Integrating privacy requirements considerations into a security requirements engineering method and tool

Nancy R. Mead*
SEI, Carnegie Mellon University,
4500 Fifth Avenue,
Pittsburgh, PA 15213, USA
E-mail: nrm@sei.cmu.edu
*Corresponding author

Seiya Miyazaki
Corporate R&D Division,
Panasonic Corporation,
1006 Kadoma, Kadoma City,
Osaka 571-8501, Japan
E-mail: miyazaki.seiya@jp.panasonic.com

Justin Zhan
Department of Computer Science,
North Carolina A&T State University,
1601 East Market Street,
Greensboro, NC 27410, USA
E-mail: justinzzhan@gmail.com

Abstract: In this paper we examine a method for identifying privacy requirements within the context of a security requirements engineering method. We briefly describe the security quality requirements engineering (SQUARE) methodology. Next we discuss our definition of privacy and the associated privacy concerns. We discuss the challenges of privacy requirements engineering and the need for incorporating privacy considerations into security requirements engineering approaches. We describe a novel modification to the SQUARE method and tool to incorporate privacy considerations, and identify future work that will lead to a more integrated method for security and privacy requirements engineering.

Keywords: software security engineering; privacy requirements; security requirements engineering; requirements engineering method; requirements engineering tool.

1 Introduction

Web personalisation allows users to obtain valuable information tailored to their needs. For instance, parents can discover their child’s current location if the child has a cell phone. However, the cost of this convenience is that users have to reveal private information such as search keywords, access logs, and global positioning system information to service providers. Since users cannot know whether their service providers are violating their privacy policies or employing inadequate security, they run the risk of having their private information become publicly available. Service providers that want users to entrust them with their personal information must demonstrate that they are willing and able to protect the privacy of that information.

A service provider developing a privacy-compliant system faces two challenges. The first is that developing such a system is technically difficult due to the need to prepare for all possible breaches of security. The second challenge is that each region may have numerous privacy laws, and those laws may vary widely from one region to another. Creating a system that complies with many different laws is a complex undertaking.

In this paper we examine a method for identifying privacy requirements within the context of a security requirements engineering method. First, we briefly describe the security quality requirements engineering (SQUARE) methodology, which has been well documented and discussed in depth elsewhere (Mead et al., 2005; Mead and Stehney, 2005; Mead et al., 2008; Caulkins et al., 2007). Next we discuss our definition of privacy
and the associated privacy concerns. We discuss the challenges of privacy requirements engineering and the need for incorporating privacy considerations into security requirements engineering approaches. We describe a novel modification to the SQUARE method and tool to incorporate privacy considerations, and identify future work that will lead to a more integrated method for security and privacy requirements engineering.

2 Security requirements engineering and SQUARE

It is well recognised in the software industry that requirements engineering is critical to the success of any major development project. Security requirements are often identified during the system lifecycle. However, the requirements tend to be general mechanisms such as password protection, firewalls, and virus detection tools. Often the security requirements are developed independently of the rest of the requirements engineering activity and hence are not integrated into the mainstream of the requirements activities. As a result, security requirements that are specific to the system and that provide for protection of essential services and assets are often neglected. The requirements elicitation and analysis that is needed to get a better set of security requirements seldom takes place.

The Software Engineering Institute’s CERT Program at Carnegie Mellon University has developed the SQUARE methodology to help organisations build security into the early stages of the production lifecycle. SQUARE consists of nine steps that generate a final deliverable of categorised and prioritised security requirements. Although SQUARE could likely be generalised to any large-scale design project, it was designed for use with information technology (IT) systems.

SQUARE involves the interaction of a team of requirements engineers and the stakeholders of an IT project. The requirements engineering team can be thought of as external consultants, though often the team is composed of one or more internal developers of the project. When SQUARE is applied, the user should expect to have identified, documented, and inspected relevant security requirements for the system or software that is being developed.

SQUARE begins with the requirements engineering team and project stakeholders agreeing on technical definitions that serve as a baseline for all future communication. Next, business and security goals are outlined. Third, artefacts and documentation are created, which are necessary for a full understanding of the relevant system. A structured risk assessment determines the likelihood and impact of possible threats to the system. Following this work, the requirements engineering team determines the best method for eliciting initial security requirements from stakeholders, which is dependent on several factors, including the stakeholders involved, the expertise of the requirements engineering team, and the size and complexity of the project. Once a method has been established, the participants rely on artefacts and risk assessment results to elicit an initial set of security requirements. Two subsequent stages are spent categorising and prioritising these requirements for management’s use in making trade-off decisions. Finally, an inspection stage is included to ensure the consistency and accuracy of the security requirements that have been generated.

Table 1 (Mead et al., 2005) details the steps involved in the SQUARE process.
Table 1  SQUARE steps

<table>
<thead>
<tr>
<th>Step 1: Agree on definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td><strong>Technique</strong></td>
</tr>
<tr>
<td><strong>Participant</strong></td>
</tr>
<tr>
<td><strong>Output</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 2: Identify assets and security goals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td><strong>Technique</strong></td>
</tr>
<tr>
<td><strong>Participant</strong></td>
</tr>
<tr>
<td><strong>Output</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 3: Develop artefacts to support security requirements definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td><strong>Technique</strong></td>
</tr>
<tr>
<td><strong>Participant</strong></td>
</tr>
<tr>
<td><strong>Output</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 4: Perform risk assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td><strong>Technique</strong></td>
</tr>
<tr>
<td><strong>Participant</strong></td>
</tr>
<tr>
<td><strong>Output</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 5: Select elicitation techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td><strong>Technique</strong></td>
</tr>
<tr>
<td><strong>Participant</strong></td>
</tr>
<tr>
<td><strong>Output</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 6: Elicit security requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td><strong>Technique</strong></td>
</tr>
<tr>
<td><strong>Participant</strong></td>
</tr>
<tr>
<td><strong>Output</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 7: Categorise requirements as to level (system, software, etc.) and whether they are requirements or other kinds of constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td><strong>Technique</strong></td>
</tr>
<tr>
<td><strong>Participant</strong></td>
</tr>
<tr>
<td><strong>Output</strong></td>
</tr>
</tbody>
</table>
Table 1  SQUARE steps (continued)

<table>
<thead>
<tr>
<th>Step 8: Prioritise requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Categorised requirements and risk assessment results</td>
</tr>
<tr>
<td>Technique</td>
<td>Prioritisation methods such as analytical hierarchy process (AHP), triage, win-win</td>
</tr>
<tr>
<td>Participant</td>
<td>Stakeholders facilitated by requirements engineer</td>
</tr>
<tr>
<td>Output</td>
<td>Prioritised requirements</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 9: Inspect requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Prioritised requirements, candidate formal inspection technique</td>
</tr>
<tr>
<td>Technique</td>
<td>Inspection method such as Fagan, peer reviews</td>
</tr>
<tr>
<td>Participant</td>
<td>Inspection team</td>
</tr>
<tr>
<td>Output</td>
<td>Initial selected requirements, documentation of decision making process and rationale</td>
</tr>
</tbody>
</table>

3  Privacy requirements engineering

3.1  Definition of privacy

Privacy can mean different things to different people. In this context, we define privacy as the ability of an individual to control his/her own information (Turkington, 1990) rather than as ‘the right to be left alone’, which was the first definition of ‘privacy rights’ published by Warren and Brandeis (1890). The reason is that using a passive definition such as the latter makes it difficult to sufficiently protect personal rights as technology advances.

The Organisation for Economic Cooperation and Development (OECD) guidelines on the protection of privacy and transborder data flows of personal Data (1980) is one of the best-known sets of fair information practice principles. The guidelines have widespread acceptance and serve as the fundamental concepts underlying legislation all over the world. However, the perception of what is private versus what is public changes with geography, time, and technology. Thus each country has different laws for protecting privacy. It is not likely a single generic privacy policy can be created to cover all personal information in the world, even though this would be ideal for information trading.

3.2  Privacy-protection laws

The principal mechanisms for ensuring privacy protection to individual data subjects are not only technical but also legislative and administrative. The significant difference between security and privacy protection is that threats to individual privacy often arise from authorised users of the system rather than from unauthorised users. In cases like this, security is not breached but privacy is. Therefore, a privacy-protection policy is needed to keep authorised users from making unauthorised use of personal information.

National-level privacy protection legislation is now in force more than 30 countries. Basically, the aim of the privacy laws is to prevent what are considered to be violations of fundamental human rights, such as the unlawful storage of personal data, the storage of
Integrating privacy requirements considerations

Inaccurate personal data, or the abuse or unauthorised disclosure of such data. In the USA, although there is no explicit, exclusive use of the term ‘privacy’ in the constitution, privacy protection is enforced through many federal laws, such as the Privacy Act, the Right to Financial Privacy Act, the Electronic Communications Privacy Act, the Computer Matching and Privacy Protection Act, the Family Educational Rights and Privacy Act, the Privacy Protection Act, the Cable Communications Policy Act, the Video Privacy Protection Act, the Health Insurance Portability and Accountability Act (HIPAA), the Children’s Online Privacy Protection Act, and the Gramm-Leech-Bliley Act. Moreover, there are more than one hundred state laws to protect privacy (NCSL, 2007). In the European Union, the European Commission’s Directive on Data Protection, which went into effect in 1998, is a more regulatory model and is more focused on civil and human rights than the USA laws. In Japan, the Act on the Protection of Personal Information (2003) has been stringently enforced since 2005.

3.3 Privacy requirements engineering

Privacy requirements should meet the needs of customers and comply with standards and also service policies. For example, the US Privacy Act includes the general requirement that “agencies shall maintain all records with such accuracy, relevance, timeliness, and completeness as is reasonably required to assure fairness to individuals in determinations.” On the other hand, a company may define its privacy requirements for a system more strictly, as, for example, “the application shall not allow unauthorised individuals or programs access to any communications” or “the system shall provide personnel information only to members of the Human Resources Department” (Haley et al., 2007).

A privacy requirement is usually defined as a part of an organisation’s security requirements. However, in most cases of requirements engineering, engineers tend to define requirements in ambiguous terms such as “The system shall protect the privacy of external communications with users” (Wood and Silver, 1989). They do this because they lack adequate guidelines or tools so that they can create complete privacy requirements. Good requirements should be verifiable and purpose-specific (Allen et al., 2008).

What happens if engineers fail to develop adequate privacy requirements? Requirements influence overall implementation. Therefore, without adequate privacy requirements, systems may not prevent violations of human rights, such as the leaking of personal information. The legal and economic penalties for such violations can be serious.

However, security requirements engineering has naturally focused more on security than on privacy. Laws, guidelines, and policies have important roles in the privacy area. Therefore, in order to help developers have a better understanding of how to build privacy functionality into their systems, requirements engineering should address security and privacy in a unified way.

3.4 Problems with privacy requirements

As implementing security in software solutions is now viewed as a responsibility of IT architects, privacy will be the next new responsibility of designers of software. New U.S. laws like Sarbanes-Oxley and HIPAA, with their heavy emphasis on data privacy, are
forcing software developers to pay closer attention to privacy protection for their products. However, most of them have difficulties in developing adequately privacy-compliant systems. For example, a study (Ponemon and IAPP, 2004) reported that although 98% of the companies in their survey have a privacy policy, 52% of them think that they do not have the resources to adequately protect privacy. As we mentioned in the previous section, privacy requirements engineering is the key to mitigate the situation. However, it is not always addressed in actual software development projects. The potential reasons are as follows.

First, privacy requirements engineering should be done with wide and deep understanding of legislative requirements, standards, and policies, but requirements engineers usually don’t have expertise in these areas. It has always been difficult to bridge the gap between legal language and computer language, especially when legal obligations have to be converted into requirements to be enforced by the IT infrastructure.

Second, the features of new IT technology applications set the stage for potential privacy protection problems. Personalisation in websites has brought up several controversial topics about privacy violation. For example, several search engines store search queries and click behaviour in order to infer a person’s favourites. As they also capture the customer’s name and credit card numbers, a security breach may cause serious privacy violations. Ubiquitous social computing systems also bring several threats to user privacy. Motahari et al. (2007) categorised these as seven threats:

- Inappropriate use by administrators. For example, the system admin sells personal data without permission.
- Legal obligations. The system administrator is forced by an organisation such as the police to reveal personal data.
- Inadequate security
- Designed invasion (poor features). For example, a cell phone application that reveals location to friends, but does this without informing the user or providing control of this feature.
- Social inference through lack of entropy
- Social inference through persistent user observation. For example, Bob is often in Alice’s office. Their relationship must be romantic.
- Social leveraging of privileged data. For example, David can’t access my location, but Jane can. David asks Jane for my location.

Murakami (2004) also points out some privacy issues by the use of RFID tags in the ubiquitous information society.

Third, there are several newly emerging privacy enhancing technologies, and software engineers often cannot keep up with these techniques. Privacy enhancing technologies include encryption, anonymity schemes, privacy preserving data mining, tools to remove, change, hide, or blur data, and tools to inform users after collection. Therefore, a new mechanism is needed for privacy requirements engineering that will be palatable to requirements engineers and suitable for use in all software development.
4 Brief results of privacy literature survey

We examined a number of existing approaches for identifying security and privacy requirements. A fuller discussion of the literature survey can be found in (Miyazaki, 2008).

4.1 Goal-based requirements analysis method

The goal-based requirements analysis method (Anton et al., 2001) is a systematic approach to identifying system and enterprise goals and requirements. It is useful for identifying the goals that software systems must achieve, managing trade-offs among the goals, and converting them into operational requirements. The method has been applied to the analysis of electronic commerce applications (Anton and Potts, 1998) and to mine privacy policies for system goals and requirements (Anton and Earp, 2001).

4.2 Pattern-based approach

A pattern-based approach (Yoder and Barcalow, 1997; Schumacher, 2003) has been incorporated into software engineering as a method for object-based reuse. With this approach, security patterns are essentially best practices presented in a template format. This format aids designers in identifying and understanding security concerns and in implementing appropriate security measures.

4.3 E-commerce personalisation approach

Cranor (2003) proposed a number of approaches that may be helpful depending on the functionality of e-commerce personalisation systems. Although the paper mentions that no simple universal formula exists for designing a privacy-protective e-commerce personalisation system, it shows some useful rules of relationship between design of a personalisation system and privacy principles.

4.4 Security requirements elicitation techniques

The following list identifies several techniques that could be considered for eliciting security and privacy requirements (Mead et al., 2005). Some have been developed specifically with security in mind, whereas others have been used for traditional requirements engineering and could potentially be extended to security and privacy requirements.

- misuse cases (Sindre and Opdahl, 2000; Alexander, 2003)
- soft systems methodology (SSM) (Checkland, 1989)
- quality function deployment (QFD) (ASI, 1986)
- controlled requirements expression (CORE) (SDS, 1985)
- issue-based information systems (IBIS) (Conklin and Begeman, 1988)
- joint application development (JAD) (Wood and Silver, 1989)
feature-oriented domain analysis (FODA) (Kang et al., 1990)

critical discourse analysis (CDA) (Checkland, 1989)

accelerated requirements method (ARM) (Hubbard et al., 2000)

reusable legal requirements.

Earlier work in developing a database of reusable legal requirements (Toval et al., 2002) and more recent work by Otto and Anton (2007) seemed promising in terms of capturing privacy requirements and creating a repository of reusable requirements that could be tailored to project needs with a suitable selection mechanism. We decided to build on this work in our approach to privacy requirements elicitation. In our case, we use a questionnaire to help select the appropriate requirements for a given legal environment, and we incorporate the approach into the existing SQUARE method, as will be seen below.

5 The privacy requirements elicitation technique and tool

Requirements engineers need a methodology that can accelerate the privacy requirements engineering process for requirements engineers and stakeholders. The methodology must help prevent requirements leaks, especially for those who are not accustomed to dealing with privacy laws and techniques.

Our objective is to propose an efficient methodology that will help software engineers and stakeholders elicit privacy requirements in order to develop adequately privacy-compliant information systems and applications. To do this, we propose using the computer-aided privacy requirements elicitation technique. This technique is based not only on the discussion between requirements engineers and stakeholders but also on the collaboration between requirements engineers, stakeholders, and a computer-aided tool. The tool uses a questionnaire to elicit development system information. After requirements engineers and stakeholders discuss privacy requirements and reach consensus regarding them, they fill out the questionnaire. The tool then searches for the appropriate privacy requirements in its privacy requirements database and displays the results. The requirements engineers and stakeholders then adapt these requirements to the particular requirements of the system.

In order to achieve the desired result of the methodology, we need to address the following two issues:

1. How do we collect privacy requirements for the database? What are adequate privacy requirements?

2. What questions does the tool need to use in order to tailor the privacy requirements to the project? What decision process is used to select the requirements?

To address the first issue, we reviewed a number of papers dealing with privacy requirements. We then reviewed some principles and laws. To address the second issue, we debated what questions the tool should ask, focusing on the legal and requirements conditions. We also used seal programs and privacy policy statement generators to create questions.
Integrating privacy requirements considerations

Our scope so far has been limited in that we cover only a limited number of privacy laws. More than 30 countries worldwide have adopted privacy legislation. In the USA, there are numerous federal states laws related to privacy. We selected representative laws and structured the methodology so that it could be expanded easily in the future.

Another limitation is that use of this method should not imply any seal of approval or privacy law compliance. Users may, however, accelerate their elicitation process and improve the quality of the resultant requirements when they use this method as part of the process of developing their privacy requirements.

5.1 Collecting privacy requirements

We surveyed privacy principles and privacy laws, as well as the existing project requirements and papers about privacy requirements. Most of the laws do not intend to define system requirements, and in most cases, they restrict or endorse the behaviour of the service provider. Therefore, in order to collect privacy requirements, we first had to research the laws and translate them into a format which would become part of a privacy requirements database. In this proof of concept study, we reviewed the derived requirements among the project members, but we did not do a rigorous public review of the derived requirements to ensure that each one was an exact match for the source of the requirement. Such a public review would be needed for this to become an operational product. For our study, we derived privacy requirements from the following sources:

- Japanese PIPA (Act on the Protection of Personal Information)
- Laws in the USA such as the Privacy Protection Act (USC, 1980), the Video Privacy Protection Act (USC, 1988), CA_SB_1386 (2002), the Family Educational Rights and Privacy Act (U.S. Congress, 2004), the HIPAA (U.S. Congress, 1996), and the Children’s Online Privacy Protection Act (U.S. Congress, 1998)
- W3C Web Services Architecture Requirements (W3C, 2004)
- Misuse cases that we created to support more specific technical requirements (Miyazaki, 2008)

We provide one example below.

5.1.1 OECD guidelines on the protection of privacy

Over the past three decades, several sets of principles have been developed for protecting privacy when using personal information, including the OECD guidelines on the protection of privacy and transborder data flows of personal data (1980). Many other sets of guidelines and privacy laws are based on these principles, so use of OECD is a reasonable starting point for a privacy-requirements database.

The eight OECD privacy principles provide a useful framework for analysing privacy issues:
Each principle can be expanded further in the context of a specific application. For example, Patrick and Kenny have performed an analysis and made detailed user interface design recommendations for an internet job search tool (2003). Cranor adopted the principles to e-commerce personalization (2003). We developed some general privacy requirements from OECD privacy principles. Note that because these are general privacy requirements, which must be further refined, they cannot necessarily be used directly on a project. Because the exact project context is absent, some of these general privacy requirements can be ambiguous and subject to further project specification. On a specific project, we would know the service provider, the nature of the personal data, what is meant by phrases such as ‘where appropriate’, ‘to the extent necessary’, and so on. Nevertheless, we believe that if we can easily identify these general privacy requirements, we will have saved the projects a considerable amount of time. Table 2, which became part of our database, show these general privacy requirements along with their corresponding principles.

### Table 2 Requirements derived from OECD privacy principles

<table>
<thead>
<tr>
<th>ID</th>
<th>OECD guidelines on the protection of privacy and transborder flows of personal data</th>
<th>Privacy requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD_PP_P7</td>
<td>There should be limits to the collection of personal data, and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject (collection limitation principle).</td>
<td>The service provider shall limit the collection of personal data and obtain such data by lawful and fair means.</td>
</tr>
<tr>
<td>OECD_PP_P8</td>
<td>Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete, and kept up-to-date (data quality principle).</td>
<td>Personal data shall be accurate, complete, and kept up-to-date, if possible.</td>
</tr>
<tr>
<td>OECD_PP_P9</td>
<td>The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose (purpose specification principle).</td>
<td>Before collecting personal data, the service provider shall specify the purpose.</td>
</tr>
</tbody>
</table>
Table 2 Requirements derived from OECD privacy principles (continued)

<table>
<thead>
<tr>
<th>ID</th>
<th>OECD guidelines on the protection of privacy and transborder flows of personal data</th>
<th>Privacy requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD_PP_P10</td>
<td>Personal data should not be disclosed, made available, or otherwise used for purposes other than those specified in accordance with purpose specification principle of the OECD Privacy Guidelines except (a) with the consent of the data subject or (b) by the authority of law (use limitation principle).</td>
<td>The service provider shall disclose personal data only with the consent of the data subject or by the authority of law.</td>
</tr>
<tr>
<td>OECD_PP_P11</td>
<td>Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification, or disclosure of data (security safeguards principle).</td>
<td>Personal data should be protected by reasonable security safeguards against such risks as loss, unauthorised access, destruction, use, modification, or disclosure of data.</td>
</tr>
<tr>
<td>OECD_PP_P12</td>
<td>There should be a general policy of openness about developments, practices, and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller (openness principle).</td>
<td>The system shall provide a mechanism by which users can verify their data.</td>
</tr>
<tr>
<td>OECD_PP_P13</td>
<td>An individual should have the right (a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him; (b) to have communicated to him data relating to him within a reasonable time, at a charge, if any, that is not excessive, in a reasonable manner, and in a form that is readily intelligible to him; (c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and (d) to challenge data relating to him and, if the challenge is successful to have the data erased, rectified, completed, or amended (individual participation principle).</td>
<td>The system shall provide a mechanism by which users can verify their data.</td>
</tr>
<tr>
<td>OECD_PP_P14</td>
<td>A data controller should be accountable for complying with measures which give effect to the principles stated above (accountability principle).</td>
<td></td>
</tr>
</tbody>
</table>

5.2 Designing a questionnaire and a requirements decision process

In this section, we describe how we developed the questionnaire that privacy requirements elicitation technique and tool (PRET) uses to choose proper requirements. Although questionnaires are not necessarily an optimal approach for general requirements elicitation (Hickey and Davis, 2004) we believe that they can be used effectively for an area such as privacy, which is fairly well bounded by laws and policies. A significant condition is the law coverage condition, such as the country and type of data controller, because most of the collected requirements in our database are derived from laws.
Privacy policy is another important factor. Policy defines what data they process, how they collect data, and so forth. We have to think about what questions are needed and can be handled with the tool.

5.2.1 Sources of privacy information

Privacy seal organisations, such as TRUSTe (2009) and PrivacyMark (2009), provide a privacy seal and a guarantee for a website. In reality, these privacy seals simply ensure that they have reviewed the licensee’s privacy policy for disclosure of the following uses of information by a website:

- What personal information is being gathered?
- How the information will be used?
- Who the information will be shared with and the choices available regarding how collected information is used?
- Safeguards in place to protect personal information from loss, misuse, or alteration?
- How individuals can update or correct inaccuracies in information collected about them?

Privacy seal organisations do not disclose their detailed criteria, but their criteria are based on OECD privacy principles.

Another source is the OECD privacy statement generator (2000). The OECD has developed the generator as a tool to provide users with useful input in the development of a privacy policy and statement. The Generator makes use of a questionnaire to learn about the website’s personal data practices. Warning flags appear where appropriate. The answers are then fed into a preformatted draft policy statement.

We referred to privacy seal programs and the OECD privacy statement generator to design PRET tool questions and a decision process. However, there are critical differences between our proposal and these references, as follows:

- These references are for dealing with the privacy policy of an existing website. The PRET tool is for generating privacy requirements for developing a software system.
- These references basically use OECD principles in their decision process. The PRET tool also uses laws and empirical requirements.

5.2.2 Designing a questionnaire

By adding and deleting questions from the sources we used, we defined ten questions:

Q1 Does the service provider process personal information?
- Yes / No

Q2 In which country (e.g., Canada) or area (e.g., EU) is the service provided?
- USA / EU / Canada / Japan / Other

Q3 What type of service provider?
- Industrial / Governmental / Academic / Other
Integrating privacy requirements considerations

Q3.1 If industrial, does the service provider belong to any of these fields?
   • Medicine / Communication / Education

Q3.2 If governmental, does the service provider belong to any of these fields?
   • Military branch / Non-military branch / Research body

Q3.3 Is the purpose of the service related to journalism, literary work, academic studies, religious activities, or political activities?
   • Yes / No

Q4 What kind of personal information does the service provider process?
   • Point of contact / Social identification / Personal identity data / Demographic information / Age, Education / Health information / Financial information / Personal information of children / Other sensitive personal data

Q5 How does the service provider obtain personal information?
   • Provided by users / Provided by third parties / Collected automatically from users / Collected automatically from third parties

Q6 Where does the service provider store personal information?
   • Client side / Server side / Third party client side / Third party server side

Q7 How long does the service provider store personal information?
   • Does not store / One transaction / Certain period of time / Forever

Q8 Does the service provider use personal information for another purpose?
   • Yes / No

Q9 Does the service provider share personal information with others?
   • Yes / No

Q10 What privacy protection level does the service provider set?
   • High / Mid / Low

5.2.3 Elicitation decision process

Based on the answers to the above questions, the tool selects a set of requirements from the database. The tool also outputs a priority level for each requirement. In the current software, we assume that a requirement derived from any law has high priority, a requirement derived from principles or misuse cases has medium priority, and a requirement derived from using privacy enhancing techniques has low priority.

6 PRET tool development

In this section, we describe the development of the privacy requirements elicitation technique tool and its integration with the SQUARE method and tool.
6.1 PRET tool requirements

6.1.1 Functional requirements

1. The system should have a process to determine the candidates for general privacy requirements.
   - This function aids users in exhaustively identifying general privacy requirements.

2. The system should be integrated into the SQUARE process.
   - This provides a well-organised and structured approach for addressing privacy concerns in the early stages of the software development lifecycle.

3. The system should have a user interface that enables users to select or modify the requirements.
   - This specification helps users to adapt the requirements to better fit their project.

6.1.2 Non-functional requirements

1. The system should handle the system information correctly.
   - This function enables users to elicit proper requirements that are appropriate for the project.

2. The system should have suitable knowledge and output requirements.
   - This function enables users to elicit proper requirements that are adequate for the project.

3. The system should provide an efficient user interface.
   - This specification helps users to adapt the requirements to better fit their project. The requirements can be categorised by levels (system, software, etc.) and prioritised (SQUARE Step 7: categorise requirements by level and Step 8: prioritise requirements).

4. It must accommodate those who are not familiar with laws, policies, or even special topics about privacy-enhancing technology.

5. The system should have an extensible database for requirements elicitation.

6.2 Integration into the SQUARE methodology

In SQUARE, the requirements engineering team uses elicitation techniques that provide detailed guidance on how to perform the elicitation. The proposed system can be one of the elicitation techniques that produce requirements in Step 6. The relationship between the SQUARE tool and the PRET tool is shown below.
7 PRET results and future plans

We did some usability testing and functional testing of the tool. We tried the tool on the following academic examples:

1 Auto insurance service
   - Suppose there is a domestic car insurance web service which interacts with a user who wants to buy his car insurance, gathers personal information from its website, stores data to its database, and then calculates customer-tailored insurance prices.

2 Smart healthcare ring
   - Suppose there is a ring that collects blood pressure and other vital information from a user every minute, and then sends data to a server. When the server finds abnormal data behaviour, it notifies a third party such as a hospital.

3 Ubiquitous video monitor system
   - Suppose there are many monitor cameras in a town that capture the images of walkers. The system provides motion pictures when its user requests them. Before providing them, the system converts the pictures in order to preserve the third person’s privacy.
We did our own review of the usability and effectiveness of the tool, and also gathered input from outside expert reviewers. Tables 3 and 4 are example inputs and results for the auto insurance service example. Once we had the inputs from the questionnaire shown in Table 3, we were able to do a table lookup to identify matching requirements in our requirements database. Table 4 provides the text of the requirement, the source in the database (derivation), the questions that were used in identifying the requirement, and the priority. Additional examples and discussion of the review process can be found in (Miyazaki, 2008).

Once these high-level requirements are obtained, the user of the tool can adopt all of the requirements, or a subset, and make them project-specific. However, a significant amount of work is saved by providing the user with a set of high-level privacy requirements in a relatively painless way.

In addition to conducting the case studies, we also consulted with user interface experts and other subject matter experts outside the project on the effectiveness of the prototype tool. Finally, we did a subjective comparison of PRET with other requirements elicitation methods, and identified potential areas for future work.

Table 3  Inputs for auto insurance service

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1  Does the service provider process personal information?</td>
<td>Yes</td>
</tr>
<tr>
<td>Q2  In which country or area is the service provided?</td>
<td>USA</td>
</tr>
<tr>
<td>Q3  What type of service provider?</td>
<td>Industrial</td>
</tr>
<tr>
<td>Q3.1 If Industrial, does the service provider belong to any of these fields?</td>
<td></td>
</tr>
<tr>
<td>Q3.2 If Governmental, does the service provider belong to any of these fields?</td>
<td></td>
</tr>
<tr>
<td>Q3.3 Is the purpose of the service related to journalism, literary work, academic studies, religious activities, or political activities?</td>
<td>No</td>
</tr>
<tr>
<td>Q4  What kind of personal information does the service provider process?</td>
<td>Point, social, demographic, age</td>
</tr>
<tr>
<td>Q5  How does the service provider obtain personal information?</td>
<td>Provided by users, provided by third parties</td>
</tr>
<tr>
<td>Q6  Where does the service provider store personal information?</td>
<td>Server side</td>
</tr>
<tr>
<td>Q7  How long does the service provider store personal information?</td>
<td>Forever</td>
</tr>
<tr>
<td>Q8  Does the service provider use personal information for another purpose?</td>
<td>No</td>
</tr>
<tr>
<td>Q9  Does the service provider share personal information with others?</td>
<td>No</td>
</tr>
<tr>
<td>Q10 What privacy protection level does the service provider set?</td>
<td>Mid</td>
</tr>
</tbody>
</table>
Table 4  Results for auto insurance service

<table>
<thead>
<tr>
<th>Privacy requirements</th>
<th>Derivation</th>
<th>Explanation</th>
<th>Priority level</th>
</tr>
</thead>
<tbody>
<tr>
<td>The service architecture shall describe privacy policy statements and enable a user to access them.</td>
<td>W3C_AR020.1,20.3</td>
<td>Personal data usage (Q1, Q2)</td>
<td>Mid</td>
</tr>
<tr>
<td>Before collecting personal data, the data controller shall specify the purpose.</td>
<td>OECD_PP_P9</td>
<td>Personal data usage (Q1)</td>
<td>Mid</td>
</tr>
<tr>
<td>The service provider shall limit the collection of personal data and obtain such data by lawful and fair means.</td>
<td>OECD_PP_P7</td>
<td>Personal data collection (Q6)</td>
<td>Mid</td>
</tr>
<tr>
<td>The system network communications must be protected from unauthorised information gathering and/or eavesdropping.</td>
<td>Misuse_case_1</td>
<td>Personal data collection (Q6)</td>
<td>Mid</td>
</tr>
<tr>
<td>The system should have functional audit logs and usage reports without disclosing identity information.</td>
<td>Misuse_case_2</td>
<td>Personal data collection (Q6)</td>
<td>Mid</td>
</tr>
<tr>
<td>The system shall have strong authentication measures in place at all system gateways and entrance points.</td>
<td>Misuse_case_3</td>
<td>Personal data storage (Q7)</td>
<td>Mid</td>
</tr>
<tr>
<td>Personal data should be protected by reasonable security safeguards against such risks as loss, unauthorised access, destruction, use, modification or disclosure of data.</td>
<td>OECD_PP_P11</td>
<td>Personal data storage (Q7)</td>
<td>Mid</td>
</tr>
<tr>
<td>Personal data shall be accurate, complete and kept up-to-date, if it is possible.</td>
<td>OECD_PP_P8</td>
<td>Personal data storage (Q7)</td>
<td>Mid</td>
</tr>
<tr>
<td>The system shall provide a mechanism by which users can verify their data.</td>
<td>OECD_PP_P13</td>
<td>Personal data storage (Q7)</td>
<td>Mid</td>
</tr>
<tr>
<td>The system shall provide a data backup mechanism.</td>
<td>Misuse_case_4</td>
<td>Personal data storage (Q7)</td>
<td>Mid</td>
</tr>
<tr>
<td>The system shall have a verification process to check whether there is a disclosure agreement between the third party and the person.</td>
<td>Misuse_case_5</td>
<td>Personal data collection from the third party (Q5)</td>
<td>Mid</td>
</tr>
<tr>
<td>The service provider shall report to all the customers if the privacy information is breached.</td>
<td>CA_SB_1386</td>
<td>Breach report in JP, USA (Q1,Q2, Q3)</td>
<td>High</td>
</tr>
</tbody>
</table>

7.1  Future plans for privacy integration with SQUARE

Although these early results are interesting, we need to get more experience with generation of privacy requirements on real projects using this approach. It is not clear to us whether the questionnaire approach will hold up, in view of the fact that the current questionnaire does not cover all possible cases, and it is based on our interpretation and consolidation of the underlying laws. Also the questionnaire runs the risk of becoming
obsolete if the underlying laws change and the questionnaire are not revised to stay current with those changes.

Another approach might be to instantiate the questionnaire and the underlying laws to be project-specific at the outset and analyse them, either manually or in some automated fashion, to produce a project-specific set of requirements.

We are especially interested in addressing the combination of security and privacy on real projects in the context of SQUARE. We also note that there are other methods for eliciting privacy requirements, such as PriS (Kalloniatis et al., 2008) that could be explored further.

We can see that there would be more benefit from fuller integration of privacy concerns with the SQUARE method and tool. At present we are investigating suitable risk assessment methods for privacy (Step 4 of SQUARE) and doing further SQUARE tool development to incorporate both security and privacy requirements engineering in a more integrated way. We expect this to be reflected back into the method and plan to do industry case studies with this combined method and tool. We hope that in the future we can offer a single integrated method to address both security and privacy.

References

Integrating privacy requirements considerations


