Using Content Analysis for Privacy Requirement Extraction and Policy Formalization

Stefanie Rinderle-Ma[1] Zhendong Ma[2] Bernhard Madlmayr[3]

Abstract:

Privacy in cyberspace is a major concern nowadays and enterprises are required to comply with existing privacy regulations and ensure a certain level of privacy for societal and user acceptance. Privacy is also a multidisciplinary and mercury concept, which makes it challenging to define clear privacy requirements and policies to facilitate compliance check and enforcement at the technical level. This paper investigates the potential of using knowledge engineering approaches to transform legal documents to actionable business process models through the extraction of privacy requirements and formalization of privacy policies. The paper features two contributions: A literature review of existing privacy engineering approaches shows that semi-automatic support for extracting and modeling privacy policies from textual documents is often missing. A case study applying content analysis to five guideline documents on implementing privacy-preserving video surveillance systems yields promising first results towards a methodology on semi-automatic extraction and formalization of privacy policies using knowledge engineering approaches.

1 Introduction

Privacy in cyberspace has become a major concern nowadays and enterprises are obliged to ensure a certain level of privacy as demanded by law [Bi08] and society [In97]. As a multidisciplinary topic, privacy is influenced by social, legal, and technical factors. The ambiguity of privacy definition, the difference in privacy perception, and the fast changing technological landscape make it very challenging for an enterprise to keep up with the privacy stipulations and expectations.

Since business processes capture activities at both human and system level within an enterprise, they often serve as the basis for privacy checks [AM14], i.e., it can be analyzed how (daily) routines in an enterprise are conducted with respect to privacy requirements and policies. As for security [Le14], business processes can be either checked for their compliance with privacy requirements (privacy ensuring) or they can be used to implement privacy policy (privacy enforcement).

Verification of privacy requirements over a system or business process can be conducted by, for example, model checking. Such approaches require the formal

[1] University of Vienna, Faculty of Computer Science, Austria, stefanie.rinderle-ma@univie.ac.at
[2] Austrian Institute of Technology, Digital Safety and Security Department, Austria, zhendong.ma@ait.ac.at
[3] University of Vienna, Faculty of Computer Science, Austria, bernhardmad@gmail.com

representation of privacy requirements as structured privacy policies. However, privacy requirements are often originated from legal documents [BA08], i.e., in natural language and hence in an unstructured way, which is subject to interpretations (e.g. by law professionals) and lacks clarity at the technical level. Often these documents are vague and generic. For example, the clause "to provide adequate privacy protection" might be sufficient for lawyers but way too ambiguous for system designers and engineers to implement. Therefore, engineers have great difficulties to understand and interpret such documents and translate them into practical technical privacy-preserving designs and practices. A recent multi-disciplinary approach to address privacy in surveillance systems found out the main difficulty in designing privacy-preserving systems is the ambiguity from the knowledge gap between technical and non-technical world [Ma14]. Hence, the ability to extract the relevant information from privacy documents and provide the extracted information in a structured (formalized) and unambiguous way (i.e. understandable and actionable technical specifications) can be very beneficiary in designing and developing privacy-preserving ICT systems. As extraction of privacy requirements can be tedious and error prone when done manually, it would be useful to employ techniques to at least derive candidates for privacy policies in a semi-automatic way. Here, we advocate the investigation of knowledge engineering techniques such as content analysis [St06] or text mining [AZ12] for their suitability to extract privacy requirements from legal documents in a semi-automatic way. For clarification of terminology, throughout the paper, we denote as privacy requirements the privacy-related information within the textual documents which are first extracted and then modeled or formalized as privacy policies.

In summary, the paper addresses two questions:

1. How to utilize knowledge engineering techniques for extracting privacy requirements from text in legal documents in a semi-automatic manner?
2. How to model the extracted information as structured privacy policies?

Many approaches have addressed privacy requirement engineering, e.g., [ANM10, BM10, BA08, Ch08, Ch11, Co07, De11, Gr12, Gü05, He03, KBG11, KS85, Le06, LYM03, MdAY14, MPZ05, MMZ11, MMZ08, PDG14, RGK13, Ri14, dRAF05]. However, as it will be shown, most of these approaches are manual or do not consider textual input. In order to underpin this claim and provide an overview of existing privacy requirements engineering approaches, the paper provides a literature review in Sect. 2, guided by the questions: What knowledge engineering technique is used? What are source and target format for privacy requirement engineering? Section 3 presents the results of applying content analysis to five documents for implementing privacy-preserving video surveillance systems. The result is a first suggestion of how knowledge engineering techniques can be utilized for privacy policy extraction and formalization and is presented in Sect. 4. As such, the proposal can be used in almost any of the existing approaches. It also discusses next steps in validation and transferability of the methodology.

2 Literature review

A literature review was conducted in order to obtain an overview of existing approaches for elicitation of privacy requirements. Specifically interesting in the context of this paper are approaches that utilize knowledge engineering techniques. The guidelines for conducting a systematic literature review were taken up in a simplified form from [Ki09].

At first, the following keywords were selected for the horizontal literature search:

`policy engineering, privacy policy engineering, privacy requirements engineering, security policy engineering, security requirements engineering, policy elicitation, privacy policy elicitation, security policy elicitation, privacy requirement elicitation, security requirement elicitation.`

As search method, the keywords were used as title search in google scholar[4] from 22 – 24 Oct 2014 as well as on 27 Oct 2014, excluding patents and citations. Table 2 shows the results of the horizontal literature search, i.e., the first column contains the keywords and the second column the number of papers found.

Keywords	*# Hits*	*Selection words*	*# selections*
policy engineering	505	privacy, security	9
privacy policy engineering	3		0 (overlap with policy engineering)
privacy requirement engineering	26	focus: privacy requirements	21 (1 overlap, 1 not available, 1 duplicate
security policy engineering	11	focus: security policies	0 (overlap with policy engineering and selection criteria)
security requirement engineering	120	64 focus: security requirements	(duplicates, unavailable, journal extension)
policy elicitation	14	privacy	0
privacy / security policy elicitation	0		0
privacy requirement elicitation	1	0	
security requirement elicitation	6		5
overall vertical	686		99

Table 1: Results of vertical literature search

Within a primary selection process, each paper title was checked for the covered area. For each keyword, selection words were defined, i.e., those words that specify and restrict the found papers for the specific area of privacy and security policy elicitation. Take, for example, keyword `policy engineering` which results in 505 found papers during the primary search. However, policy engineering might also refer to other policies than privacy and security policies. Hence, the found papers

[4] `scholar.google.com`

were scanned through their title and abstract whether or not they refer to the privacy and security area. resulting in 9 papers. On top of these content-related selection criteria, general selection criteria such as availability of paper, written in English, and scientific paper were applied.

The result of the vertical literature search, i.e., a list of the primarily selected number of 99 publications can be found at[5]. The primary literature list was reduced within an expert discussion based on the following criteria: lack of focus on privacy, model-driven approaches, lack of linkage to knowledge / requirements engineering methods. In addition, similar approaches, specifically from the same group of authors on the same topics were aggregated by considering a selection of their papers.

The reduction resulted in 27 papers. Based on these papers, snowballing was conducted, resulting in $27+6+5=38$ papers[6]. In addition, snowballing led to a new keyword, i.e., *extraction* which was combined with keywords *privacy policy* and *privacy requirement* when conducting another round of vertical search. However, the keywords did not yield any results.

These core papers were analyzed along the following research questions:

1. Is a knowledge engineering method suggested / applied? If yes, which ones?
2. Which sources are used?
3. What is the target format?

The first question was used as a reduction criteria, i.e., if an approach was neither proposing nor applying a knowledge engineering method it was excluded from further analysis. Out of the 38 papers, 25 approaches were found during horizontal and vertical search that suggest usage of knowledge engineering method(s): [AM14, AE00, ANM10, BM10, BVA06, BA08, Ch08, Ch11, Co07, De11, Gr12, Gü05, He03, KBG11, KS85, Le06, LYM03, MdAY14, MPZ05, MMZ11, MMZ08, PDG14, RGK13, Ri14, dRAF05]. 4 papers provide an overview of existing security requirements engineering / modeling / elicitation techniques themselves [El11, Fa10, Me10, SK12] and were hence not considered in the further analysis. The remaining 9 papers did not suggest any elicitation method and were hence discarded from further investigation.

With respect to the research questions set out in the introduction, the 25 resulting papers were analyzed whether they (a) employ a manual or (semi-)automatic engineering technique, (b) take text as input format, and (c) produce an output format that can be utilized for business process compliance checking. Results:

1. The only approach (from 1985) that suggests a (semi-)automatic approach is [KS85]. All other approaches propose, extend, or employ manual methods.

[5] `http://cs.univie.ac.at/fileadmin/user_upload/fak_informatik/RG_WST/documents/Rinderle-Ma/PrimarySearch_SEC15_MaRi.pdf`

[6] Again, papers of the same group were considered in an aggregated way, i.e., with the most current or comprehensive paper.

Some of these methods are tool-supported, i.e., PRET [MMZ08] and the method proposed in [Gr12] supported by Objectiver. It is worth taking a look what is exactly supported by tools, the extraction or the modeling or both.

2. Several approaches extract privacy requirements from textual sources, i.e., PRET [MMZ08], [BA08] specifically for HIPAA, [BVA06] in form of Unrestricted Natural Language Statements (UNLR), using Secure Tropos on law by [MPZ05], and specifically analyzing DITSCAP [Le06]. The other approaches range from business process models [AM14] and stakeholder knowledge [Gü05, De11, dRAF05, KBG11, ANM10], to requirements [Ch11, PDG14]. The other approaches remain either unspecific, e.g., by stating "various" information sources or information systems.

3. Regarding the last question of the target format, most approaches provide some structured format, i.e., requirements, policies or rules, patterns, XML, and ontologies. By contrast, [AM14, MdAY14] have text as target format.

Overall, none of the approaches fits the requirements set out in the introduction, i.e., provides a (semi-)automatic methodology for extracting structured privacy requirements from text. Overall, most of the approaches aim at comprehensive methodology for guiding the entire engineering process from identifying relevant documents or other artifacts until privacy policies are specified. In particular, most of the approaches include the users, e.g., domain experts. This is for sure an important issue. This paper does not suggest to replace an overall methodology and inclusion of users, but aims at support of ONE specific step of the overall methodology, i.e., the extraction and formalization step as discussed in the next section.

3 Preliminary study: Content analysis

Methods for the extraction of information from text are proposed and applied in different areas. Knowledge Engineering [SBF98] deals more generally with the construction of Knowledge-based Systems and comprises the extraction of information as one step next to other steps such as modeling and derivation. Information extraction also plays a crucial role in web environments where often (semi-)structured data is the basis to extraction [Sa08]. Specifically geared towards information extraction from text are, for example, text mining [AZ12], qualitative content analysis [St06], and Natural Language Processing (NLP) [Fr11].

The purpose of this preliminary study was to evaluate the suitability of knowledge engineering methods based on the example of content analysis for the extraction of privacy requirements from text or unstructured data such as regulatory documents or laws. Qualitative Content Analysis (QCA) has a manual component as documents must be unitized, categorized, and coded. Support is provided by tools such as QDA Miner[7] and Atlas.ti[8]. Particular advantages of QCA are reliability and maintain-

[7] http://provalisresearch.com/products/qualitative-data-analysis-software/
[8] http://atlasti.com/

ability. We have gathered positive experience with QCA in deriving the teaching process at the University of Vienna based on interview transcripts [KRM11].

The case study was focused on privacy in video surveillance. As a widely deployed technology for protecting humans and property in public and private spaces, video surveillance has always been a privacy concern and a subject of debate. Moreover, due to technological advancement, video surveillance systems are becoming more powerful and hence more privacy-intrusive, in which multiple information sources can be aggregated and video images can be analyzed automatically in large scales. Due to the privacy concern around video surveillance, a large amount of regulations and guidelines exist. However, similar to many other privacy-related documents, they often lack the clarity and precision that are important for compliance check and system design at the technical level. The case study was based on the following guidelines on implementing privacy-preserving video surveillance systems.

1. *The EDPS Video-Surveillance Guidelines* contains guidelines "for European institutions and bodies on how to design and operate their video-surveillance system"[9].
2. *OECD Privacy Guidelines* "govern[...] the protection of privacy and transborder flows of personal data"[10].
3. *Guidelines for Public Video Surveillance* provided by an initiative for protecting "civil liberties" in America[11].
4. *Data protection and privacy ethical guidelines*[12] address data and privacy issues in the context of EU FP7 projects.
5. *Operational Guidance on taking account of Fundamental Rights in Commission Impact Assessments*[13] issued by the European Commission.

Due to experience and availability we opted for using QDA Miner. The QCA was conducted by one analyst. In a first round, the analyst read through the above documents and obtained a general overview of the content and the relation between the documents.

As the target format is process-structured, the two basic categories to be extracted from the text are `Actors` and `Activities`. Focusing on `Actors` and `Activities` as a first step corresponds to the idea of analyzing sentences finding verbs and objects as featured in, e.g., Friedrich et al. [Fr11] extracting actors and actions from sentences.

In a second round, the analyst read through the documents again highlighting relevant phrases from the document that fit into those two categories. Examples for

[9] `https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Supervision/Guidelines/10-03-17_Video-surveillance_Guidelines_EN.pdf`

[10] `http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm`

[11] `http://www.constitutionproject.org/wp-content/uploads/2012/09/54.pdf`

[12] `http://ec.europa.eu/research/participants/data/ref/fp7/89827/privacy_en.pdf`

[13] `http://ec.europa.eu/justice/fundamental-rights/files/operational-guidance_en.pdf`

actors are `Government`, `Child`, and `Employee` and for processes `Impact assessment`, `Monitor Area`, and `Install System`.

Based on evaluating statistics on word frequencies, the documents were coded along the categories `Actors` and `Activities`. The code base for QDA Miner can be found here: `http://cs.univie.ac.at/fileadmin/user_upload/fak_informatik/RG_WST/documents/Rinderle-Ma/Privacy.ppj`. Figure 1 shows the code book for the five documents. Note that codes abstract from different terms and phrases in the documents. One example is activity `Consultation` which represent, for example, phrase `Consult DPO`. The coding was aggregated and reviewed several times in order to overcome errors and to provide the coding at an adequate abstractions level.

Figure 1: Coded Actor and Activity Hierarchy (Code Book Produced Using QDA Miner, Optimized Presentation)

Let us first take a look at the `Actors`. Here different categories can be organized into sub-categories, e.g., category `Civilian` has sub-categories `Individuals`, `Group`, `Applicant`, `Community`, and `Family`. The code hierarchies for `Actor` can be transferred and modeled as, for example, organigram in order to connect the organizational information with the processes to be derived. The model shown in Fig. 2 was modeled using Signavio.

Category `Actor` was used during QCA. Organigrams usually offer more meta model elements to capture organizational information such as `Roles`, `Organizational Units`, and `Persons`. Hence, in principle, two design decisions can be made. Either more categories are considered during QCA or the categories that are coded are mapped onto different meta model elements. In this example, the second option was chosen, i.e., category `Actor` was mapped onto `Roles`, `Organizational Units`, and `Persons`. The mapping was done manually.

At the end of this step, an organigram exists that captures the information from all documents and can be directly used in processes that express privacy requirements.

In a second step, the coded activities (cf. Fig. 1) are to be combined into a process model. We gained positive experience with expressing medical guidelines with BPMN, the standard process modeling language [Du11]. Thus, in the following, process models are derived from the code book activities in BPMN.

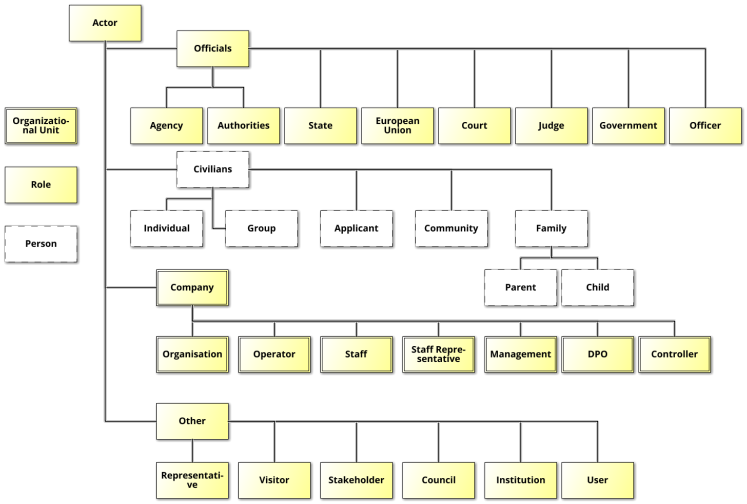


Figure 2: Transformation into Organigram (Using Signavio)

The identification of which codes belong to the same process model is based on co-occurences and proximity of codes. Both can be analyzed by comparing overlapping code segments. Co-occurence, frequency, and proximity can be measured by different indexes, e.g., the Jaccard's coefficient as for the dendrogram depicted in Fig. 3.

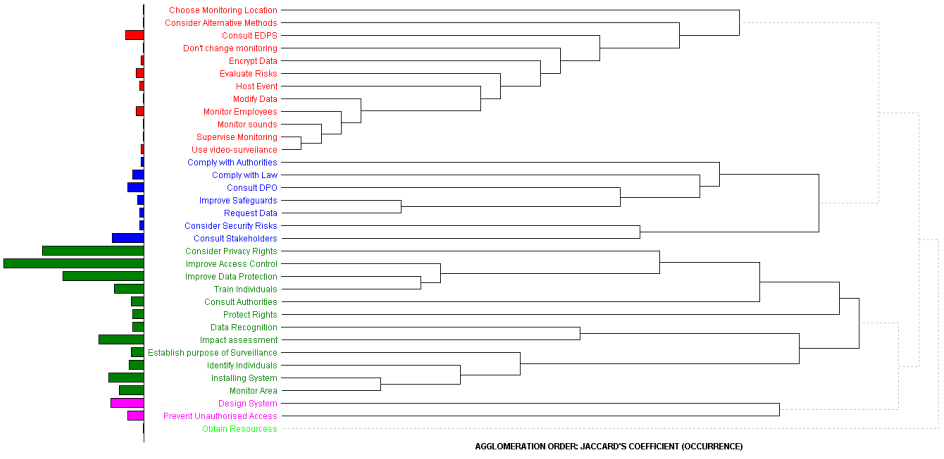


Figure 3: Dendrogram: Co-Occurrence of Codes (Using QDA Miner)

The dendogram is produced with 5 clusters by QDA Miner expressed by the color of the bars. One noticeable cluster is the green one where specifically activities `Consider Privacy Rights`, `Improve Access Control`, and `Improve Data Protection` show a high similarity (degree of co-occurence). This impression is supported by the proximity plot in Fig. 4 for activity `Consider Privacy Rights` which

shows the a proximity of 1.0 with activities `Improve Access Control`, `Improve Data Protection`, and `Train Individual`.

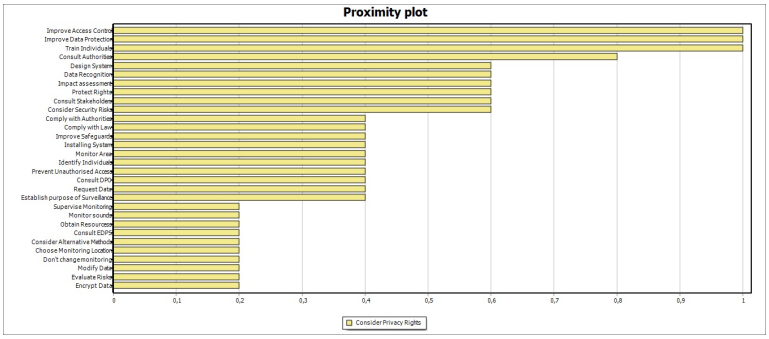


Figure 4: Proximity Plot for Activity `Improve Access Control` (Using QDA Miner)

It is also possible to analyze Code Sequences in QDA Miner, for example, the frequencies and probability of an activity A followed by another activity B. This analysis yields, for example, that activity `Consider Privacy Rights` is followed by `Improve Data Protection` in 12.5% of the cases.

The above analysis results provide an overview of the relations between coded activities. It is difficult to directly derive process models from these analysis as codes may occur multiple times and the context of each occurrence must be taken into consideration before creating a model. Hence, the analysis results can be taken as hints for candidates when revisiting the coded text again. Selecting code `Consider Privacy Rights` and comparing the coded text fragments with the analysis results, the fragment depicted in Fig. 5 is considered a candidate for a process model reflecting a privacy requirement.

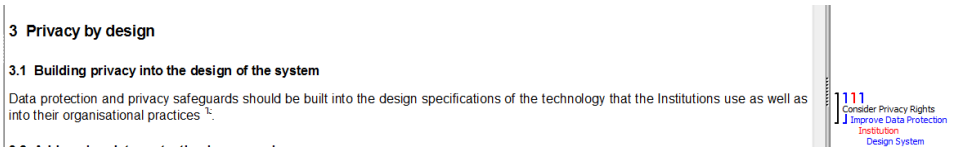


Figure 5: Text Fragment and Codes: `Institution, Consider Privacy Rights, Improve Data Protection, Design System`

More precisely, the fragment contains the codes `Institution, Consider Privacy Rights, Improve Data Protection, Design System` whereof the three activities `Consider Privacy Rights, Improve Data Protection, Design System` are related under co-occurrence (cf. Fig. 3), proximity (cf. Fig. 4), and (partly) code sequence probability. The latter shows that `Improve Data Protection` has some probability to follow `Consider Privacy Rights`. The Frequency Matrix shows that `Design System` seems to be not in a sequence with any other activity. Thus, it can be concluded that `Design System` occurs together with `Consider Privacy Rights` and `Improve Data Protection`, but in no specific order, whereas `Consider Privacy Rights` occurs in sequence with `Improve Data Protection`. The process model in Fig. 6 describes these orders, particularly, the parallel ordering of `Design System` with the other activities.

In the text, activities `Consider Privacy Rights`, `Improve Data Protection`, `Design System` are connected with actor `Institution`. Proximity analysis shows that `Design System` has a proximity of 0.67 and `Consider Privacy Rights` has a proximity of 0.4 (Jaccard coefficient). This assignment is reflected by positioning these two activities in the lane `Institution`. The lane where `Improve Data Protection` is positioned has been marked with ? as the assigned actor must be further investigated. Proximity analysis shows potential candidates such as `Individual`, `Group`, `Officer`, `State`, and `Staff` with a proximity of 1.0. These candidates must be again checked against the text fragments and codes. Due to space restrictions we abstain from details here. However, all lanes can be positioned in pool `Actor` according to the organigram in Fig. 2.

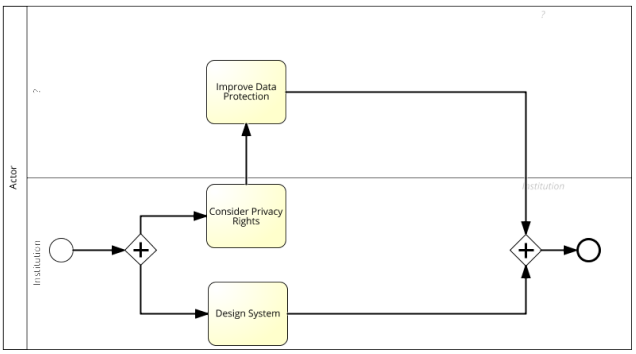
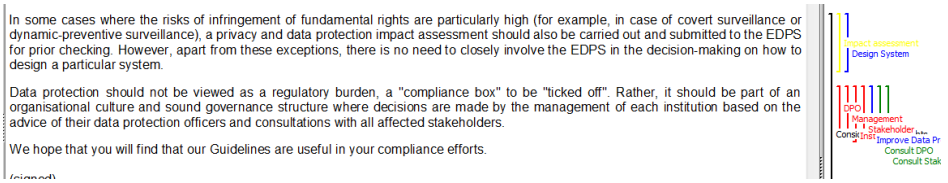


Figure 6: Guideline Example Derived from Text Fragment and Codes in Fig. 5 (Modeled in BPMN Using Signavio)

The coded text fragment depicted in Fig. 5 is relatively simple. An interesting question is how to deal with more complex text fragments and codes as shown, for example, in Fig. 7.


In some cases where the risks of infringement of fundamental rights are particularly high (for example, in case of covert surveillance or dynamic-preventive surveillance), a privacy and data protection impact assessment should also be carried out and submitted to the EDPS for prior checking. However, apart from these exceptions, there is no need to closely involve the EDPS in the decision-making on how to design a particular system.

Data protection should not be viewed as a regulatory burden, a "compliance box" to be "ticked off". Rather, it should be part of an organisational culture and sound governance structure where decisions are made by the management of each institution based on the advice of their data protection officers and consultations with all affected stakeholders.

We hope that you will find that our Guidelines are useful in your compliance efforts.

Design System
DPO
Management
Stakeholder
Improve Data Pr
Consult DPO
Consult Stak

Figure 7: Example Text Fragment and Codes for `Consider Privacy Rights`

4 Conclusion and Discussion

The first case study shows that a QCA is in principle an interesting knowledge engineering approach to derive business process models reflecting privacy requirements from text such as regulatory documents. It also shows that a structured methodology is necessary. As a first proposal, we suggest:

PE-QCA Methodology - first draft

1. Code the documents; categories `Actor` and `Activity`
2. Derive organigram from `Actor` hierarchy
3. Apply co-occurrence, proximity, and code sequence analysis to activities and select candidates for process task elements
4. Go back to text and codes, select phrases and codes for candidates
5. Apply sequence analysis to selected activities and derive process model
6. Select attached actors and check proximity for each activity candidate
7. Add pool and lanes respectively
8. **Discuss with experts**

The last item is crucial to validate the feasibility of the process models and is present in most of the existing methodologies. Moreover, most likely the PE-CQA methodology has to be applied iteratively. Probably, for each candidate set of activities all associated text fragments should be considered. We see process as a glue to connect human and technology as well as a vehicle to preserve and enhance privacy in various information systems. Process models can be used to facilitate many aspects of privacy engineering. Especially, they can be used to capture and present the privacy requirements and define privacy-preserving process in system design and operation. As a targeted format of knowledge engineering of privacy requirements, once created, process models can be shared, extended, and verified by domain experts (e.g. law professional, ethical experts, and system engineers) based on reusable models and reproducible procedure and techniques. As next steps, the methodology will be applied to further case studies from the privacy domain. Moreover, the case studies will be repeated with other knowledge engineering techniques such as text mining. The results of the different case studies and of the application of the different techniques will be taken as evaluation of the method proposed above. We think that the most promising way will be a combination of different techniques as all of them have specific advantages.

Another interesting question is how the findings can be transferred to other areas such as health care. Here the extraction and modeling of medical guidelines plays an important role as well [Du11]. The same holds for compliance requirements in general [Ly15]. In order to provide a comprehensive analysis of the transferability of the proposed methodology in the context of privacy requirements, at first, the literature review must be extended to cover the area of compliance requirement engineering and approaches from other domains such as medical guidelines. For the application of content analysis, the methodology seems to be quite generic and not confined to privacy requirements. However, this statement, must be underpinned with respective case studies which will be part of our future work. Finally, it would be beneficiary to derive entire process models from textual description as process elicitation and modeling can be a tedious and costly job [KRM11]. Friedrich et al. [Fr11] provide an approach based on NLP for the derivation of process models (in BPMN) from text. It will be part of future work to apply a comprehensive analysis

and comparison of existing approaches for establishing a methodology for privacy requirement elicitation.

Acknowledgments

This work was partly funded by the EC through the project PrivAcy pReserving Infrastructure for Surveillance (PARIS) (FP7-SEC-2012-1-312504).

References

[AE00] Antón, A.; Earp, J.: Strategies for developing policies and requirements for secure electronic commerce systems. In: E-commerce security and privacy (2):29–46, 2000.

[AM14] Ahmed, N.; Matulevicius, R.: A Method for Eliciting Security Requirements from the Business Process Models. In: CAISE Forum, 2014.

[ANM10] Abu-Nimeh, S.; Mead, N.: Combining Privacy and Security Risk Assessment in Security Quality Requirements Engineering. In: AAAI Spring Symposium: Intelligent Information Privacy Management, 2010.

[AZ12] Aggarwal, C.; Zhai, C.: Mining text data. Springer Science & Business Media, 2012.

[BA08] Breaux, T; Antón, A.: Analyzing regulatory rules for privacy and security requirements. IEEE TSE 34(1):5–20, 2008.

[Bi08] Birnhack, M.: The EU data protection directive: an engine of a global regime. Computer Law & Security Review 24(6):508–520, 2008.

[BM10] Bijwe, A.; Mead, N.: Adapting the square process for privacy requirements engineering. Techn. Rep. CMU/SEI-2010-TN-022, Carnegie-Mellon, 2010.

[BVA06] Breaux, T.D.; Vail, M.W.; Anton, A.I.: Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations. In: Requirements Engineering, pp. 49–58, 2006.

[Ch08] Chiasera, A.; Casati, F.; Daniel, F.; Velegrakis, Y.: Engineering privacy requirements in business intelligence applications. In: Secure Data Management, pp. 219–228. Springer, 2008.

[Ch11] Chikh, A.; Abulaish, M.; Nabi, S.; Alghathbar, K.: An ontology based information security requirements engineering framework. In: Secure and Trust Computing, Data Management and Applications, pp. 139–146. Springer, 2011.

[Co07] Compagna, L.; Khoury, P.; Massacci, F.; Thomas, R.; Zannone, N.: How to capture, model, and verify the knowledge of legal, security, and privacy experts: a pattern-based approach. In: P.Artificial intelligence and law. pp. 149–153, 2007.

[De11] Deng, M.; Wuyts, K.; Scandariato, R.; Preneel, B.; Joosen, W.: A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requirements Engineering, 16(1):3–32, 2011.

[dRAF05] da Rocha, S.; Abdelouahab, Z.; Freire, E.: Requirement Elicitation Based on Goals with Security and Privacy Policies in Electronic Commerce. In: WER. pp. 63–74, 2005.

[Du11] Dunkl, R.; Fröschl, K.; Grossmann, W.; Rinderle-Ma, S.: Assessing Medical Treatment Compliance Based on Formal Process Modeling. In: USAB 2011. Springer, pp. 533–546, 2011.

[El11] Elahi, G.; Yu, E.; Li, T.; Liu, L.: Security requirements engineering in the wild: A survey of common practices. In: IEEE Computer Software and Applications Conference. pp. 314–319, 2011.

[Fa10] Fabian, B.; Gürses, S.; Heisel, M.; Santen, T.; Schmidt, H.: A comparison of security requirements engineering methods. Requirements engineering, 15(1):7–40, 2010.

[Fr11] Friedrich, F.; Mendling, J.; Puhlmann, F.: Process Model Generation from Natural Language Text. In: CAiSE, pp. 482-496, 2011

[Gr12] Graa, M.; Cuppens-Boulahia, N.; Autrel, F.; Azkia, H.; Cuppens, F.; Coatrieux, G.; Cavalli, A.; Mammar, A.: Using requirements engineering in an automatic security policy derivation process. In: Data Privacy Management and Autonomous Spontaneus Security, pp. 155–172. Springer, 2012.

[Gü05] Gürses, S.; Jahnke, J.; Obry, C.; Onabajo, A.; Santen, T.; Price, M.: Eliciting confidentiality requirements in practice. In: Conf. of the Centre for Advanced Studies on Collaborative research. pp. 101–116, 2005.

[He03] He, Q.; Antón, A. et al.: A framework for modeling privacy requirements in role engineering. In: Proc. of REFSQ 3, pp. 137–146, 2003.

[In97] Introna, L.: Privacy and the computer: why we need privacy in the information society. Metaphilosophy, 28(3):259–275, 1997.

[KBG11] Kalloniatis, C.; Belsis, P.; Gritzalis, S.: A soft computing approach for privacy requirements engineering: The PriS framework. Applied Soft Computing, 11(7):4341–4348, 2011.

[Ki09] Kitchenham, B.; Pearl Brereton, O; Budgen, D.; Turner, M..; Bailey, J.; Linkman, S.: Systematic literature reviews in software engineering–a systematic literature review. Information and software technology, 51(1):7–15, 2009.

[KRM11] Kabicher, S.; Rinderle-Ma, S.: Human-centered process engineering based on content analysis and process view aggregation. In: Advanced Information Systems Engineering. pp. 467–481, 2011.

[KS85] Kowalski, R.; Sergot, M.: Computer Representation of the Law. In: IJCAI. pp. 1269–1270, 1985.

[Le06] Lee, S.; Gandhi, R.; Muthurajan, D.; Yavagal, D.; Ahn, G.: Building problem domain ontology from security requirements in regulatory documents. In: Workshop on Software engineering for secure systems. pp. 43–50, 2006.

[Le14] Leitner, M.; Rinderle-Ma, S.: A systematic review on security in Process-Aware Information Systems - Constitution, challenges, and future directions. In: Information & Software Technology 56(3): 273-293, 2014

[Ly15] Ly, L.T.; Maggi, F.M.; Montali, M.; Rinderle-Ma, S.; van der Aalst, W.M.P.: Compliance monitoring in business processes: Functionalities, application, and tool-support. Information Systems, 2015. (in press).

[LYM03] Liu, L.; Yu, E.; Mylopoulos, J.: Security and privacy requirements analysis within a social setting. In: IEEE Requirements Engineering Conference. pp. 151–161, 2003.

[Ma14] Ma, Z. et al: Towards a Multidisciplinary Framework to Include Privacy in the Design of Video Surveillance Systems. In: 2nd Annual Privacy Forum - Privacy Technologies and Policy. pp. 101–116, 2014

[MdAY14] Martin, Y.; del Alamo, J.; Yelmo, J.: Engineering privacy requirements valuable lessons from another realm. In: Evolving Security and Privacy Requirements Engineering. pp. 19–24, 2014.

[Me10] Mellado, D.; Blanco, C.; Sánchez, L.; Fernández-Medina, E.: A systematic review of security requirements engineering. Computer Standards & Interfaces, 32(4):153–165, 2010.

[MMZ08] Miyazaki, S.; Mead, N.; Zhan, J.: Computer-aided privacy requirements elicitation technique. In: IEEE Asia-Pacific Services Computing Conf. pp. 367–372, 2008.

[MMZ11] Mead, N.; Miyazaki, S.; Zhan, J.: Integrating privacy requirements considerations into a security requirements engineering method and tool. Int'l Journal of Information Privacy, Security and Integrity, 1(1):106–126, 2011.

[MPZ05] Massacci, F.; Prest, M.; Zannone, N.: Using a security requirements engineering methodology in practice: the compliance with the Italian data protection legislation. Computer Standards & Interfaces, 27(5):445–455, 2005.

[PDG14] Paja, E.; Dalpiaz, F.; Giorgini, P.: STS-Tool: Security Requirements Engineering for Socio-Technical Systems. In: Engineering Secure Future Internet Services and Systems, pp. 65–96. Springer, 2014.

[RGK13] Radics, P.; Gracanin, D.; Kafura, D.: Preprocess before You Build: Introducing a Framework for Privacy Requirements Engineering. In: Social Computing (SocialCom). IEEE, pp. 564–569, 2013.

[Ri14] Riaz, M.; King, J.; Slankas, J.; Williams, L.: Hidden in plain sight: Automatically identifying security requirements from natural language artifacts. In: Requirements Engineering Conference. pp. 183–192, 2014.

[Sa08] Sarawagi, S.: Information extraction. Foundations and trends in databases, 1(3):261–377, 2008.

[SBF98] Studer, R.; Benjamins, V.; Fensel, D.: Knowledge engineering: principles and methods. Data & knowledge engineering, 25(1):161–197, 1998.

[SK12] Salini, P; Kanmani, S: Survey and analysis on security requirements engineering. Computers & Electrical Engineering, 38(6):1785–1797, 2012.

[St06] Strijbos, J.; Martens, R.; Prins, F.; Jochems, W.: Content analysis: What are they talking about? Computers & Education, 46(1):29–48, 2006.