-A Financial Industry Case Study. Enterprise Modelling and Information Systems Architectures–International Journal of Conceptual Modeling, 10(1), 67-88. (2015)
16. Islam, S., Mouratidis, H., Jürjens, J.: A framework to support alignment of secure software engineering with legal regulations. Software and Systems Modeling, 10(3), 369-394. (2011)
17. Cherdantseva, Y., Hilton, J., Rana, O.: Towards SecureBPMN-Aligning BPMN with the information assurance and security domain. In International Workshop on Business Process Modeling Notation (pp. 107-115). Springer (2012).
18. Gan, M. F., Chua, H. N., Wong, S. F.: Personal Data Protection Act Enforcement with PETs Adoption: An Exploratory Study on Employees' Working Process Change. In IT Convergence and Security 2017 (pp. 193-202). Springer. (2018)
19. DLA Piper Data Protection, https://www.dlapiperdataprotection.com/
20. Knackstedt, R., Braeuer, S., Heddier, M., Becker, J.: Integrating Regulatory Requirements into Information Systems Design and Implementation. (2014)
21. Altuhhova, O., Matulevičius, R., Ahmed, N.: An extension of business process model and notation for security risk management. International Journal of Information System Modeling and Design (IJISMD), 4(4), 93-113 (2013).
22. DARPA Brandeis, https://www.darpa.mil/program/brandeis