Privacy Pattern Catalogue: A Tool for Integrating Privacy Principles of ISO/IEC 29100 into the Software Development Process

Olha Drozd

Privacy Pattern Catalogue: A Tool for Integrating Privacy Principles of ISO/IEC 29100 into the Software Development Process

Olha Drozd

Vienna University of Economics and Business, Vienna, Austria
olha.drozd@wu.ac.at

Abstract. A proper integration of privacy patterns into a software development process enables development of reliable and privacy-friendly software products. While previous work has identified some loosely connected privacy patterns, there exists no comprehensive privacy pattern catalogue that is specifically designed for the application by software architects during the software development process. To address this gap an interactive online privacy pattern catalogue was developed using patterns obtained from interviews with privacy experts as well as from existing privacy patterns work. The catalogue classifies patterns according to the description of the privacy principles of the international standard ISO/IEC 29100:2011 (E) and is, therefore, internationally applicable.

Keywords. Privacy patterns · Privacy principles · ISO/IEC 29100:2011 (E) · Privacy by design · Privacy pattern catalogue

1 Introduction

Any software that processes personally identifiable information (i.e. any information that can be used to identify the natural person to whom such information relates, or is or might be directly or indirectly linked to that person [1]) should protect the privacy of data subjects. In order to develop qualitative privacy-friendly software, privacy should be integrated into software at early stages of the software development process. At the design stage of the software development cycle one can utilize the privacy patterns to facilitate privacy integration [2]. "A pattern is a piece of literature that describes a design problem and a general solution for the problem in a particular context [3]."

Some attempts were made to collect privacy patterns, but those collections were limited to a very specific context. For example, Hafiz concentrates on patterns for the design of anonymity systems [4]. The Privacy and Identity Management in Europe for Life pattern collection provides a list of human-computer interaction patterns [5]. The University of California (UC) Berkley School of Information's collection [6] broadens the application context of privacy patterns, but describes only 9 privacy patterns. PRIPARE project [7] presents the newest attempt to collect privacy patterns. Howev-

er, their list of patterns is also limited. They added some new patterns but omitted some old ones that were mentioned in the previous work.

While reviewing the literature, no pattern classifications, catalogues or tools providing a structured approach to privacy pattern integration into the software development process in connection with the ISO (the International Organization for Standardization) /IEC (the International Electrotechnical Commission) 29100 standard were found. To fill this gap an interactive online privacy pattern catalogue was developed. This structured solution-oriented representation of patterns that collects numerous patterns in one place and allows end user to easily navigate through the catalogue, could convince companies to adopt privacy by design approach.

The target audience of the catalogue is software architects and software developers. The catalogue enables them to efficiently integrate the privacy principles of ISO/IEC 29100 into the software development process. The ISO/IEC 29100 privacy principles were selected because they comprehensively cover the domain of privacy requirements and because the standard is internationally applicable.

The paper is divided into 5 parts. The background part provides the background information on the design patterns and the ISO/IEC 29100 standard. The method section describes the structured-case method employed in this research project. The part following the methodology section explains the idea of the privacy pattern catalogue and describes its functionality. The discussion part provides an overview of the issues that appeared during the compilation of the catalogue and describes how those issues were addressed in the catalogue as well as how they might be solved in the future work. Finally, the conclusion summarizes the main points of the paper.

2 Background

Privacy by design aims to integrate privacy requirements into every stage of the software development process [2]. This paper addresses the problem of privacy integration at the design stage of the system development process by using design patterns as one of the main components of the catalogue. The descriptions of 11 privacy principles from the ISO/IEC 29100 standard were selected as the source of the privacy requirements.

2.1 Design Patterns

As it was mentioned before, privacy patterns may help to integrate privacy at the design stage of the software development process [2]. There exist many definitions of patterns. For instance, in building and architecture pattern describes an iterative problem in a specific environment and a reusable solution to it [8]. In software engineering "patterns codify reusable design expertise that provides time-proven solutions to commonly occurring software problems that arise in particular contexts and domains [9]." Another definition of the term pattern in software engineering field describes pattern as "a description of communicating objects and classes that are customized to solve a general design problem in a particular context [10]."

All the above-mentioned definitions suggest describing patterns in terms of problem description, solution to the problem and context where this problem occurs. The existent lists of privacy patterns describe them in different ways. Hafiz provides a very detailed description of patterns and uses the following sections: intent, also know as, motivation, context, problem, forces, solution, design issues, consequences, known uses, related patterns [4]. PRIPARE project uses a similar approach in its description of privacy patterns. Every pattern here is described with the help of summary, problem, context, goals, motivating example, solution, constraints and consequences, known uses, tags, categories and technology readiness level. Patterns collected at the UC Berkeley School of Information use a less detailed description template. The sections intent, context, problem, solution and examples are used in that project. The template of the PrimeLife project contains the following sections: problem, solution, use when, how, why, related patterns [5]. The sections problem, solution and context seem to be a universal way to describe patterns because they are always present, in various formulations, in the definitions as well as in the pattern lists. That is why for the purpose of this research the patterns were explained with the help of those three sections. The consequences section was added to the description to provide a better understanding of the results after implementation of the pattern. This section proved to be useful in the above-described pattern lists as well.

For example, the pattern Data Track from the PrimeLife project is described in the catalogue as follows: [5]

Table 1. Description of the Data Track pattern.

Section	*Description*
Problem	Users may lose an overview of what kind of data they disclosed to whom under which conditions
Solution	Provide an end-user transparency tool that provides the user with a detailed overview of all the user's personal data releases to communication partners
Context	Implement when personal data are released
Consequences	Easier recollection of where, when and under what conditions the user posted which data

2.2 ISO/IEC 29100 as the Source of the Privacy Requirements

The aim of the catalogue is to facilitate the integration of privacy requirements at the design stage of the software development process. There are a number of data protection laws that could be used as a requirement sources but they are often specific to every country and are, unfortunately, slightly outdated. For example, "Privacy Online: Fair Information Practices in the Electronic Marketplace" [11] is specific to the United States, "Organization of Economic Co-Operation and Development Guidelines on the Protection of Privacy and Transborder Flows of Personal Data" [12] is somewhat outdated, General Data Protection Regulation [13] covers only the European Union and still needs to be finalized.

The ISO/IEC 29100 standard was chosen as the source of the privacy requirements, as it is an international standard that "provides a high-level framework for the protection of personally identifiable information (PII) within information and communication technology (ICT) systems [1]." Moreover, the ISO/IEC 29100 standard was compiled to help, inter alia, architect, design and develop ICT systems or services in a privacy-friendly way [1]. In addition to the advantageous characteristics mentioned above, being international, this standard could be equally applied in different countries. It also comprehensively covers the domain of privacy requirements.

In the standard the requirements are presented in the form of 11 privacy principles: consent and choice; purpose legitimacy and specification; collection limitation; data minimization; use, retention and disclosure limitation; accuracy and quality; openness, transparency and notice; individual participation and access; accountability; information security; privacy compliance [1]. Each principle is then described in more detail with the help of the list of bullet points. Every bullet point explains, in the form of an instruction, what adhering to this or that principle means. For instance, to adhere to the collection limitation principle one should limit "the collection of PII to that which is within the bounds of applicable law and strictly necessary for the specified purpose(s) [1]." Other principles are explained in a similar way. The number of bullet points varies depending on the privacy principle.

The privacy principles of ISO/IEC 29100 and the instructions form the first and the second level of the catalogue hierarchy respectively. The third level of the catalogue is filled with the privacy patterns that help (directly or indirectly) to implement the corresponding privacy principle instruction (Figure 1).

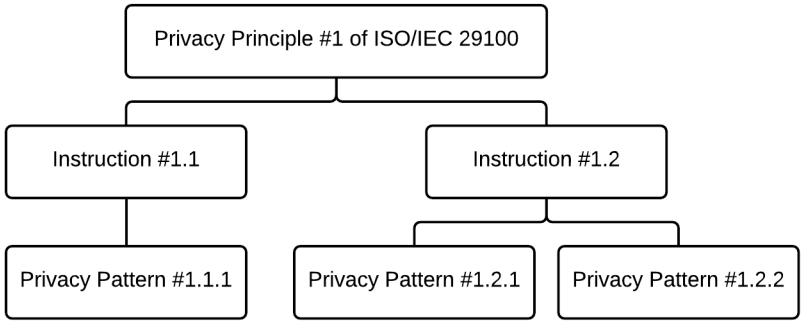

Fig. 1. Catalogue concept

3 Method

The catalogue was compiled by applying a structured-case methodological framework for building theory in information systems research [14]. The structured-case consists of 3 structural components, namely "the conceptual framework, the research cycle and the literature-based scrutiny of the theory built [14]." The general idea of the structured-case research method is shown in Figure 2.

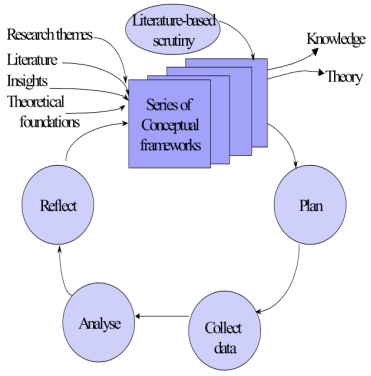

Fig. 2. The structured-case research method (Source: Carroll, J.M., Swatman, P. A.: Structured-case: a methodological framework for building theory in information systems research. Eur. J. Inf. Syst. 9, 235–242 (2000))

3.1 Conceptual Framework

According to Miles and Huberman, "a conceptual framework explains either graphically or in narrative form, the main things to be studied – the key factors, constructs or variables – and the presumed relationships among them [15]."

Figure 1 depicts the concept of the privacy pattern catalogue – the elements of the catalogue and the hierarchical relationship between them. The first list of privacy patterns for the conceptual framework was compiled from the reviewed privacy pattern literature. As it is allowed, or even welcomed, in structured-case methodology to update the framework if valuable knowledge is gained during the research process [14], new patterns were added and the description of some patterns was refined in the course of the project.

3.2 Research Cycle

Each research cycle was divided into four stages: plan, collect, analyse and reflect [14].

Plan. 11 interviews with privacy experts (PhD candidates, PhDs, professors and professionals in the field of data protection) from Austria, Germany, Greece, Ireland, Sweden and the USA were planned. The three main goals of those interviews were to classify the patterns according to the privacy principle instructions of ISO/IEC 29100, to expand the set of privacy patterns derived from the literature review and, if necessary, to update the description of patterns. To make the process of interviewing easier, the online questionnaire was developed. The questionnaire consisted of 55 privacy principle instructions and 28 privacy patterns with the descriptions. The patterns and their descriptions were derived from the literature review.

Each instruction of the privacy principle formed a separate question. For example, the consent and choice principle is described with the help of a bulleted list of 5 instructions [1]. Therefore, there were 5 questions concerning the consent and choice principle in the questionnaire. The privacy patterns were described in terms of what context they can be used in, what problem they solve, what solution they offer and what consequences should be expected after the implementation of the patterns. That is, in the same way as it was planned to describe the patterns in the catalogue. The description was shown upon mouseover on the 'i' icon. Figure 3 illustrates the cropped version of the first page of the questionnaire.

It was decided that the results of this research project would be presented in the form of the interactive online privacy pattern catalogue.

Collect. 11 interviews were conducted at the collection stage of the research cycle. The privacy experts were asked to choose patterns that, in their opinion, could implement the instructions of the privacy principles. They also chose what connection (direct or indirect) the pattern had to the privacy principle instruction. The interviewees were asked to explain their decisions briefly. The questionnaire also presented a possibility to add privacy patterns if the experts suggested patterns that were missing from the list. The interviewees also commented on the descriptions of patterns and on the questionnaire in general. The answers from the questionnaire were saved into the database and the experts' comments were recorded.

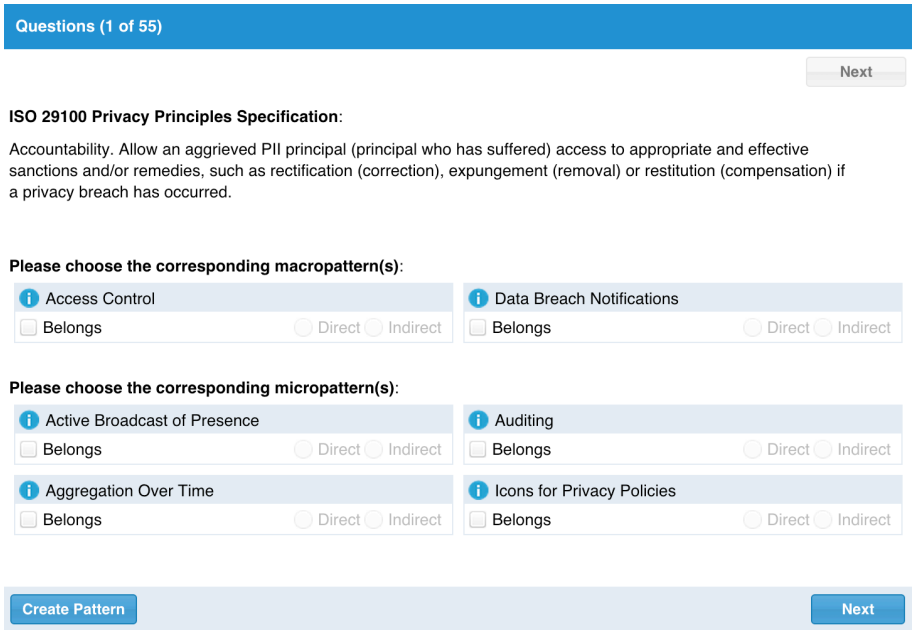

Fig. 3. The first page of the questionnaire (cropped version)

Analyse. The data from the database were organized into 55 bar charts. One of the bar graphs is shown in Figure 4. The important comments and improvement suggestions from the recordings were transcribed and, if confirmed during the reflection stage, implemented for the next research cycle.

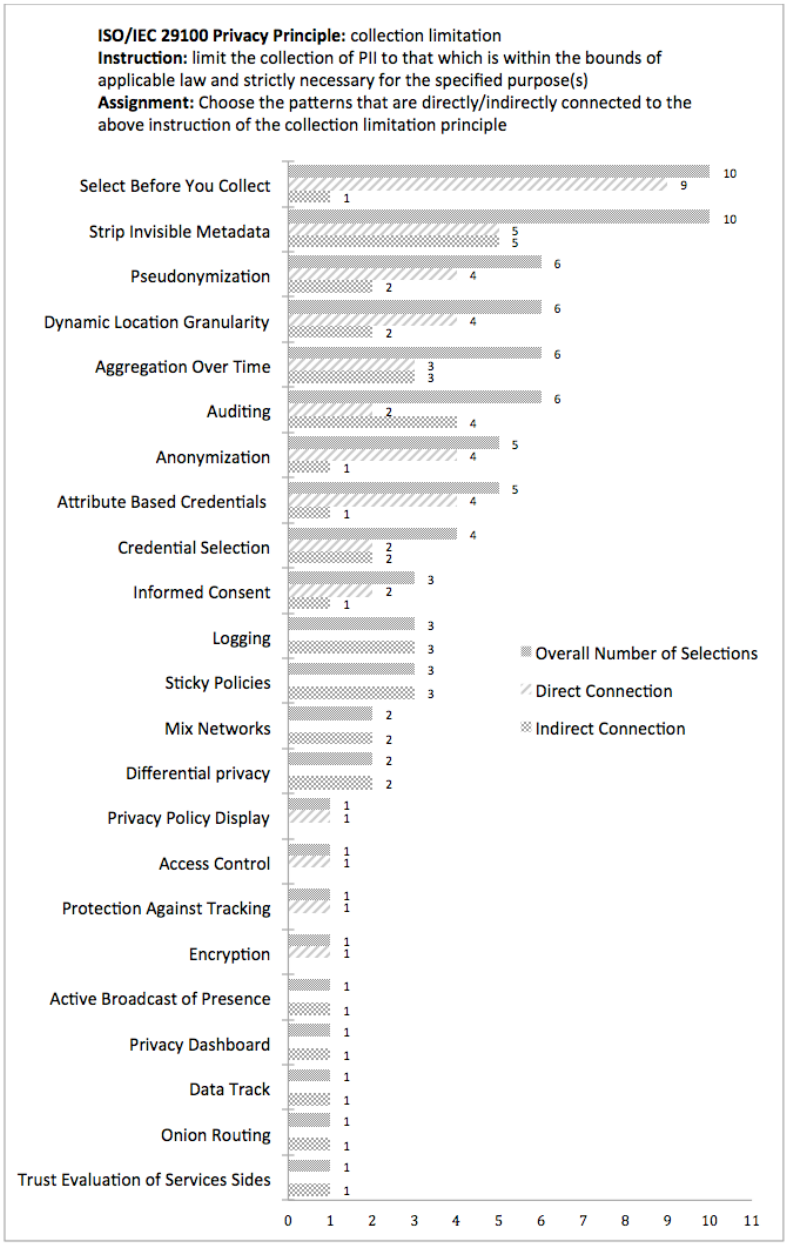

Fig. 4. The results for the 'collection limitation' principle

Reflect. The fellow researchers evaluated the results of the analysis and reflected on the interviewees' comments regarding the research process and the conceptual framework. The conceptual framework was updated with the new knowledge acquired at previous stages of the research cycle.

3.3 Theory Building

The research cycles were stopped at 11 interviews because the amount of new data, improvement suggestions and ideas received from the experts were low in the last interviews.

The last component of the structured-case methodology requires the results to be compared to the existent literature.

The findings were compared with the scarce literature on privacy patterns as well as with the technology descriptions suggested during interviews. To the best of my knowledge there were no attempts made to classify privacy patterns according to the privacy principles of ISO/IEC 29100, so no comparison was performed on the findings concerning this matter.

4 Privacy Pattern Catalogue

The interactive online privacy pattern catalogue (Figure 5) [16] has been developed to present the results in a comprehensible and usable way. This online tool presents 40 privacy patterns in a structured way by grouping them according to the privacy principle instructions.

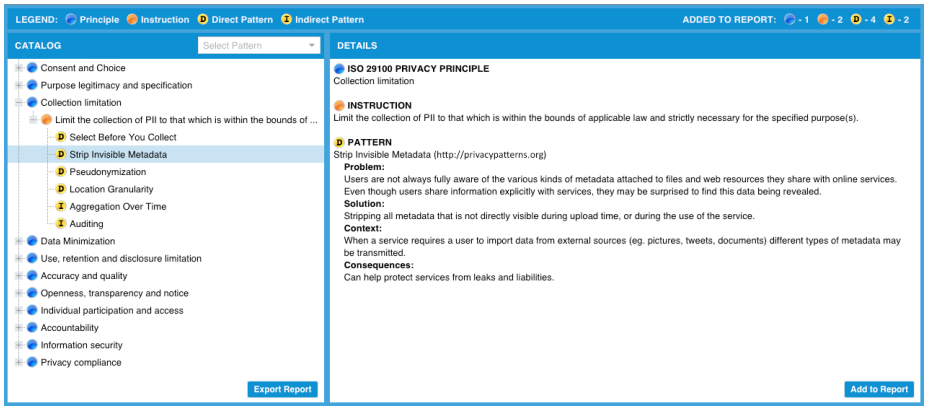

Fig. 5. The interactive privacy pattern catalogue

The catalogue could be useful for software architects and software developers in the projects where ISO/IEC 29100 certification is required. Depending on the status of the project, software architects can use the catalogue in both top-down and bottom-up directions. A top-down approach is used to identify which patterns implement a

specific ISO/IEC 29100 privacy principle or instruction. A bottom-up approach provides the information on the ISO/IEC 29100 privacy principle and the corresponding instruction implemented by the chosen privacy pattern. Additionally, one can utilize the catalogue for the training purposes.

Two extra functionalities were integrated into the catalogue:

- Search by privacy pattern
- Export the report

The first feature gives a possibility to view what instructions and privacy principles are (to some extent) covered by the chosen privacy pattern.

The second feature could be very useful for the top-down approach. Because of the large amount of possible combinations, it could be difficult and time-consuming for a software architect to document the principles, instructions and patterns that are relevant to the project. To address this issue the system allows software architects to select required elements, stores all chosen items in a database through the whole selection process and offers a possibility to generate a report (Figure 6) that contains all selected elements.

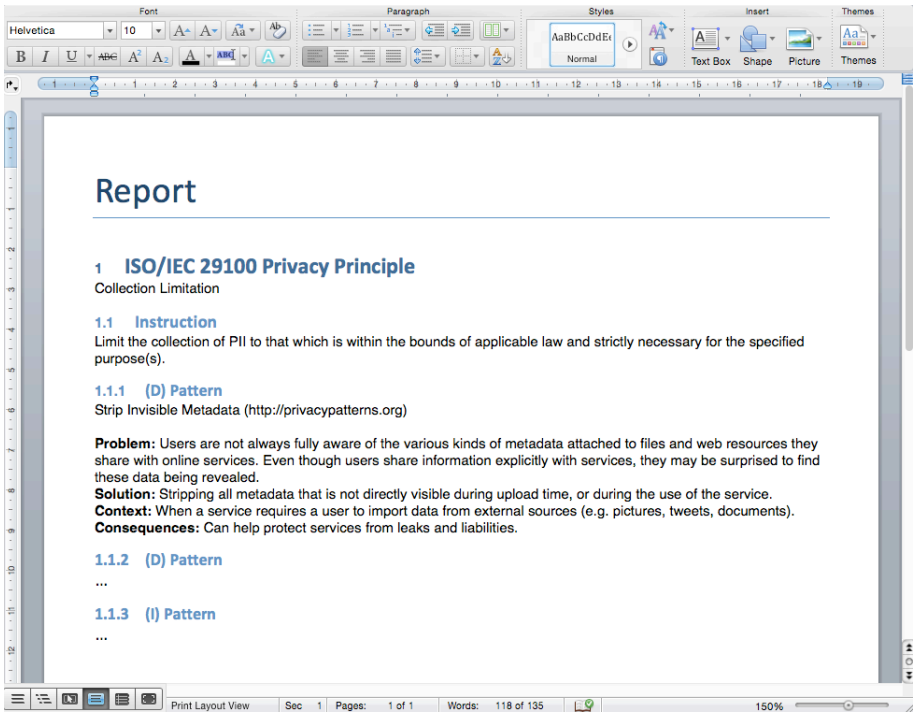

Fig. 6. Generated report (shortened version)

5 Discussion

While compiling the above-described catalogue a number of issues occurred.

Although the interviews are considered to be one of the powerful methods for gaining knowledge, there are some problems connected with this way of gathering information. In the case of this project the interviews lasted up to 4.5 hours and the interviewees mentioned the problem of time pressure and the need to complete the questionnaire as quickly as possible. Indeed, the lack of time can cause two problems: either the information gathered will be incomplete or the interviewees will generate more input than they usually would do in the normal situation but the obtained information could be unreliable [17]. In order to examine how this issue might have influenced the results, the existent catalogue could be compared to the results obtained from the answers to the questionnaire that was filled out without time pressure.

The conceptual framework illustrates a clean tree structure of the catalogue. However, the privacy pattern instructions of the ISO/IEC 29100 overlap in some cases. This makes it possible that the same pattern could be assigned to different instructions. To mitigate this issue and to give a better overview of what instructions could be covered by one and the same pattern, the catalogue offers the 'search by privacy pattern' functionality. By using this functionality the user can obtain a summary of all the privacy principle instructions and corresponding privacy principles that are (to a certain extent) implemented by the chosen pattern.

Another issue mentioned by some interviewees was that sometimes the name and the description of the pattern were formulated in a very broad or very narrow manner. This may explain why abstract patterns appeared more often in the catalogue compared to the concrete ones. To partially solve this issue the comments from the interviews will be used to extend the catalogue by categorizing patterns into different dimensions and adding various angles of view in terms of context. This should bring even more structure to the pattern collection.

Additionally, the patterns could be described in more detail using more sections in the description template. One area that is currently under investigation is the compilation of a standardized template for privacy patterns. After the template is finalized the patterns in the catalogue should be updated according to it.

Another issue arises due to the fact that there are some privacy principle instructions, which do not have corresponding direct privacy patterns. This research project showed that "technical" patterns cover only a part of the ISO/IEC 29100 requirements. Some requirements could only be implemented by the organizational measures and processes. This means that the privacy patterns in the field of information technology governance should be identified and described.

The catalogue could also be extended by adding another hierarchy level (Figure 7) to cater for privacy enhancing technologies (PETs) that would be assigned to the corresponding patterns.

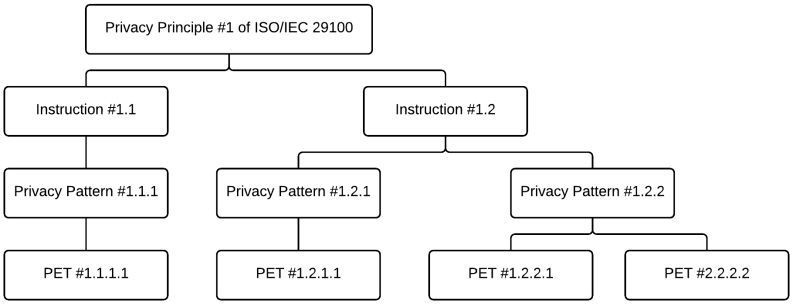

Fig. 7. Concept of an extended catalogue

6 Conclusion

This work presents an interactive online privacy pattern catalogue for software architects and software developers drawing both on a review of existing privacy patterns and on the interviews with privacy experts.

In the catalogue, privacy patterns are matched with the privacy principle instructions of the ISO/IEC 29100 standard, i.e. the users can view the list of patterns that implement a particular privacy principle instruction and read detailed information about those patterns. The users can also search by a particular pattern to see what privacy principles and their instructions could be implemented by that pattern. The catalogue provides a possibility to select patterns that are relevant to the project, from the software architect's point of view, and then automatically generate a report that presents all the selected items in a structured manner.

The process of privacy pattern classification showed that "technical" privacy patterns cover only a part of the privacy principle instructions of the ISO/IEC 29100 standard. As a result, the catalogue contains privacy principle requirements without corresponding direct privacy patterns. Those principle instructions could be implemented with the help of organizational processes.

References

1. International Standard ISO/IEC 29100:2011(E) Information technology — Security techniques — Privacy framework (2011).
2. Hoepman, J.: Privacy design strategies. arXiv Prepr. arXiv1210.6621. 12 (2012).
3. Coplien, J.O.: Software Patterns. Lucent Technologies, Bell Labs Innovations, New York (1996).
4. Hafiz, M.: A collection of privacy design patterns. Proc. 2006 Conf. Pattern Lang. programs - PLoP '06. 1 (2006).
5. Fischer-Hübner, S., Köffel, C., Pettersson, J.-S., Wolkerstorfer, P., Graf, C., Holtz, L.E., König, U., Hedbom, H., Kellermann, B.: HCI Pattern Collection – Version 2. Priv. Identity Manag. Eur. Life. 61 (2010).
6. Privacy Patterns, privacypatterns.org, access date: 14.07.2015.

7. Privacypatterns.eu - Collecting Patterns for Better Privacy, https://privacypatterns.eu/, access date: 14.07.2015.
8. Alexander, C., Ishikawa, S., Silverstein, M.: A Pattern Language: Towns, Buildings, Construction (1977).
9. Schmidt, D.C., Buschmann, F.: Patterns, frameworks, and middleware: their synergistic relationships. 25th Int. Conf. Softw. Eng. 2003. Proceedings (2003).
10. Gamma, E., Helm, R., Johnson, R.E., Vlissides, J.: Design patterns: elements of reusable object-oriented software. Design. 206, 395 (1995).
11. Anthony, S.F., Thompson, M.W., Swindle, O., Leary, T.B.: Privacy Online : Fair Information Practices in the Electronic Marketplace a Report To Congress. Security (2000).
12. Organisation of Economic Co-Operation and Development: OECD guidelines governing the protection of privacy and transborder flows of personal data (1980).
13. Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (2012).
14. Carroll, J.M., Swatman, P. a: Structured-case: a methodological framework for building theory in information systems research. Eur. J. Inf. Syst. 9, 235–242 (2000).
15. Miles, M.B., Huberman, A.M.: Qualitative Data Analysis (1994).
16. Privacy Pattern Catalogue, privacypatterns.wu.ac.at, access date: 01.11.2015.
17. Myers, M.D., Newman, M.: The qualitative interview in IS research: Examining the craft. Inf. Organ. 17, 2–26 (2007).