Bart Preneel
Demosthenes Ikonomou (Eds.)

LNCS 8319

Privacy Technologies and Policy

First Annual Privacy Forum, APF 2012
Limassol, Cyprus, October 2012
Revised Selected Papers

Springer

Lecture Notes in Computer Science 8319

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Bart Preneel Demosthenes Ikonomou (Eds.)

Privacy Technologies and Policy

First Annual Privacy Forum, APF 2012
Limassol, Cyprus, October 10-11, 2012
Revised Selected Papers

Springer

Volume Editors

Bart Preneel
KU Leuven and iMinds
Department of Electrical Engineering-ESAT
Kasteelpark Arenberg 10 Bus 2452, 3001 Leuven-Heverlee, Belgium
E-mail: bart.preneel@esat.kuleuven.be

Demosthenes Ikonomou
ENISA
Information Security & Data Protection Unit
1 Vasilissis Sofias, Marousi, 15124 Athens, Greece
E-mail: demosthenes.ikonomou@enisa.europa.eu

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-54068-4 e-ISBN 978-3-642-54069-1
DOI 10.1007/978-3-642-54069-1
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013958070

CR Subject Classification (1998): K.4, K.6, H.4

LNCS Sublibrary: SL 4 – Security and Cryptology

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The First Annual Privacy Forum (APF 2012) was held in Limassol, Cyprus, during October 10–11, 2012. The forum was co-organized by the European Network and Information Security Agency (ENISA) and the European Commission Directorate General for Communications Networks, Content and Technology (DG CONNECT), with the support of the Department of Computer Science of the University of Cyprus. APF 2012 was endorsed as an official event of the Cyprus Presidency of the Council of the European Union.

As ICT technologies develop, they put a rapidly growing number of services and tools in the hands of users, companies, and governments. This trend started accelerating with the widespread use of the Internet, developed with the Web 2.0 concept, and is currently evolving toward the Internet of Things (also known as ambient intelligence, pervasive computing, or ubiquitous computing). The hot topic today in business events is "big data"; this refers to the massive availability of data on all aspects of society. There is no doubt that these developments will transform society, with many beneficial effects on the quality of our lives. However, one of the main concerns is that these developments have a tendency to systematically erode our privacy. Addressing this challenge requires a deep understanding of the political, legal, sociological, psychological, and technical aspects of these problems.

While there are many scientific events dedicated to privacy and privacy technologies, there is a need for an event at a European level that brings together key decision-makers and scientists to discuss the latest developments. In order to achieve this mix, the program of APF 2012 had an unusual combination: it consisted of invited speakers and panels interleaved with a set of contributed papers that had undergone a scientific review process. But in contrast with most scientific events, researchers were encouraged to submit positions papers or overview papers that offered a broader perspective on their research.

As a result of the Call for Papers, 26 papers were submitted; after a thorough review by the members of the scientific Program Committee, 20 papers were accepted for presentation at APF 2012. Following the completion of the Forum, 13 papers were revised by the authors and selected for inclusion in these proceedings.

Among the recommendations of the Forum are the need for more privacy-respecting tools, that research should seek innovative tools to empower users by enhancing transparency, that empirical understanding of data flows should be the starting point for actors' agendas, and also that data protection agencies should analyze market failures in privacy technology and intervene with scientific and economic precision.

Several people have contributed to the success of APF 2012. First we would like to thank all the presenters, as well as the authors who submitted their work. We sincerely thank all the Program Committee members, who volunteered to review the papers and discuss the comments. APF 2012 would not have been such a success without the tireless contribution of the staff of ENISA. We would also like to thank the colleagues at the European Commission DG CONNECT as well as the Computer Science Department of the University of Cyprus and in particular Prof. Marios Dikaiakos and Mrs. Maria Poveda for their continuous support and collaboration throughout the organization of this event. Our gratitude is also extended to the Cyprus Presidency of the EU Council for endorsing APF 2012 as one of the official events of the Presidency. Finally we want to express our gratitude to the Business Software Alliance (BSA), Austrian Airlines, and the Cyprus Telecommunication Authority (CYTA) for their support. We would also like to thank our partners NESSOS, CEPIS, the Cyprus Computer Society and EGI.

We hope that this forum can play a stimulating role in the European and international privacy community — offering a forum for the exchange of views and ideas between policymakers, research communities, and industry.

October 2012

Demosthenes Ikonomou
Bart Preneel

Organization

Annual Privacy Forum
Limassol, Cyprus, October 10–11, 2012

Organized by

European Network and Information Security Agency (ENISA)
European Commission Directorate General for Communications Networks, Content and Technology (DG CONNECT)
Department of Computer Science of the University of Cyprus

General Co-chairs

Giuseppe Abbamonte	European Commission (DG CONNECT Unit H4 Trust and Security)
Demosthenes Ikonomou	ENISA
Marios Dikaiakos	University of Cyprus

Organizing Committee

Santiago Alvarez	ENISA
Daria Catalui	ENISA
Slawomir Gorniak	ENISA
Martin Mühleck	DG CONNECT
Maria Poveda	University of Cyprus
Rodica Tirtea	ENISA

Program Chair

Bart Preneel	KU Leuven and iMinds

Program Committee

Alessandro Acquisti	Carnegie Mellon University, USA
Andreas Albers	University of Frankfurt, Germany
Elisa Bertino	University of Purdue, USA
Rainer Böhme	University of Münster, Germany
Caspar Bowden	Independent expert
Jacques Bus	University of Luxembourg, Luxembourg
Jan Camenisch	IBM Zurich Research Laboratory, Switzerland
Claude Castelluccia	Inria, France
George Danezis	Microsoft Research Cambridge, UK
Claudia Diaz	COSIC KU Leuven, Belgium
Paul de Hert	University of Tilburg, The Netherlands and Vrije Universiteit Brussels, Belgium
Ioanna Dionysiou	University of Nicosia, Cyprus
Tassos Dimitriou	Athens Information Technology – AIT, Greece
Elena Ferrari	University of Insubria, Italy
Simone Fischer-Hübner	University of Karlstad, Sweden
Paul Francis	Max Planck Institute for Software Systems, Germany
Sotiris Ioannidis	FORTH, Greece
Nicola Jentzsch	DIW Berlin, Germany
Sokratis Katsikas	University of Piraeus, Greece
Florian Kerschbaum	TU Dresden, Germany
Eleni Kosta	ICRI KU Leuven, Belgium
Daniel Le-Metayer	Inria, France
Giannis Marias	Athens University of Economics and Business, Greece
Evangelos Markatos	FORTH, Greece
Periklis Papakonstantinou	Tsinghua University, China
Nick Papanikolaou	HP Labs Bristol, UK
Aljosa Pasic	Atos Research, Spain
Kai Rannenberg	University of Frankfurt, Germany
Stefan Schiffner	TU Darmstadt – CASED, Germany
Rodica Tirtea	ENISA, Greece
Carmela Troncoso	Gradiant, Spain
Claire Vishik	Intel, UK
Nick Wainwright	HP Labs Bristol, UK
Alma Whitten	Google, USA

External Reviewers

Danny De Cock	Stephan Heim
Harald Gjermundrød	Aleksandra Korolova
Seda Gürses	Anja Lehmann

Andreas Pashalidis
Suksant Sae Lor
Rula Sayaf
Jessica Staddon
Markus Tschersich
Fatbardh Veseli

Table of Contents

Modelling

A Problem-Based Approach for Computer-Aided Privacy Threat Identification 1
Kristian Beckers, Stephan Faßbender, Maritta Heisel, and Rene Meis

Conceptual Framework and Architecture for Privacy Audit 17
Ksenya Kveler, Kirsten Bock, Pietro Colombo, Tamar Domany, Elena Ferrari, and Alan Hartman

Privacy by Design

Privacy-Preserving Computation (Position Paper) 41
Florian Kerschbaum

Designing Privacy-by-Design 55
Jeroen van Rest, Daniel Boonstra, Maarten Everts, Martin van Rijn, and Ron van Paassen

Enhancing Privacy by Design from a Developer's Perspective 73
Christoph Bier, Pascal Birnstill, Erik Krempel, Hauke Vagts, and Jürgen Beyerer

A Solution, But Not a Panacea for Defending Privacy: The Challenges, Criticism and Limitations of Privacy by Design 86
Demetrius Klitou

Identity Management

Integrating Anonymous Credentials with eIDs for Privacy-Respecting Online Authentication 111
Ronny Bjones, Ioannis Krontiris, Pascal Paillier, and Kai Rannenberg

Federated Identity as Capabilities 125
Harry Halpin and Blaine Cook

Privacy Preserving Course Evaluations in Greek Higher Education Institutes: An e-Participation Case Study with the Empowerment of Attribute Based Credentials 140
Vasiliki Liagkou, George Metakides, Apostolis Pyrgelis, Christoforos Raptopoulos, Paul Spirakis, and Yannis C. Stamatiou

Case Studies

Collection and Storage of Personal Data: A Critical View on Current Practices in the Transportation Sector 157
Eleni Kosta, Hans Graux, and Jos Dumortier

ICT and Privacy: Barriers 177
Antonio Kung

A Method for Analysing Traceability between Privacy Policies and Privacy Controls of Online Social Networks 187
Pauline Anthonysamy, Phil Greenwood, and Awais Rashid

Electronic Footprints in the Sand: Technologies for Assisting Domestic Violence Survivors 203
Martin Emms, Budi Arief, and Aad van Moorsel

Author Index 215

A Problem-Based Approach for Computer-Aided Privacy Threat Identification*

Kristian Beckers, Stephan Faßbender, Maritta Heisel, and Rene Meis

paluno - The Ruhr Institute for Software Technology – University of Duisburg-Essen
firstname.lastname@paluno.uni-due.de

Abstract. Recently, there has been an increase of reported privacy threats hitting large software systems. These threats can originate from stakeholders that are part of the system. Thus, it is crucial for software engineers to identify these privacy threats, refine these into privacy requirements, and design solutions that mitigate the threats.

In this paper, we introduce our methodology named Problem-Based Privacy Analysis (ProPAn). The ProPAn method is an approach for identifying privacy threats during the requirements analysis of software systems using problem frame models. Our approach does not rely entirely on the privacy analyst to detect privacy threats, but allows a computer aided privacy threat identification that is derived from the relations between stakeholders, technology, and personal information in the system-to-be.

To capture the environment of the system, e.g., stakeholders and other IT systems, we use problem frames, a requirements engineering approach founded on the modeling of a machine (system-to-be) in its environment (e.g. stakeholders, other software). We define a UML profile for privacy requirements and a reasoning technique that identifies stakeholders, whose personal information are stored or transmitted in the system-to-be and stakeholders from whom we have to protect this personal information. We illustrate our approach using an eHealth scenario provided by the industrial partners of the EU project NESSoS.

Keywords: privacy, threat analysis, problem frames, requirements engineering.

1 Introduction

Identifying privacy threats to a software system is difficult, because of a lack of structured approaches for identifying stakeholders that have privacy requirements in a system. In addition, finding methods to fulfill these requirements, and fulfilling the functional requirements of the system-to-be at the same time, is even more challenging.

Westin defines privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" [19]. The privacy specification in the ISO 15408 standard - Common

* This research was partially supported by the EU project Network of Excellence on Engineering Secure Future Internet Software Services and Systems (NESSoS, ICT-2009.1.4 Trustworthy ICT, Grant No. 256980).

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 1–16, 2014.

Criteria for Information Technology Security Evaluation (or short CC) [14] defines four privacy goals. These goals can be refined into privacy requirements for a given software system. *Anonymity* means that a subject is not identifiable within a set of subjects, the anonymity set. *Unlinkability* of two or more items of interest (IOI) means that within a system the attacker cannot sufficiently distinguish whether these IOIs are related or not. *Unobservability* of an IOI means that an IOI is not detectable by any subject uninvolved in it and anonymity of the subject(s) involved in the IOI even against the other subject(s) involved in that IOI. A pseudonym is an identifier of a subject other than one of the subject's real names. Using pseudonyms means *pseudonymity*.

In this paper, we introduce our methodology named Problem-Based Privacy Analysis (ProPAn). The ProPAn method provides assistance for the initial steps of any given privacy analysis, which is to figure out those parts of the system, where personal information, we have to protect, can be disclosed by counterstakeholders. We use the problem frame [15] requirements engineering approach to model the machine (system-to-be) in its environment (e.g. stakeholders, other software). We extend the UML4PF framework [5] with a UML profile for privacy requirements and a reasoning technique. This reasoning technique identifies the domains, in which personal information is stored or to which personal information is transmitted. Additionally, our technique identifies the domains, to which counterstakeholders have access. From these identified domains, our technique derives the possible privacy threats of the system-to-be. We illustrate our approach using an eHealth scenario provided by the industrial partners of the EU project *Network of Excellence (NoE) on Engineering Secure Future Internet Software Services and Systems (NESSoS)*[1].

A number of guidelines for privacy are available. *The Fair Information Practice Principles* (– or short FIPs) [18] – are widely accepted. They state that a person's informed consent is required for the data that is collected, collection should be limited for the task it is required for and erased as soon as this is not the case anymore. The collector of the data shall keep the data secure and shall be held accountable for any violation of these principles. The FIPs were also adapted in the Personal Information Protection and Electronic Documents Act in Canada's private-sector privacy law. In the European Union, the *EU Data Protection Directive*, *Directive 95/46/EC* does not permit processing personal data at all, except when a specific legal basis explicitly allows it or when the individuals concerned consented prior to the data processing [9]. The U.S. have no central data protection law, but separate privacy laws, e.g., the Gramm-Leach-Bliley Act for financial information, the Health Insurance Portability and Accountability Act for medical information, and the Children's Online Privacy Protection Act for data related to children [11]. These legal guidelines must be implemented by any given software system for which the guidelines apply. Our work supports privacy threat analysis, which has to be performed in order to comply with any of these regulations.

The rest of the paper is organized as follows. Section 2 presents the problem frame approach and our support tool, and Sect. 3 presents the eHealth case study. We introduce our method in Sect. 4, and illustrate its application to the case study in Sect. 5. Section 6 contains related work, and Sect. 7 concludes.

[1] `http://www.nessos-project.eu/`

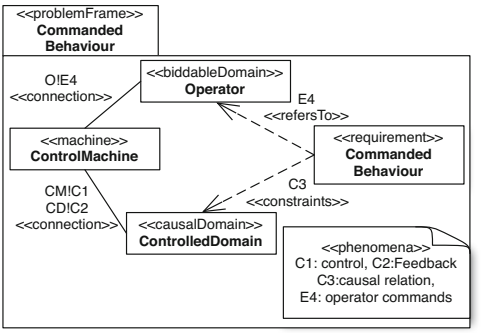


Fig. 1. *Commanded Behaviour* problem frame

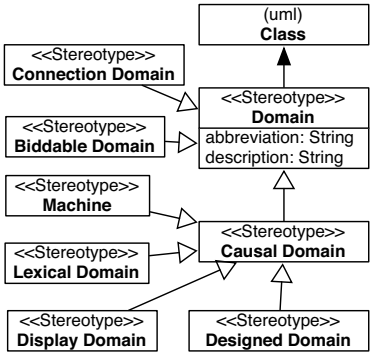


Fig. 2. Inheritance structure of different domain types

2 Background

We use the problem frames approach to build our privacy threat identification on, because problem frames are an appropriate means to analyze not only functional, but also dependability and other quality requirements [12,1].

Problem frames are a means to describe software development problems. They were proposed by Jackson [15], who describes them as follows: *"A problem frame is a kind of pattern. It defines an intuitively identifiable problem class in terms of its context and the characteristics of its domains, interfaces and requirement."*. It is described by a *frame diagram*, which consists of domains, interfaces between them, and a requirement. We describe problem frames using class diagrams extended by stereotypes as proposed by Hatebur and Heisel [13]. All elements of a problem frame diagram act as placeholders, which must be instantiated to represent concrete problems. Doing so, one obtains a problem description that belongs to a specific class of problems.

Figure 1 shows an example of a problem frame. The class with the stereotype ≪machine≫ represents the thing to be developed (e.g., the software). The classes with some domain stereotypes, e.g., ≪causalDomain≫ or ≪biddableDomain≫ represent *problem domains* that already exist in the application environment. Jackson distinguishes the domain types *causal domains* that comply with some physical laws, *lexical domains* that are data representations, and *biddable domains* that are usually people. We use the formal meta model [13] shown in Fig. 2 to annotate domains with their corresponding stereotype.

Domains are connected by interfaces consisting of shared phenomena. Shared phenomena may be events, operation calls, messages, and the like. They are observable by at least two domains, but controlled by only one domain, as indicated by an exclamation mark. For example, in Fig. 1 the notation *O!E4* means that the phenomena in the set *E4* are controlled by the domain **Operator**. These interfaces are represented as associations, and the name of the associations contain the phenomena and the domains controlling the phenomena.

In Fig. 1, the **ControlledDomain** domain is constrained and the **Operator** is referred, because the **ControlMachine** has the role to change the **ControlledDomain** on behalf of the **Operator**'s commands for achieving the required **Commanded Behaviour**. These relationships are modeled using dependencies that are annotated with the corresponding stereotypes.

Problem frames support developers in analyzing problems to be solved. They show what domains have to be considered, and what knowledge must be described and reasoned about when analyzing the problem in depth. Other problem frames besides the commanded behavior frame shown in Fig. 1 are *required behaviour*, *simple workpieces*, *information display*, and *transformation* [15].

Software development with problem frames proceeds as follows: first, the environment in which the machine will operate is represented by a *context diagram*. Like a frame diagram, a context diagram consists of domains and interfaces. However, a context diagram contains no requirements. An example is given in Fig. 3. Then, the problem is decomposed into subproblems. If ever possible, the decomposition is done in such a way that the subproblems fit to given problem frames. To fit a subproblem to a problem frame, one must instantiate its frame diagram, i.e., provide instances for its domains, phenomena, and interfaces. The instantiated frame diagram is called a *problem diagram*. Examples are given in Fig. 4, 5, and 6.

Since the requirements refer to the *environment* in which the machine must operate, the next step consists in deriving a *specification* for the machine (see [16] for details). The specification describes the machine and is the starting point for its construction.

The UML4PF framework provides tool support for this approach. A more detailed description can be found in [5].

3 Case Study

To illustrate our approach for identifying privacy threats, we use a scenario taken from the health care domain. It concerns managing **Electronic Health Records (EHR)**s and is provided by the industrial partners of the EU project NESSoS. EHRs contain any information created by health care professionals in the context of the care of a patient. Examples are laboratory reports, X-ray images, and data from monitoring equipment. The information stored in the EHR shall only be accessed with the consent of the patient. The only exception is a medical emergency, in which case the patient's physical status may prevent her from giving the consent. In addition, the information in the EHR supports clinical research.

In Fig. 3 we present a context diagram of the electronic health system (**EHS**). The **EHS** is the machine to be built and the lexical domain **EHR** is directly connected to it. The **EHS** is also connected to the **Patient** using the **Browser Patient**, a **Mobile Device**, which is further connected to **Sensor**s that are in turn attached to the **Patient**. A **Monitor** or the **Browser Care Provider**s connects the machine to the health care professionals **Nurse** and **Doctor**. The **Researcher** uses the **Browser Researcher** to access the **Research Database Application**, which is in turn connected to the **EHS**.

We identified 19 preliminary functional requirements for the EHS, which were refined into 34 functional requirements and corresponding problem diagrams. For reasons

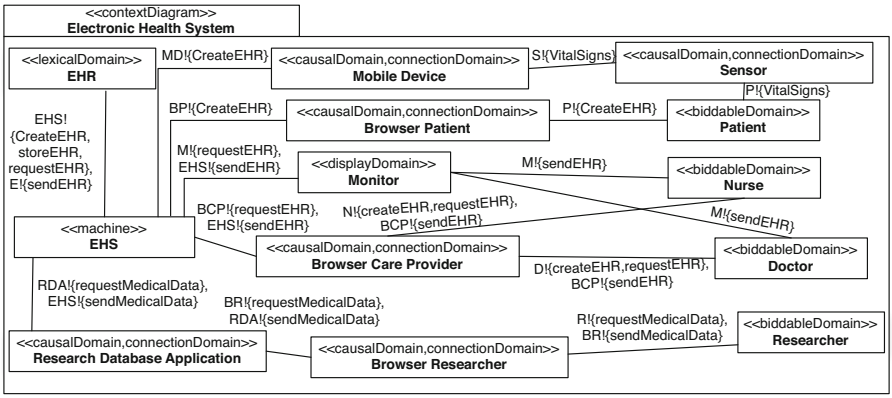


Fig. 3. Context Diagram

of space, we focus on the following requirements for the remainder of the paper, which define the basic functionality of a EHS and include a potential privacy threat:

RQ1.1 Store **EHR**, which are created by care providers.
RQ1.2 Display **EHR** to care providers as needed
RQ15.1 Send alarms, appointments or instructions using **EHR** from **Doctor** to **Patient**
RQ15.2 Send automated alarms or instructions based on **EHR** analysis to **Patient**
RQ16 Release medical data to **Researcher**s

The problem diagram for RQ1.1 describes creating and storing of **EHR**s (depicted in Fig. 4). The **Patient** is connected to a **Sensor** that reports the **Patient**'s vital signs to the **EHR Create & Store Machine** using a **Mobile Device**. The machine stores the **EHR**. In addition, the **Patient** can use the **Browser Patient** to create an **EHR**. **Doctor**s and **Nurse**s can use the **Browser Care Providers** to command the machine to create **EHR**s.

The problem diagram for RQ1.2 shown in Fig. 5 describes how care providers can access the **EHR**. **Doctor**s and **Nurse**s can either use the **Monitor** or the **Browser Care Providers** to request **EHR**s from the **EHR Load Machine**. The machine can access the **EHR** and display it on either the **Monitor** or the **Browser Care Providers** to the **Doctor**s and **Nurse**s.

The release of medical information to researches described in RQ16 and depicted in the problem diagram in Fig. 6. **Researcher**s can use the **Browser Researcher** to request medical data from the **Research Database Application**. This application requests the data in turn from the **ReleaseMedicalDataMachine**, which releases it the **Research Database Application**. The application sends the information to the browser, where it is shown to the **Researcher**s.

The problem diagrams for RQ15.1 and RQ15.2 are drawn in a similar manner.

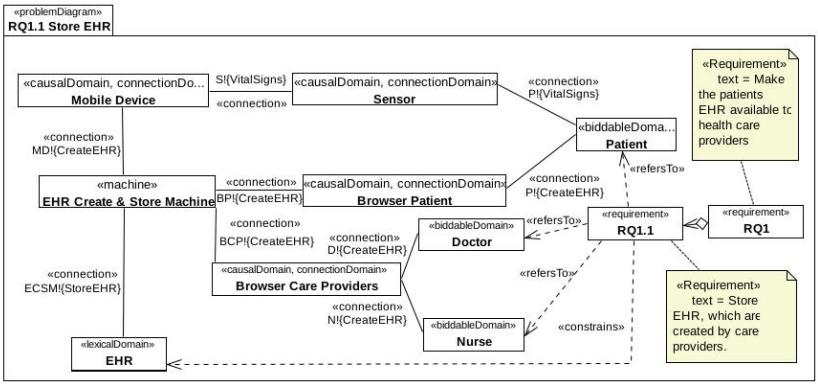


Fig. 4. Problem Diagram for Requirement RQ1.1

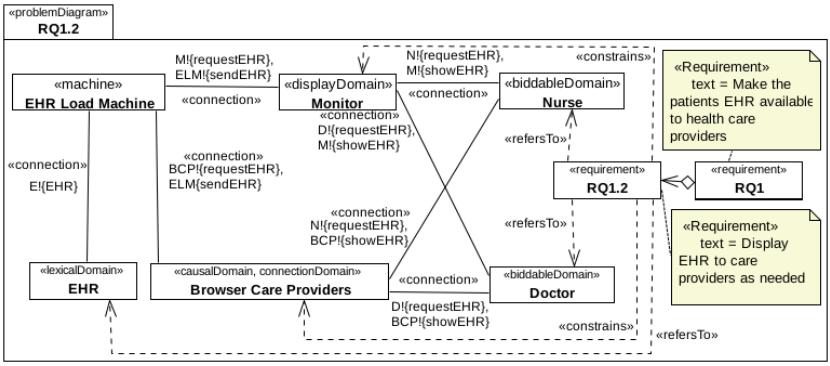


Fig. 5. Problem Diagram for Requirement RQ1.2

4 Method

An overview of the ProPAn method is shown in Fig. 7. It consists of four steps Draw context diagram and problem diagrams, Add privacy requirements to model, Generate privacy threat graphs, and Analyze privacy threat graphs that will be explained in detail in the following.

4.1 Creation of the Model

The first step of the ProPAn method is to Draw context diagram and problem diagrams for the given Set of functional requirements. For this purpose, we use the UML4PF tool. The result of this step is a Model containing context diagram and problem diagrams. This step follows the basic principles of requirements engineering using the problem frames approach as explained in Sect. 2.

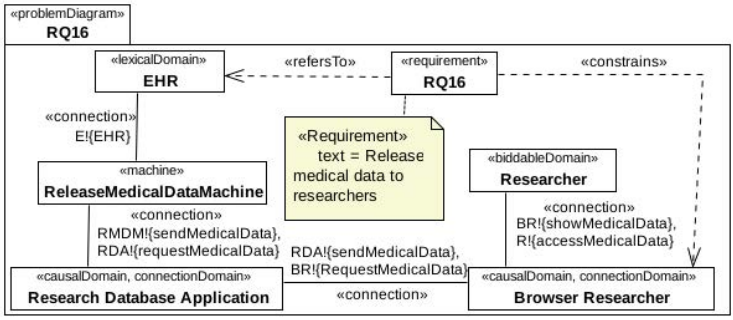


Fig. 6. Problem Diagram for Requirement RQ16

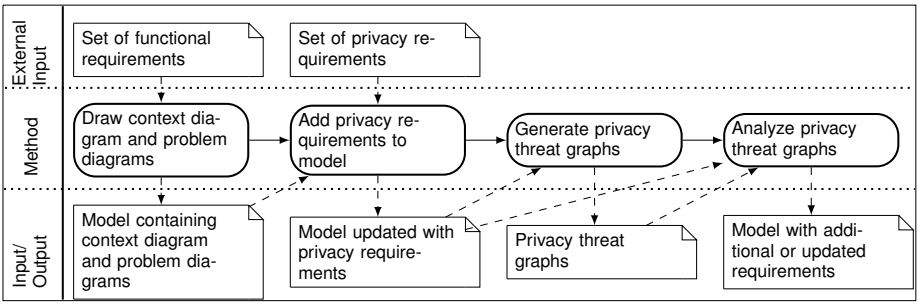


Fig. 7. Illustration of the ProPAn method

4.2 Privacy Requirements

As a second step we consider the Set of privacy requirements the system-to-be shall enforce. To Add privacy requirements to model we introduce a new stereotype. In Fig. 8 the UML profile is shown. We derived this stereotype from the CC's privacy specification. All privacy specifications follow the pattern that they describe whose privacy shall be protected from whom. For our privacy threat identification it is not necessary to distinguish between the different privacy goals, such as anonymity, unlinkability, unobservability, and pseudonymity (see Sect. 2), but the profile offers the opportunity to easily refine our general privacy stereotype. The general privacy stereotype has three attributes. The attribute stakeholder is the biddable domain whose privacy shall be protected against the domain given in the attribute counterStakeholder. In the attribute Description a more detailed textual description of the privacy requirement can be given. Our threat identification focuses on internal counterstakeholders, i.e. domains that occur in a problem diagram. In contrast to the term *"attacker"* a *"counterstakeholder"* may obtain sensitive data about the stakeholder involuntarily.

The Model containing context diagram and problem diagrams is updated to obtain the Model updated with privacy requirements. For our example, we add the privacy

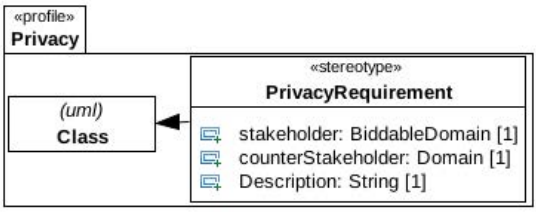


Fig. 8. Privacy Profile

Fig. 9. Privacy Requirement

requirement Preserve anonymity with the stakeholder Patient and the counterstakeholder Researcher (see Fig. 9).

4.3 Graph Generation

All graphs $(\mathcal{V}, \mathcal{E})$ that we use for our threat identification in the ProPAn method are labeled and directed. The set of vertexes is a subset of the domains occurring in the model, formally $\mathcal{V} \subseteq Domain$. The edges are annotated with problem diagrams and point from one domain to another, formally $\mathcal{E} \subseteq Domain \times ProblemDiagram \times Domain$. In the following we describe a graph $(\mathcal{V}, \mathcal{E})$ only by its edges $\mathcal{E}$.

Global Information Flow Graph. To Generate privacy threat graphs we use the Model updated with privacy requirements as an input. From this static model we create the global information flow graph $\mathcal{G}$, which is an over-approximation of the information flow occurring in the system-to-be. An edge (d_1, p, d_2) is in $\mathcal{G}$, iff the domains d_1 and d_2 are not equal, are both part of the problem diagram p, and the domain d_2 is constrained in p. This is expressed using the following formula.

$$\mathcal{G} = \{(d_1, p, d_2) : Domain \times ProblemDiagram \times Domain \mid d_1 \neq d_2 \wedge d_1, d_2 \in p \wedge constr(d_2, p)\}$$

Because of the annotation of the edge we keep the information which problem diagram causes the information flow. Thus, our information flow graph contains traceability links that are used in our further analysis. The semantics of an edge $(d_1, p, d_2) \in \mathcal{G}$ is that in problem diagram p there is possibly an information flow from domain d_1 to domain d_2.

Stakeholder Information Flow Graph. We now want to determine where data of the stakeholder s, whose privacy shall be protected against the counterstakeholder c, is possibly processed or stored. Using the global information flow graph $\mathcal{G}$, we can compute the stakeholder information flow graph $\mathcal{S}_s \subseteq \mathcal{G}$. The algorithm for the computation of $\mathcal{S}_s$ is given in Listing 1.1.

The algorithm operates on four sets. During the algorithm the set D contains all domains from whom possibly an additional information flow can occur. D is initialized as the singleton set containing the stakeholder s. The domains, which have already been

```
1  var D, U : Set[Domain]; P : Set[ProblemDiagram];
2  var E, S_s : Set[Domain × ProblemDiagram × Domain];
3  D := {s}; U := ∅; P := {p : ProblemDiagram}; S_s := ∅;
4  repeat
5    E := ∅;
6    foreach d ∈ D do
7      foreach p ∈ P do
8        E := E ∪ {d' : Domain | (d, p, d') ∈ G • (d, p, d')}
9      endforeach
10   endforeach;
11   U := U ∪ D ;
12   D := {d : Domain | d ∉ U ∧ ∃d' : Domain; p : ProblemDiagram • (d', p, d)∈ E};
13   P := P \ {p : ProblemDiagram | ∃d, d' : Domain • (d,p,d') ∈ E};
14   S_s := S_s ∪ E
15 until D = ∅ ∨ P = ∅ endrepeat
```

Listing 1.1. Algorithm for the computation of the stakeholder information flow graph

used by the algorithm, are collected in the set U, which is initially empty. The set P consists of all problem diagrams, which are not yet part of an information flow in the stakeholder information flow graph. The set P is initialized as the set of all problem diagrams. The set E is used to temporally store the edges, which will be added to the stakeholder information flow graph at the end of each iteration. The set $\mathcal{S}_s$ is the resulting stakeholder information flow graph and is initialized as the empty set of edges. Inside the repeat loop, we firstly initialize the set of new edges E as empty (line 5). We then iterate all domains of D (line 6) and all problem diagrams of P (line 7). All edges from $\mathcal{G}$, which start from a domain in D and are annotated with a problem diagram in P, are added to E (line 8). After the iteration of D and P, all domains from D are added to the used Domains U (line 11). Then the set of domains D is updated to the set of domains, which are reachable from the new edges E, but are not in U, i.e., they have not yet been used (line 12). The set of problem diagrams P is reduced by the set of problem diagrams that occur as annotations in the set of new edges E (line 13). At last, the new edges E are added to the stakeholder information flow graph $\mathcal{S}_s$ (line 14). The algorithm terminates when one of the sets D or P is empty (line 15). This is ensured, because each domain and each problem diagram is considered for at most one execution of the repeat loop.

It is sufficient to consider each problem diagram for at most one execution of the repeat loop, because all information flows that would be added later, are redundant to the existing information flows.

Counterstakeholder Graph. To determine which information the system-to-be provides to the counterstakeholder c, we generate the counterstakeholder graph $\mathcal{C}_c$. An edge (c, p, d) is in $\mathcal{C}_c$ iff the counterstakeholder c and the domain d both occur in the problem diagram p. We express this using the following formula.

$$\mathcal{C}_c = \{(d_1, p, d_2) : Domain \times ProblemDiagram \times Domain \mid d_1 = c \wedge d_1, d_2 \in p\}$$

Please note that the counterstakeholder graph $\mathcal{C}_c$ is not a subgraph of the global information flow graph $\mathcal{G}$. The semantics of an edge $(c, p, d) \in \mathcal{C}_c$ is that the counterstakeholder c may gain information from the domain d in the problem diagram p, which differs from the semantics of an edge in $\mathcal{G}$.

Privacy Threat Graph. Finally, we automatically generate the privacy threat graph $\mathcal{T}_{s,c}$ for the stakeholder s and the counterstakeholder c. This graph connects the stakeholder information flow graph $\mathcal{S}_s$ with the counterstakeholder graph $\mathcal{C}_c$. From $\mathcal{C}_c$ the privacy threat graph $\mathcal{T}_{s,c}$ contains only the edges from $\mathcal{C}_c$, which point to domains, which possibly provide information about the stakeholder s. We call this counterstakeholder subgraph $\mathcal{C}_{c,s} \subseteq \mathcal{C}_c$. The privacy threat graph $\mathcal{T}_{s,c}$ contains only the edges from $\mathcal{S}_s$, which are part of a path from the stakeholder s to a domain that possibly provides information to the counterstakeholder c. We call this stakeholder information flow subgraph $\mathcal{S}_{s,c} \subseteq \mathcal{S}_s$. $\mathcal{T}_{s,c}$ is then the union of $\mathcal{S}_{s,c}$ and $\mathcal{C}_{c,s}$. Formally we have:

$$
\begin{aligned}
\mathcal{S}_{s,c} =& \{(d, p, d') : \mathcal{S}_s \mid \exists (e, p', e') : \mathcal{C}_c; (d_1, p_1, d_2), \ldots, (d_{n-1}, p_{n-1}, d_n) : \mathcal{S}_s \bullet \\
& \qquad d' = d_1 \wedge d_n = e'\} \\
\mathcal{C}_{c,s} =& \{(e, p, e') : \mathcal{C}_c \mid \exists (d, p', d') : \mathcal{S}_s \bullet e' = d'\} \\
\mathcal{T}_{s,c} =& \mathcal{S}_{s,c} \cup \mathcal{C}_{c,s}
\end{aligned}
$$

Thus we generate in the Phase Generate privacy threat graphs one Privacy Threat Graph for each privacy requirement, which is part of the Model updated with privacy requirements. All graph generations are performed automatically using OCL expressions. The details of the automatic graph generation are described in Sect. 4.5.

4.4 Analysis

For the Analyze privacy threat graphs we have to note that in the Privacy threat graphs we can distinguish two kinds of edges:

1. Edges (c, p, d) have the semantics that the counterstakeholder c may gain information from the domain d in problem diagram p.
2. All edges (d_1, p, d_2) with $d_1 \neq c$, i.e. the edges do not start from the counterstakeholder c, have the semantics that information is possibly transferred from domain d_1 to d_2 in problem diagram p.

In the analysis we have to take a closer look at all domains d for which a problem diagram p exists such that $(c, p, d) \in \mathcal{T}_{s,c}$. For each such domain d, we may have found a privacy threat against the stakeholder s from the counterstakeholder c. The threat graph $\mathcal{T}_{s,c}$ provides us two different kinds of information according to the two different kinds of edges in the graph as mentioned above.

1. We provide the information which requirements may allow the counterstakeholder c to gain information from the domain d about the stakeholder s. These requirements are contained in the problem diagrams p with $(c, p, d) \in \mathcal{T}_{s,c}$.

2. How the information about the stakeholder s got to the domain d is described through the path $(s, p_1, d_1), \ldots, (d_{n-1}, p_n, d)$ where the requirements in the problem diagrams p_i should explain which information is transferred to which domain.

Since our approach is an over-approximation of the system's actual information flow, there can be edges that actually do not represent an information flow in the system. These edges can be identified manually by studying the problem diagram and the requirement expressed in it. Edges that after the removal of the above-mentioned edges are no longer part of a path starting from the stakeholder, can be removed. To solve the threats in the refined threat graph $\mathcal{T}'_{s,c}$ there are again two possibilities, which also can be used in combination.

1. For an edge $(c, p, d) \in \mathcal{T}'_{s,c}$, the requirement in the problem diagram p can be modified or an additional requirement can be added, expressing that the counterstakeholder c is not able to gain personal information of the stakeholder s via the domain d.
2. For an edge $(d_1, p, d_2) \in \mathcal{T}'_{s,c}$ with $d_1 \neq c$, the requirement in the problem diagram p can be modified or an additional requirement can be added, expressing that no personal information of the stakeholder s is processed or stored by the domain d_2.

The outcome of our last step Analyze privacy threat graphs is the Model with additional or updated requirements created from the Model updated with privacy requirements and the Privacy threat graphs.

4.5 Technical Realization

For the generation of the graphs explained in Sect. 4.3 we developed the ProPAn tool[2]. This tool generates the four kinds of graphs, namely the global information flow graph, the stakeholder information flow graph, the counterstakeholder access graph, and the privacy threat graph, from a UML model. This model has to contain problem diagrams and has to be annotated with the stereotypes presented in Sects. 2 and 4.2. The graphs are generated automatically without user interaction for all privacy requirements, stakeholders, and counterstakeholders, respectively using OCL expressions.

The generation of the global information flow graph $\mathcal{G}$ is done by the OCL expression shown in Listing 1.2. This OCL expression first defines the sets of all domains doms (lines 1-3) and all problem diagrams pds (lines (4-6) of the model. For the generation of $\mathcal{G}$, the expression iterates the set of all problem diagrams pds (line 8). For each domain $p \in$ pds the sets pdoms (lines 9-12) and pcdoms (lines 13-17) are generated. The set pdoms consists of all domains that are part of the problem diagram p. The set pcdoms includes all constrained domains of the problem diagram p. Then the elements of pcdoms and pdoms are iterated (lines 18-19). For all $c \in pcdoms$ and $d \in pdoms$ with $d \neq c$ an edge (d, p, c) is added to the global information flow graph $\mathcal{G}$ (line 20).

For the generation of the global information flow graph $\mathcal{G}$, we iterate all problem diagrams and for each problem diagram, we iterate all domains that are part of it and all constrained domains in it. Hence, $\mathcal{G}$ contains at most $\#ProblemDiagram \cdot$

[2] http://www.uni-due.de/swe/propan.shtml

```
1  let doms: Set(Class) =
2    ProblemFrames::Domain.allInstances().base_Class
3  in
4  let pds: Set(Package) =
5    ProblemFrames::ProblemDiagram.allInstances().base_Package
6  in
7  let G: Set(Tuple(d1 : Class, pd : Package, d2 : Class)) =
8    pds→iterate(p; A: Set(Tuple(d1 : Class, pd : Package, d2 : Class)) = Set{} |
9      let pdoms: Set(Class) =
10       p.member→select(oclIsTypeOf(Association)).member.type→asSet()
11       →intersection(doms)→asSet()
12     in
13     let pcdoms: Set(Class) =
14       p.member→select(oclIsTypeOf(Dependency))
15       →select(getAppliedStereotypes().name→includes('constrains'))
16       .target→select(oclIsTypeOf(Class))→asSet()
17     in
18     A→union(pcdoms→iterate(c; B: Set(Tuple(d1 : Class, pd : Package, d2 : Class)) = Set{} |
19       B→union(pdoms→iterate(d; C: Set(Tuple(d1 : Class, pd : Package, d2 : Class)) = Set{} |
20         if d≠c then C→including(Tuple{d1=d, pd=p, d2=c}) else C endif)))))
21 in G
```

Listing 1.2. OCL Expression for Global Information Flow Graph

$\#Domain^2$ many edges. Thus, the asymptotic time complexity for the generation of all privacy threat graphs is in $\mathcal{O}(\#ProblemDiagram \cdot \#Domain^2)$, because for the generation of all other graphs, the global information flow graph is used.

For the generation the ProPAn tool uses the Eclipse Platform [7], the Acceleo model to text framework [8] and the graph layout tool GraphViz [4]. Each generator is realized as Acceleo template, which generates a dot file. An Acceleo template uses OCL to query a model and transforms the query results in a file format defined by the template. In our case it is the dot file format which can be read by Graphviz to generate different kinds of graphical representations. These templates are exposed as plug-ins to the Eclipse Front-End. So they directly integrate with modeling tools like Papyrus [3] and UML4PF [5].

5 Application on the Case Study

In this section, we will apply our method described in Sect. 4 on the NESSoS EHS case study introduced in Sect. 3.

5.1 Graph Generation

Figure 10 shows the global information flow graph, which the ProPAn tool generated from the problem diagrams of the 5 requirements mentioned in Sect. 3. For a better readability, we joined the edges starting from and ending at the same domain and annotate the resulting edge with the set of problem diagrams, the joined edges were annotated with.

To analyze if researchers may gain information about patients in the EHS, we added a privacy requirement, which is shown in Fig. 9 in Sect. 4.2. Figure 11 shows the threat graph for the stakeholder patient and the counterstakeholder researcher, which

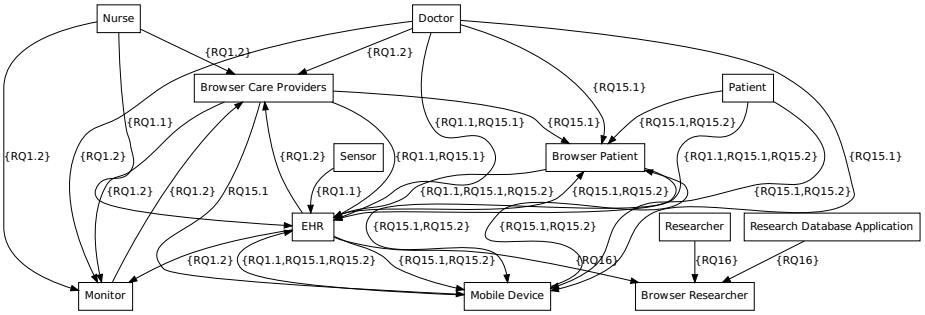


Fig. 10. Global Information Flow Graph

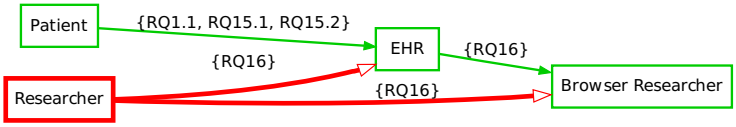


Fig. 11. Threat Graph for Stakeholder Patient and Counterstakeholder Researcher

the ProPAn tool generated from the global information flow graph and the privacy requirement. The red part with bold edges and white arrowheads depicts the counterstakeholder graph, and the green parts depict the stakeholder graph. To improve readability, we again aggregated all edges that start from the same domain and also end at the same to domain to one edge, annotated with the set of problem diagrams of the aggregated edges.

5.2 Analysis of the Threat Graph

As mentioned in Sect. 4.4, we have two possibilities to solve the privacy threats that the threat graph identifies. We can consider the edges starting from the counterstakeholder and restrict the information the counterstakeholder can access or we consider the other edges of the threat graph and restrict the information flow between the domains.

The threat graph $\mathcal{T}_{\mathsf{Patient},\mathsf{Researcher}}$ in Fig. 11 has only two edges from the counterstakeholder Researcher to the other domains. From these red, bold edges with white arrowheads, we can see that in the problem diagram for requirement RQ16 (see Fig. 6) the researcher may gain information from the lexical domain EHR and the connection domain Browser Researcher. One possibility to resolve this threat would be to modify requirement RQ16 in such a way that the medical data released to researchers has to be anonymized or pseudonymized.

The other possibility to resolve the privacy threat would be to update all requirements that lead from the Patient to the lexical domain EHR in such a way that they forbid to write personal information about the patient into it. Which problem diagrams have to be considered for this, is shown by the annotations of the green edges that are part of

a path from the Patient to the EHR. In Fig. 11, there is only one path from the Patient to the EHR, which is annotated with the requirements RQ1.1, RQ15.1, and RQ15.2. To solve the privacy threat we could update those requirements such that the medical data of the patient is stored in an anonymized or pseudonymized way in the electronic health record.

However, since the electronic health record of a patient has to contain personal information, we restrict the information the Researcher gets in requirement RQ16. We change requirement RQ16 to: "Release medical data pseudonymized to researchers."

6 Related Work

Deng et al. [6] present a threat analysis framework for privacy based upon the threat categories: linkability, identifiablitiy, non-repudiation, detectability, information disclosure, content unawareness, and policy/consent noncompliance. These threats are modeled for the elements of an information flow model, which has data flow, data store, processes and entities as components. Privacy threats are described for each of these components. Hence, privacy threat identification for an existing data flow model is simplified, because for each data flow element in a model only the threats shown in the tree need to be considered. In a last step, the authors relate the privacy threats to privacy enhancing technologies, which can be used to mitigate them. The work differs from our own, because the privacy threat identification has to be carried out manually.

The PriS method [17] elicits privacy requirements in the software design phase. Privacy requirements are modeled as organizational goals. Furthermore, privacy process patterns are used to identify system architectures, which support the privacy requirements. The PriS method starts with a conceptual model, which also considers enterprise goals, stakeholders, privacy goals, and processes. It is based upon a goal-oriented requirements engineering approach, while our work uses a problem-based approach as a foundation. The difference is that our work focuses on a description of the environment as a foundation for the privacy analysis, while the PriS method uses organizational goals as a starting point. In addition, the PriS method has to be carried out manually.

Hafiz [10] describes four privacy design patterns for the network level of software systems. These patterns solely focus on anonymity and unlinkability of senders and receivers of network messages from protocols, e.g., http. The patterns are specified in several categories. Among them are intent, motivation, context, problem and solution, as well as forces, design issues and consequences. This work focuses on privacy issues on the network layer and can complement our work in this area.

Asnar et al. [2] present a computer-aided approach to detect security threats based upon SI* models. The authors present patterns that can be used to identify specific areas in the models that present a security threat. The authors investigate access control permissions and search for roles in the models that have more permissions than they require to fulfill their goal. This analysis is carried out semi-automatically using graph patterns. The difference to our work is that we focus on privacy threat detection and we base our work upon the problem frames approach instead of SI*.

7 Conclusions

In this paper, we have presented the ProPAn method. The ProPAn is a problem-based approach for semi-automatic identification of privacy threats during the requirements analysis of software systems. The privacy threats are derived from potential access of counterstakeholders to personal information that are part of the system-to-be. We have extended the problem frames approach with the ProPAn tool[3], which consist a UML profile for privacy and a privacy threat graph generator.

The privacy threat graph helps to determine where personal information of a stakeholder may be processed and stored across the boarders of the problem diagrams, which correspond to subproblems of the overall development task. It also shows from which domains the counterstakeholder may gain information about the stakeholder. In our example the threat graph showed us that the requirement RQ16 contains a privacy threat violating the privacy requirement Preserve Anonymity.

Our graphs have formal semantics, which are strongly related to the problem frames approach. The generation of the threat graph for a privacy requirement, which is formulated with our introduced stereotype. The generation is performed automatically by the ProPAn tool from the problem diagrams, which represent the requirements the system-to-be has to fulfill. Our privacy threat identification is independent of the actual privacy goal, such as anonymity, unlinkability, unobservability, and pseudonymity, and gives guidance to detect possible privacy threats as early as possible in the software engineering process.

In summary, the ProPAn method has the following advantages:

- The privacy threat identification is re-usable for different projects.
- The privacy threat graph are generated automatically using our proposed tool chain.
- The identified privacy threats can be traced to a specific (sub-)problem of the system-to-be.
- The (sub-)problems can be enhanced with privacy requirements that constrain the functional requirements. Thus, we provide guidance where to apply privacy enhancing technologies.

Since our approach relies on the information flow inside the system-to-be, it can only detect those privacy threats that stem from an information flow starting from a stakeholders to a counterstakeholders, who both are part of system-to-be.

In the future, we plan to elaborate more on the later phases of software development. For example, we want to apply our approach to software components that were developed with the ProPAn method. We aim to identify privacy threats for existing architectures and propose solutions for these problems. The knowledge gathered during the usage of the approach might lead to the discovery of privacy threat patterns. Moreover, we want to extend the ProPAn method to support counterstakeholders that are not part of the system-to-be, but external attackers. We plan to provide extensions of the ProPAn method and tool that allows one to model the capabilities of these attackers and engineer a reasoning method for deciding if a privacy mechanism can protect a system against a certain attacker.

[3] `http://www.uni-due.de/swe/propan.shtml`

References

1. Alebrahim, A., Hatebur, D., Heisel, M.: A method to derive software architectures from quality requirements. In: Thu, T.D., Leung, K. (eds.) Proceedings of the 18th Asia-Pacific Software Engineering Conference (APSEC), pp. 322–330. IEEE Computer Society (2011)
2. Asnar, Y., Li, T., Massacci, F., Paci, F.: Computer aided threat identification. In: Proceedings of the 2011 IEEE 13th Conference on Commerce and Enterprise Computing, CEC 2011, pp. 145–152. IEEE Computer Society (2011)
3. Atos Origin: Papyrus UML Modelling Tool (February 2011), http://www.papyrusuml.org/
4. AT&T and Bell-Labs: Graphviz - Graph Visualization Software (June 2012), http://www.graphviz.org
5. Côté, I., Hatebur, D., Heisel, M., Schmidt, H.: UML4PF – a tool for problem-oriented requirements analysis. In: Proceedings of the International Conference on Requirements Engineering (RE), pp. 349–350. IEEE Computer Society (2011)
6. Deng, M., Wuyts, K., Scandariato, R., Preneel, B., Joosen, W.: A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requir. Eng. 16, 3–32 (2011)
7. Eclipse Foundation: Eclipse - An Open Development Platform (2011), http://www.eclipse.org/
8. Eclipse Foundation: Acceleo - transforming models into code (June 2012), http://www.eclipse.org/acceleo/
9. EU: Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Tech. rep., European Community(EU) (1995), http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML
10. Hafiz, M.: A collection of privacy design patterns. In: Proceedings of the 2006 Conference on Pattern Languages of Programs, PLoP 2006, pp. 7:1–7:13. ACM (2006)
11. Hansen, M., Schwartz, A., Cooper, A.: Privacy and Identity Management. IEEE Security & Privacy 6(2), 38–45 (2008)
12. Hatebur, D., Heisel, M.: A foundation for requirements analysis of dependable software. In: Buth, B., Rabe, G., Seyfarth, T. (eds.) SAFECOMP 2009. LNCS, vol. 5775, pp. 311–325. Springer, Heidelberg (2009)
13. Hatebur, D., Heisel, M.: A UML profile for requirements analysis of dependable software. In: Schoitsch, E. (ed.) SAFECOMP 2010. LNCS, vol. 6351, pp. 317–331. Springer, Heidelberg (2010)
14. ISO and IEC: Common Criteria for Information Technology Security Evaluation – Part 2 Security functional components. ISO/IEC 15408, International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) (2009)
15. Jackson, M.: Problem Frames. Analyzing and structuring software development problems. Addison-Wesley (2001)
16. Jackson, M., Zave, P.: Deriving specifications from requirements: an example. In: Proceedings 17th Int. Conf. on Software Engineering, Seattle, USA, pp. 15–24. ACM Press (1995)
17. Kalloniatis, C., Kavakli, E., Gritzalis, S.: Addressing privacy requirements in system design: the PriS method. Requir. Eng. 13, 241–255 (2008)
18. OECD: OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Tech. rep. Organisation for Economic Co-operation and Development (OECD) (1980), http://www.oecd.org/document/18/0,3746,en_2649_34255_1815186_1_1_1_1,00&&en-USS_01DBC.html
19. Westin, A.F.: Privacy and Freedom. Atheneum, New York (1967)

Conceptual Framework and Architecture for Privacy Audit

Ksenya Kveler[2], Kirsten Bock[1], Pietro Colombo[3], Tamar Domany[2], Elena Ferrari[3], and Alan Hartman[2]

[1] Unabhaengiges Landeszentrum fuer Datenschutz (ULD)
[2] IBM Israel - Science and Technology LTD
[3] University of Insubria, Department of Theoretical and Applied Science

Abstract. Many ICT applications involve the collection of personal information or information on the behaviour of customers, users, employees, citizens, or patients. The organisations that collect this data need to manage the privacy of these individuals. In many organisations there are insufficient data protection measures and a low level of trust among those whose data are concerned. It is often difficult and burdensome for organisations to prove privacy compliance and accountability especially in situations that cross national boundaries and involve a number of different legal systems governing privacy. In response to these obstacles, we describe instruments facilitating accountability, audit, and meaningful certification. These instruments are based on a set of fundamental **data protection goals** (DPG): availability, integrity, confidentiality, transparency, intervenability, and unlinkability. By using the data protection goals instead of focusing on fragmented national privacy regulations, a well defined set of privacy metrics can be identified recognising **privacy by design** requirements and widely accepted certification criteria. We also describe a novel conceptual framework and architecture for defining **comprehensive privacy compliance metrics** and providing **assessment tools for ICT applications and services** using as much automation as possible. The proposed metrics and tools will identify gaps, provide clear suggestions and will assist audit and certification to support informed decisions on the trustworthiness of ICT for citizens and businesses.

1 Introduction

Rapid developments in IT technology are constantly offering new IT products and services that involve personal data processing. An enormous amount of digital information is collected, stored, and shared all over the world. Alongside the benefits, new risks arise when privacy concerns are not properly addressed during the development process. The worldwide exchange of personal data, electronic surveillance possibilities, and the discriminatory use of personal information for actions such as profiling and identity theft impose advanced privacy concerns for individuals and significant economic and reputational risks for businesses.

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 17–40, 2014.

The fundamental right to the protection of personal data is recognized in Article 8 of the Charter of Fundamental Rights of the European Union and is set forth in the national data protection acts implementing the European Data Protection Directive 95/46/EC and the ePrivacy Directive 2002/58/EC. European national data protection acts recognize the rights of data subjects and impose obligations on data controllers, providing sanctions and remedies in cases of law infringements.

Privacy regulations apply to all the processes that relate to personal information. Personal information is any information relating to an identified or identifiable natural person (data subject) including secondary personal information such as log data. Public and private organisations are not always aware of the amount and existence of such personal identifiable data and therefore do not fully apply data protection regulations. Thus data protection non-compliance is a common problem in the EU. Data protection authorities who are responsible to safeguard the protection of the individual under the Data Protection Directive have neither the financial means, staff, nor powers to ensure full data protection compliance.

The processing of personal data is subject to rapid technical innovation. Law makers are therefore more and more refraining from regulating specific technical requirements for the processing of personal data. Instead principles of fair data processing and technical and organisational measures have been identified and put into legislation. Recently, regulators have introduced data protection goals (DPG) as a more comprehensive and adequate way to ensure data protection and privacy [DSK].

Compliance audits are one of the key mechanisms of the accountability principle and data protection regulations enforcement. Moreover, voluntary product audit privacy certifications are becoming more prevalent, providing competitive advantages and fostering user trust. Instead of sanctioning violations of privacy, promising market advantages offer positive incentives for implementing and observing privacy. For example, the EuroPriSe privacy seal [EuroPriSe] certifies that a product or service is compliant to regulations, based on an evaluation provided by privacy experts using a variety of time consuming legal and technical validation steps. Automatic tools are needed to assist auditors in assessing data protection compliance in an efficient and reliable way, improving the likelihood that the data protection goals defined by regulations are indeed met.

The difficulty in data protection is to produce comparable results in the absence of reliable privacy compliance indicators. This is partly due to the fact that data protection measures and compliance requirements are subject to legal decisions produced by the weighing of contrary principles such as data availability and data confidentiality. Thus, the validation and final result whether a specific data processing application is compliant with data protection law cannot be achieved automatically using information technologies alone; however, the evaluation process can be supported by privacy compliance indicators and an automatic toolset that allows for better and faster assessment of data protection compliance prerequisites and metrics. In doing so privacy compliance indicators support any data protection monitoring or assessment, such as Privacy Impact Assessments (PIA) [PIA] or an accountability program resulting in third party certification.

Our objective is to define a comprehensive set of privacy compliance metrics and create a set of assessment tools. Those tools will enable an audit of how an organisation performs based on those metrics, using as much automation as possible, and will provide clear suggestions for improvements. These metrics will be mapped to the protection goals, thus providing a means to assess which data protection principles are violated and why. The metrics will be defined and implemented with a set of privacy preserving techniques for their computation using a privacy by design approach.

Our privacy compliance assessment initially focuses on four main contexts: privacy policy compliance in general, and then specific compliance in the data storage, data sharing, and web sites operated by the organisation under audit. Other areas will be considered in future work.

The transformation of legal data protection requirements into technical metrics and the assessment of compliance poses a serious research challenge. New technologies are challenging lawmakers by introducing more complex systems and services which may be in conflict with the law. Current approaches in PIA and other assessment methodologies lack systematics and focus on the risks, based on the perspective of the organization either from a technology, an economic, or a legal point of view with each one of them demanding priority. They lack an explicit and systematic coverage of the protection of the interests of the data subjects. Data protection and data security operate from different perspectives and consider different risk sources: the attacker-model of data security aims to protect the operation of the organization primarily from persons who can pose special risks because they act as (former) employees, unfair or fraudulent citizens, as customers, or hackers. Data protection, by contrast, models organisations and their processing operations as potential attackers on the integrity and privacy of persons who are the data subjects in their roles of citizen, customer, client, patient, etc. Audit tools incorporating data protection goals go beyond data security assessment by operationalisation of data protection requirements which allow focusing on a common approach towards legal requirements, technical implementation and economic calculation without one of these domains dominating the other [Rost2012].

The paper is organised into three main sections where we discuss Data Protection Goals, Privacy Assessment, and the proposed Assessment Tool Architecture.

2 Data Protection Goals

The right to protection of personal data is established by Article 8 of the Charter and Article 16 TFEU and in Article 8 of the European Convention on Human Rights (ECHR) as well as by national laws. It imposes an obligation on private and public organisations to observe the right to privacy when processing personal data. The European Data Protection Directive only provides very general guidance on how data protection shall be implemented by technical and organisational measures. Due to the rapid developments in ICT, regulations on specific requirements are quickly outdated. The latest approach of privacy legislators [DSK] is to refrain from regulating specific technical security requirements; instead the regulators introduce Data Protection

Goals (DPG) as a more comprehensive and adequate way to ensure the protection of the individual [LDSG-SH]. The classification and applicability of protection goals has recently been elaborated in several articles and studies [AAL][GB2012][ZH2012]. The specific function and merit of "goals" is their ability to express a compulsory directive (normative ought), and their ability to address aspects of rules of operation, particularly of system applications. Having the same goals the different domains may be addressed coherently; experts may pursue the same direction and thus the same goal. The DPG provide a standardised approach to data protection investigations and audits [Rost 2012]. Data protection can be refined into specific fundamental protection goals: availability, integrity, confidentiality, transparency, intervenability, and unlinkability [RP2009]. The "classic" security goals of data security, availability, integrity, and confidentiality, focus primarily on guaranteeing the safe and secure maintenance of operation and infrastructure of an organisation. Data protection, by contrast, specifies these demands from the perspective of the data subjects (more precisely: citizens, customers, users, and patients) and augments this perspective with further demands derived from the basic rights of individuals. The specific demands can also be shaped into protection goals.

The following definitions are from Secttion 5 paragraph 1 of LDSG-SH. Availability is ensured if processes are timely available and can be used according to the rules. Integrity is ensured if data remain undamaged, complete, attributable, and up to date. Confidentiality requires that only authorized access is possible. Transparency means that the processing of personal data can be reproduced, verified and reviewed with reasonable. Unlinkability is ensured if personal data cannot or can only with unreasonably high efforts be collected, processed or used for another than its defined purpose. Intervenability requires a process to be designed in such a way that the data subject can exercise her rights effectively.

Availability, integrity, and confidentiality are classic, best practice IT-security protection goals since the 1990s.Data protection- goals also address transparency – as a prerequisite for the governance and regulation of technical-organisational processes– unlinkability – as an operationalisation of purpose bindingness/purpose separation – and intervenability – to operationalise data subject rights and the requirement on operators of systems to demonstrate that the data subjects have control over their information and are not dominated by the system. These goals comprehensively and explicitly address all relevant data protection aspects in a processing operation with respect to the data itself, the system and the procedures implemented [BM2012]. Using the same best practice methodology as in IT-security reduces translation errors between legal requirements and technical implementation [RB2012] and provides the methodology for privacy by design [RB2011].The data protection goals allow for the implementation of objective-specific protection measures which are technically and organisationally viable and controllable [Probst 2012].

With respect to governance the COBIT-framework [COBIT] offers a best practice to address regulation and controlling of processes. Key performance indicators and key risk indicators offered by this framework can be utilized as a regulative variable to implement and enforce data protection compliant processes in organisations. Some of the risk indicators have been specified for the RFID PIA Framework [PIA] and

were further put into more concrete form by the German Federal Office for Information Security (BSI) "Privacy Impact Assessment Guideline" [BSI]. The PIA Framework strives to address potential security and privacy risks and proposes measures to mitigate risks in the context of RFID. Nevertheless, this framework does not provide a systematic approach to specific privacy indicators (e.g., use of encryption to ensure confidentiality) or metrics (e.g., to determine the encryption level of a concept to secure confidentiality: in an organisation, how much information is sent encrypted?) to actually determine the scope of legal compliance. The PIA risk approach fails to identify those legal requirements which have not been implemented and whose non-implementation causes a potential threat. Privacy compliance indicators developed from DPG will cover not only risk based indicators but also performance and requirement based indicators. The privacy indicators focus on the performance of an IT product or service as compared to user and provider requirements and values, as defined in the legal requirements.

Data protection goals guide the systematic assessment of all privacy aspects of data processing for emerging technologies as well as for audits of running systems. Any audit will start with a description of its target of evaluation (ToE). This requires a comprehensive analysis of the processing operations, the players involved, and the data that is processed. Only an accurate analysis of who collects and processes personal data at what moment allows determining the applicable regulations and requirements.

The protection goals have been tested first on an implementation of ambient assisted living [AAL] technologies which aims to assist elderly or challenged individuals to live a more independent life. The use of AAL-technology provides a good example to illustrate a data protection audit scenario. To determine all relevant aspects a cube-model is used. The DPG CUBE (see [DSK], below) allows integrating and considering all parties involved and determines the work space which is the target of privacy evaluation (ToE).

In trying to determine the ToE one has to consider the data, the IT-systems, and the processes used. Relevant data groups in an AAL-scenario are intervention data (e.g., remote-medication by setting an injection or securing doors), vital signs (e.g., blood sugar, weight, temperature), behavioural data (sleeping-, eating-, working-, and resting times), technical infrastructure data, measurement and environmental data (e.g., temperature, lighting, humidity, sound level), triggered data (e.g., alarm contact, on-off switch). With these data in mind the IT-technology producing the data comes into focus as well as the processing, transmitting, archiving, and deletion.

In any use-case we see three process domains that are important in the processing of data: 1. Processes involving the data subject e.g. patient, 2. Processes involving the organisation or service provider (e.g., doctors office, hospital, public administration, insurance company), and 3. Infrastructural processes involving service providers of the organisation under 2. (e.g., data centres, access and content providers, but also controlling authorities and research institutes). Each process belongs to a process-owner who needs to be responsible and accountable for its design. This is where most often responsibility gaps are detected.

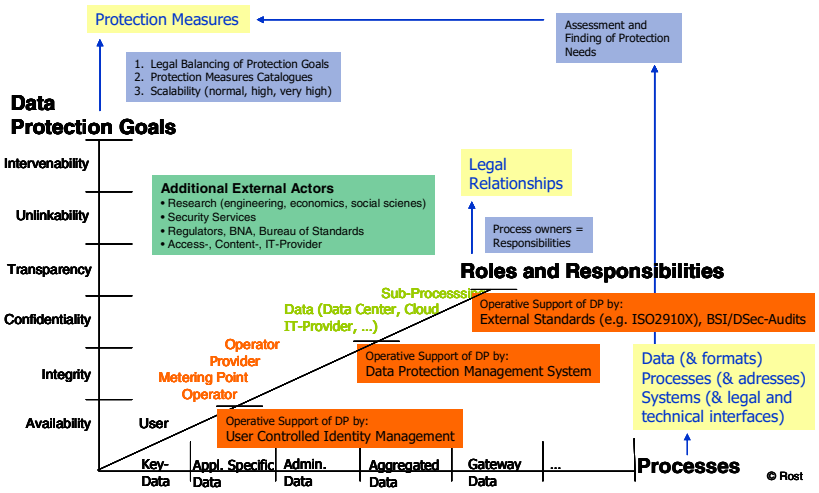


Fig. 1. DPG CUBE - differentiated risk assessment [Rost2011]

For example, if a person suffers from an early stage of Alzheimer and keeps forgetting to turn off the stove, close and lock the door at night or receives remote medication, AAL-technology can assist this person to control electrical items or the door by the use of sensor-systems which monitor status. Monitoring facilitates remote caretaking and allows the person to live at home and postpone admission to an institution. However, monitoring may entail a harmful degree of surveillance and deprivation of freedom if data protection is not observed. Lack of acceptance and trust may deter people from using such assistance.

The process-domains addressed in our AAL-example distinguish between the person concerned who needs to understand the system and must be able to control it (e.g., set-off an alarm); The process-owner, e.g. a home care organisation which is responsible for the functioning of the care system and its control mechanisms and interventions (e.g., what is to be done if an alarm is set off?); and on the infrastructure domain level a data centre which is processing the data collected at the home. By focusing on processes, the organisational structures, legal relationships and their justification, and also the responsibilities come into focus. By applying the DPG the extent of processing and access to data becomes visible and facilitates finding the appropriate legal basis or identifying an illegal process.

The three descriptive dimensions of the cube in our example are: 1) the processed data: e.g. door open – door closed, time stamp; 2) the basic processes and their owners that play a role in data processing: the person concerned opening and (not) closing the door, the care service monitoring the activities and offering the intervention service, the insurance paying for the service and the data centre hosting the IT; and 3) the protection goals to be applied to the complete use-case operationalising data protection norms [BM2012].

In our example the home would be equipped with technology to monitor the door status (closed/open) and send out an alarm in case the door stays open longer than

usual or for more than a specified time. With regard to the DPG we consider what kind of data is accessible/sent (transparency) to whom (integrity, confidentiality) for what reason (unlinkability) and how the person concerned and the organisation can or cannot intervene (intervenability). The care taking organisation needs to decide whether to send to the control room:

1. the exact time when the door was opened and closed (highly intrusive because habits and behaviour can be monitored),
2. the amount of time the door is open without a time stamp (less intrusive because no tracking of behaviour is logged), or
3. only the alarm-signal from a box taking the time the door is open at the home of the patient (privacy-friendly because no monitoring of behaviour takes place).

The care taking organisation also needs to decide whether the same access rights should be granted to the insurance company or whether it is sufficient to report the number of alarms or no data at all; and also how the IT-system integrity and confidentiality is ensured and whether redundant system architecture is provided. Finally, how to deactivate the system in case, e.g. during repair work, the craftsmen need an open door so that an alarm should not be set off.

The cube provides a model which enables us to describe the three dimensions relevant for data protection compliance analysis. The dimensions allow different view- and entry points. The advantage of the model consists of its ability to identify process responsibilities crucial for determining a ToE. All too often, only data security is addressed when IT-technology is assessed. The three dimensions of the cube ensure and enable all relevant stakeholders in an IT application to identify data types, processes and systems relevant for an assessment of the data protection performance of an IT application. The positioning of DPGs in a differentiated risk assessment for technical and organisational measures is illustrated above in [DSK].

The cube-model approach clarifies that data is always produced by a specific process which is carried out by specific application of IT-technology. When it comes to compliance every process must be covered by a legal basis which governs the activities of the person or technology involved. This is often overlooked when designing and evaluating ICT. The protection goals address compliance requirements in a comprehensive yet abstract manner. A catalogue of technical and organisational protection measures is deployed to implement the protection goals, as in the German Federal Office for Information Security [BSI] baseline protection [Probst2012].

On the dimensions of data and processes, we analyse which functional objectives are to be achieved by the process, what kind of data is required, which technical measures should have been chosen, and who is responsible in which role. Based on the processes, we can distinguish functions and determine the purposes that will generate the necessary data. In an assessment of a technical system, demonstrating what kind of data can be generated is of great importance. The data must be concretely defined and categorised in a manner that is relevant to its content.

The main focus of the third dimension of the DPG cube is the protection goals—the regulation and controllability of organisational processes. Here, process-organisational paradigms such as [ITIL, CoBIT], or processes based on ISO-oriented quality management are fairly well known. DPGs are expected to augment these paradigms. The approach, framework and toolset described in this paper will help determine the status of a system and match it to the target state, determined in the DPG-CUBE model. This procedure will support the data protection management of an organisation in continually monitoring data protection compliance.

In summary, contemplating data and processes is necessary to determine the purpose and necessity of a data processing. Contemplating data and protection goals leads to the analysis and determination of the protection demand of the data at hand, and governs the choice of technical and organisational protection measures. Contemplating processes and protection goals visualises processes and their regulation in the data protection management of the organisations involved. The generic DPG CUBE allows determining the protection demands of data and IT systems, measured processes, legal relationships and responsibilities, as well as the legally-weighted protection goals and protection measures. By addressing these relationships in a systematic manner, we address privacy compliance in a holistic way.

3 Privacy Assessment

The first step towards the development of privacy metrics and the associated assessment tools is the availability of a conceptual model for privacy onto which the assessed system can be mapped. Roughly speaking, it should define how the privacy-aware system is ideally supposed to be implemented. Privacy is a complex property characterised by numerous structural features, such as processing actions, data, purposes, obligations, users, authorisations to perform processing actions, and so on. All these concepts have to be properly formalised and composed to form the conceptual model for the privacy domain.

The conceptual model should be built around the principles of the DPG CUBE, and thus it has to support data specification, organisation and aggregation, as well as role-based data manipulation processes, protection goals and privacy policies, data protection mechanisms and adversary models. Although some proposals exist for languages to specify privacy policies, such as XACML [XACML], our analysis of the literature revealed the absence of a conceptual model that comprehensively considers all the concepts needed. Therefore, we first introduce a proper conceptual model. We approached this task by considering a subset of core privacy elements, originally formalised in [BL08], centred on the concept of purpose and related purpose-based access control policies. The preliminary version of the conceptual model, which is represented in the UML Class diagram in Figure 2, is discussed in [CF2012].

A key element of the model in Figure 2 is the concept of Purpose, which specifies the reasons for data collection and use. The other main components of the model are explained through our running example. Suppose that the AAL system manages data for *assisting*, *marketing* and *analysis* purposes. This can be modelled by *PS1,* an

instance of element PurposeSet, which groups these purposes. At any point in time the system administrator may decide to add a new Purpose to a PurposeSet, to remove or modify an existing Purpose. Therefore, an instance of PurposeSet can change dynamically over time. Data owners accessing the AAL system grant consent to use their data for specific purposes and prohibit their processing for other purposes. This is reflected in our conceptual model through the IntendedPurpose element which models a collection of allowed (intended) purposes (*aip*) and prohibited (intended) purposes (*pip*), which are bound to data (for simplicity, in Figure 2 we assume that data are organised according to the relational model). The purposes collected by an IntendedPurpose element must belong to the same PurposeSet. In our example, suppose that Bob, who suffers from Alzheimer, granted consent to process the state of the door of his apartment, which is modelled by means of an instance of element Data called *doorState,* for *assisting* purposes only. Bob also specified that his data cannot be processed for *marketing* purposes. These privacy requirements can be modelled through an IntendedPurpose *IP1*, including *assisting* in the *aip* component and *marketing* in the *pip* component, respectively, which is then assigned to *doorState.* In contrast, the element AccessPurpose collects the access purposes that are assigned to ProcessingActions that access and manipulate data. For instance, in our example, suppose we have *AP1*, an instance of AccessPurpose including purpose *assisting*, and that *AP1* is assigned to the ProcessingActions *monitoring, openDoor*, *closeDoor*, and *sendingAlarm*. A required (but not sufficient) condition to allow a ProcessingAction to process a set of data is that the purposes grouped by the IntendedPurposes assigned to the data and those collected by the AccessPurpose associated with the ProcessingAction belong to the same PurposeSet and are compliant.

Let us suppose that at a given point in time the system administrator introduces the new ProcessingAction *homeMonitoring*, which checks the state of windows and doors of the apartment where a patient lives for *security* reasons. Accordingly, the administrator 1) adds Purpose *security* to PurposeSet *PS1*, 2) introduces AccessPurpose *AP2* which includes *security* in the declared purposes, and 3) assigns *AP2* to *homeMonitoring*. As a consequence, Bob, who wants to benefit from the homeMonitoring service, 1) includes *security* in the *aip* set of *IP1* assigned to *doorState*, 2) introduces *IP2*, which is an instance of IntendedPurpose whose *aip* set consists of the purpose *security*, and 3) assigns *IP2* to *windowState*. The required condition for the execution of *homeMonitoring* on *doorState* and *windowState* is satisfied. On the other hand, *monitoring, openDoor*, *closeDoor*, and *sendingAlarm* cannot process *windowState* since the access purposes assigned to the ProcessingActions and the IntendedPurposes assigned to the data are not compliant (the *aip* set of *IP2* assigned to *windowState* does not include the Purpose *assisting*).Besides the straightforward concepts of User, the model in Figure 2 introduces the concepts of Role and ConditionalRole as a way to fine tune the administration of access rights. Role is composed of a set of attributes that characterise the role properties. For instance, a Role *domiciliaryAssistant* may be specified to model employees of the AAL service provider characterised by the *city* where they work, and the *zone* of the city. The attributes of a Role are initialized when the Role is assigned to a User. For instance, if the Role *domiciliaryAssistant* is assigned to user Mary, *zone* is set to 'Rosemont', whereas if it is assigned to Alice,

zone is set to 'Mont Royal'. To allow a more fine-grained management of access rights, element ConditionalRole extends Role with a condition that constrains when the Role can be assigned to a User. Such a constraint is a Boolean predicate defined in terms of the Role's attributes. For instance, the ConditionalRole *domiciliaryAssistantCR* extends *domiciliaryAssistant* with the constraint *"zone= 'Mont Royal'"*. This allows the assignment of the role to Alice, and prevents the assignment to Mary, whose zone of competence is Rosemont.

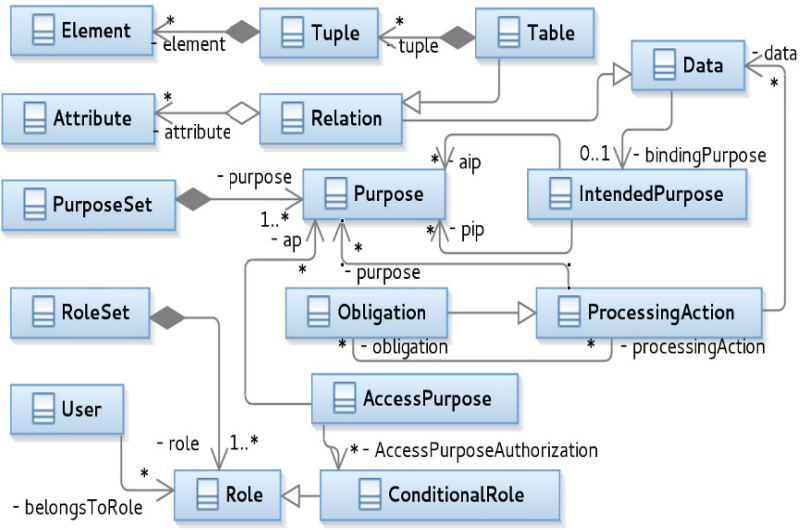


Fig. 2. A preliminary version of the privacy conceptual model

The conceptual model supports the specification of various types of privacy requirements that will then be used to drive the assessment phase. Privacy requirements are specified by means of a set of predicates that constrain the properties of the model elements. Such constraints capture all the properties that must be assessed for the system under analysis. For instance, an example of a privacy requirement is the privacy policy authorisation, which authorises access to data for actions associated with access purposes that comply with the intended purposes for which data owners granted their consent. This implies that, when an AccessPurposes AP is authorised for a ConditionalRole CR that extends a Role R, users that are authorised to play R and satisfy the constraint of the conditional role are allowed to execute processing actions with AccessPurpose AP on data D, only if AP complies with the IntendedPurpose IP specified for D. For instance, we can authorise domiciliaryAssistantCR to execute processing actions with AccessPurpose AP1. The authorisation is implicitly granted to all users who are authorised to play the Role domiciliaryAssistant and satisfy the domiciliaryAssistantCR constraint. Based on previous assumptions, in our scenario Alice is the only User with those characteristics. Let us consider the effect of the authorisation for the processing of Bob's data doorState. AP1 includes assisting as

access purpose and IP1, assigned to doorState, includes assisting in the set of allowed purposes. Therefore, AP1 complies with IP1. As such, Alice is allowed to execute the ProcessingActions monitoring, openDoor, closeDoor and sendingAlarm (associated with AP1) on Bob's doorState.

The conceptual model also supports the specification of data minimisation requirements, i.e., requirements that specify the largest set of data that can be collected and processed by the system under analysis. The elements *Relation* and *Attribute* (see Figure 2) are used to define the scheme of the data that can be collected. The minimisation requirement specifies that the tables schemes of the system under analysis must not have more attributes than those specified in the corresponding Relation element. The specification is performed by setting the Boolean attribute *minimisation* of element *Relation* to true. A similar mechanism is used to require the minimisation of the processed data. More specifically, the element *ProcessingAction* allows the specification of the *Relations* and the *Attributes* that can be involved in the execution of the *ProcessingAction*.

Requirements belonging to different categories (e.g., privacy policies and data minimisation) are integrated into a unified instance of the conceptual model (that is, as constraints on the corresponding entities). According to the DPG cube principles, multiple views can then be defined on the expressed requirements, allowing analysts to look at the requirements from the perspective of different data protection goals.

After specifying the requirements, the second step for the definition of the assessment mechanisms is the mapping of the components of the system under analysis to elements of the privacy conceptual model. This practice requires the analysis of the access control component, the database structure, and the configuration options of the assessed system.

Metrics are defined to evaluate the compliance of the mapped system to the reference requirements and, in case of non-compliance, to determine the non-compliant components (e.g., obligation support, role and purpose management). The output of such metrics evaluation will then be used by the assessment toolset to provide a set of recommendations on how to improve the system under assessment. A privacy metric expresses the similarity measure of the expected system state with the actual state of the system. Every policy involves a distinct set of conceptual model elements; therefore, a metric is required for every significant combination of conceptual elements.

For instance, suppose the AAL system supports Role-based access control. In this case, the mapping of AAL roles to conceptual roles is straightforward. In contrast, if roles are not supported, a possible countermeasure is to introduce a guest role with basic authorisations, and to assign it to all the users.

As another example, suppose that *doorState* is a column of the table *Patient* that collects patient's data. Suppose that the access to *doorState* is performed by the SQL query *extractDoorState*. Based on the conceptual model, an access purpose must be assigned to each processing action, whereas allowed and prohibited intended purposes must be assigned to data. If the AAL system does not record this information for *doorState* and/or *extractDoorState*, a warning message is returned by the assessment toolset along with proper recommendations to improve the system (e.g., information on the missing purposes).

The requirements constrain how the mapped elements must be related and configured. For instance, constraints can require

1. every action to be associated with an access purpose,
2. every data with an intended purpose, and
3. every user with a role.

Each constraint will have a weight that specifies its relevance. For instance, suppose that constraints 1 and 2 have weight 5, since the purpose is the key concept of the system, while the weight of constraint 3 is 1, since if no role is assigned to a user, he/she cannot be authorised to perform any action. These constraints can be used to achieve a compliance measure of the mapped model with the conceptual model. The number of satisfied constraints along with their weights will provide a compliance measure.

Furthermore, several dimensions of the mapped model can be measured and used to analyse the effectiveness of privacy protection. For instance, an analyst may be interested in counting and deriving:

1. all access purposes that comply with an intended purpose (e.g., in our example, considering *IP1* the resulting measure is *<1,{AP1}>*),
2. all roles that are authorised for an access purpose (e.g., considering *AP1*, we derive *<1,{ domiciliaryAssistant }>*),
3. all users that belong to a conditional role (e.g., for *domiciliaryAssistantCR*, we get *<1,{Alice}>*),
4. all access purposes that are granted to a user based on the conditional role the user belongs to (e.g., in case of Mary we derive *<0, Ø >*).

Dedicated metrics will also be defined to analyse the compliance of the system with data minimisation requirements. For instance, suppose the mapped elements include the Table Patient, which collects personal and sensitive data of all the patients, and the Relation PatientRM, which represents the scheme of Patient. Let us suppose that the AAL requirements specify the Relation PatientRR, whose minimisation attribute is set to true, and the ProcessingAction extractDoorStateR, which requires the processing of Attribute doorState of PatientRR under a minimisation constraint. Since the attribute minimisation of PatientRR is set to true, the relation PatientRM, which is derived from table Patient, can include only Attributes corresponding to those of PatientRR. A metric can check this constraint and count the number of attributes of PatientRM that are not included in PatientRR. Similarly, the SQL query extractDoorState is traced back to the processing action extractDoorStateR for which a data minimisation requirement is defined. Therefore, it is required to check that the set of data fields of extractDoorState is a subset of extractDoorStateR, and to count the number of possibly exceeding fields.

Even in the case of 100% compliance, further measures may be needed, because having all the necessary components for enforcing privacy-preserving access control, does not necessarily mean that the current access control configuration correctly enforces the desired privacy requirements.

Our conceptual model includes key concepts that are required for specifying general privacy requirements and supporting the assessment of existing systems with respect to these requirements. Moreover, the proposed metrics provide a quantifiable measure of different privacy aspects of the system under analysis. However, specific application domains may require additional conceptual elements that are not included in the current version of the conceptual model. Therefore, we cannot argue that the proposed conceptual model is complete. A direct parall can be traced with software testing. Testing cannot prove that the developed system satisfies the specification, but helps developers to increase the quality of the developed systems.

We support the assessment of implemented software systems both at run time and post execution. The metrics that can be evaluated at runtime are those associated with privacy policies expressed in terms of the current system state and/or previous states, whereas post-execution metrics involve events and states that refer to current, past and future points in times.

As far as the run-time assessment is concerned, metrics computation requires the analysis of the system behaviour with respect to the privacy requirements that express invariant properties of the system, such as the policy Authorisation introduced above. The derived measures are used to constrain the system execution by allowing, forbidding, or obligating the execution of operations associated with the involved events.

For instance, in case of invocation of an SQL query (e.g., openDoor), the run-time assessment metrics verify its authorisations. In case of non-compliance, the metric assessor will determine the non-compliant components. As an example, suppose that Mary executed action openDoor accessing Bob's data. The system will inform the proper controller that Mary does not satisfy the ConditionalRole constraint and therefore she is not authorised to execute the processing action openDoor. The output of the query evaluation will also provide a set of recommendations on how to improve/correct the configuration of the access control mechanism in place.

In contrast, privacy policies that can be checked by post-execution metrics are those that refer to current, past and future points in times. This allows for specifying complex trace execution constraints that involve retention conditions and obligations that refer to a future point in time.

For instance, suppose that the AAL system is used to monitor diabetic patients. Patients are required to periodically measure their glycaemia with a device that automatically informs the AAL system of the measured value. The entire measuring and notification process is modelled by means of the ProcessingAction monitoringNotification. The policy GlycaemiaAlarm states 'if the glycaemia exceeds a certain value, a physician must be contacted by phone within 5 minutes and informed of the measured value of the involved patient'. The called user should have activated the ConditionalRole doctor, and this role should be among those the patients gave the consent to access their personal data for assisting purposes. In this case, referring to the components of our conceptual model, glycaemia data are sensitive Patient data collected by the system for the Purpose of assisting the patient, doctor is a Role, and the phone-call is an Obligation associated with monitoringNotification. GlycaemiaAlarm is a policy that can be evaluated only by the post execution assessment, since checking the obligation requires delaying the analysis till after the query execution. Therefore, in our

example a phone call trace must be included in the AAL system log that should be checked five minutes after the invocation of monitoringNotification.

4 Assessment Tool Architecture

This section discusses the proposed architecture for the assessment toolset that will support the privacy audit process. The toolset will facilitate privacy metrics computation, identify gaps and provide compliance improvement recommendations. The intended users of the toolset are experienced external privacy auditors and internal privacy reviewers, so we assume a certain level of expertise and maturity. We also do not intend the tools to be totally automated, they are intended to make the work of these specialists more efficient, but not to replace them.

The assessment tool architecture which we propose is driven by two guiding principles, transparency and extensibility. It is designed to provide easy plugging of privacy compliance metric assessor components, each providing user interface (UI), analysis, and reporting capabilities for the particular technical metrics. This approach enables a gradual delivery process, starting with a limited set of assessor plugins, while targeting additional assessment techniques, areas, and privacy goals at a later stage. The modular approach also allows customized toolset packaging depending on particular target customer needs and audit type. The transparency principle aims to show the users how assessment decisions have been made, and what evidence has been collected during the analysis. The transparency is enforced by a required interface for all assessor plugins. This interface assures that each assessor provides evidence of the compliance or non-compliance and also advises for the improvement of the metric performance of the assessed artefact.

A conceptual architectural diagram of assessment tooling is shown in Figure 3. It comprises three main modules, the Audit Engine, the set of Metric Assessor Plugins, and the Administration module.

The Audit Engine is the core of the system, responsible for audit planning, privacy compliance analysis execution and report generation. The assessment process begins with the Planning component, which collects all the necessary information needed to understand the scope, plan and execute the analysis. Users might be asked to provide target privacy requirements for the system under assessment, in terms of the conceptual model described above, specify the desired assessment categories and supply any other metric-specific inputs. The Execution component performs the actual compliance analysis, according to the selections made and to information collected during the planning stage. Depending on the particular assessment logic and needs, if complete automation cannot be achieved, execution might be interrupted to collect additional user inputs. The Reports and Analysis component generates and presents detailed reports for the completed audit together with recommendations for improvements. In particular, Evidence Reports provide a record of all non-compliance evidence found. The Advisor generates recommendations for compliance improvements based on analysis results. Recommendations might be derived from a particular set of metrics that has been executed together with the higher level conclusions drawn from the

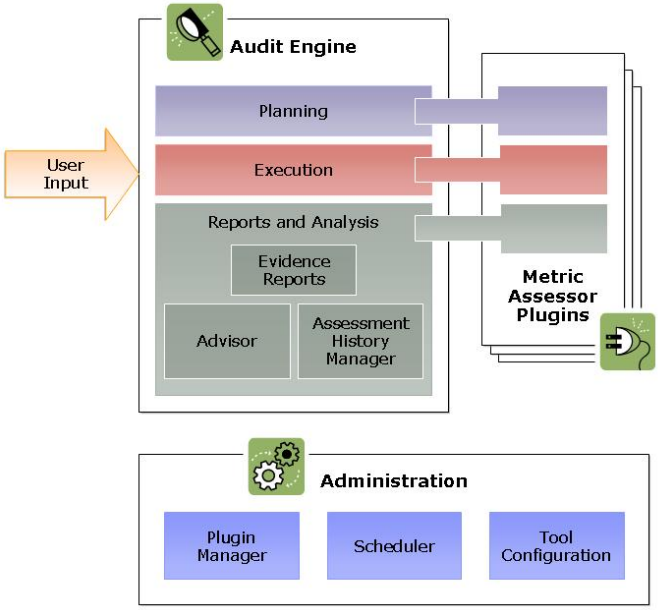


Fig. 3. Conceptual Privacy Audit tooling architecture

accumulative results of multiple metrics. The Assessment History Manager tool provides capabilities for viewing and for the analysis of previously executed audits. This enables tracking of compliance progress and improvement history.

A simplified sample audit report is shown in Figure 4. In reality, metrics will be more complicated and the resulting report will provide more details and capabilities. We plan to present audit results in a display compatible with the DPG CUBE model, providing different views on privacy compliance, from the perspectives of data protection goals, roles, processes and data. This visualisation capability is an item for future research.

The main audit report will show a general summary of the compliance analysis with grades and brief details for each metric that has been assessed. It will also allow rerunning of certain assessments to re-evaluate the metric after fixes have been applied; enable comparison to previous assessment results and drilling down into more detailed reports. The detailed reports will include more information about the data protection goals the particular metric is linked to and the concrete analysis steps that have been performed. They will also present the collected evidence of any non-compliance found and recommendations, such as guidelines on concrete measures to implement for compliance improvement.

The Metric Assessor Plugins are a set of pluggable components, each encapsulating everything that is needed in order to plan for, execute and report about a particular technical metric. Each metrics assessor contains:

1. its contribution to the planning UI,
2. its assessment execution code together with related user input UI if needed, and
3. its metrics-specific report generation capabilities.

The latter will include support for assembling and presenting metrics-specific assessment details, evidence and recommendations. Therefore, as shown on the conceptual architecture diagram in Figure 3, each metrics assessor possibly contributes to each of the three Audit Engine main modules. Each metrics assessor will also include specific information to facilitate its later use within an audit report, such as the assessment category it relates to (e.g., data store assessment) and the particular protection goals it is linked to.

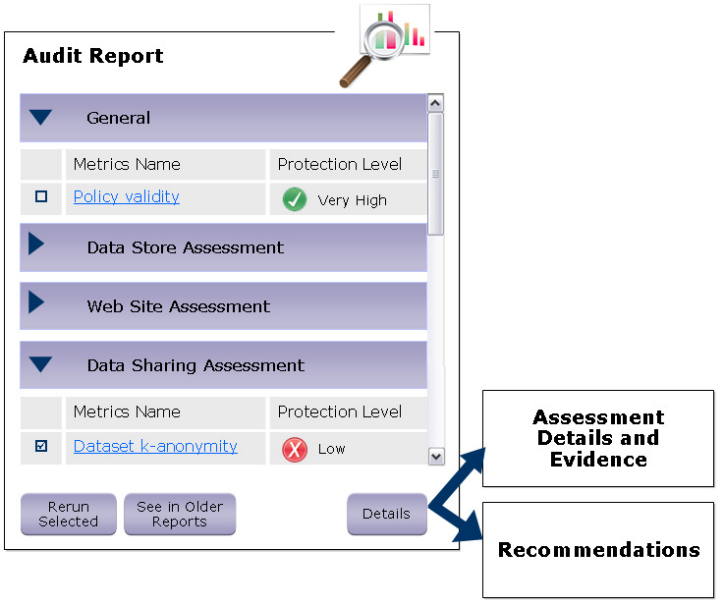


Fig. 4. Sample audit report

For example, one of the possible metrics can deal with overall validity, consistency and usability of the target privacy policy, specified in terms of the conceptual model described in the previous section. In other words, this metric should validate that the target privacy requirements make sense and would lead to a usable, conflict-free system. Following the AAL scenario above, the related assessor plugin can, for example, perform quantitative analysis and warn if there are too many users with a ConditionalRole doctor and too few with the ConditionalRole domiciliaryAssistantCR. It can also detect that according to the policy no AccessPurpose will be granted to some user based on his ConditionalRole (e.g., the case of Mary and the privacy policy Authorisation) or that there are no roles authorised for a certain AccessPurpose.

As described in the previous section, the system under analysis is mapped to the elements of the privacy conceptual model, by analysing the access control component, the database structure, and the configuration options of the assessed system. Thus another possible metric assessor plugin could examine the access policies that are actually used by the system (e.g., written using XACML) and verify their consistency with the desired target conceptual privacy policies and model. Moreover, yet another plugin could verify that the elements of the mapped system are related among them and configured according to the conceptual model constraints, for example, that every user is associated with a role.

Another metric assessor plugin could provide anomaly detection capabilities for identifying unauthorised or non-compliant database access by potential hackers, privileged insiders or other end-users. Patterns that do not conform to an established normal behaviour, and are thus considered suspicious, can be extracted by examining logs (post-execution assessment) and SQL queries intercepted at runtime. For example, in the AAL scenario, we could detect that while most of the time Alice triggers *extractDoorState* SQL query for accessing the database table *Patient*, which holds personal and sensitive data of patients, she also occasionally sends *extractPhoneNumber* query without being restricted to do so by the system. It might indicate, for example, that Alice is using patient's data (including the data of Bob) for marketing purposes without having an appropriate consent for that. This plugin would need to contribute to the planning UI, by providing plugin-specific screens where log location and database access parameters can be specified by the assessment tool users.

In our example, the AAL system manages data not only for *assisting*, but also for the purpose of *analysis*. In this case, it is necessary to assess anonymity of the datasets being shared with external parties (e.g., for statistical analysis). The Privacy Audit tool could provide a plugin that performs such an assessment based on anonymity metrics for privacy-preserving microdata release, for example *k*-anonymity [Sweeney2002]. It will contribute a dataset upload screen to the planning UI and its assessment result will appear under the Data Sharing Assessment category, as shown in the sample report in Figure 4. The plugin will respond with the result protection level ("Low" in Figure 4) according to the value of *k* for which k-anonymity is guaranteed and the particular value of *k* will be shown within the plugin-specific assessment details view.

A plugin for data minimisation assessment could support cases like the AAL door status and alert sending scenario described in the previous sections. The plugin will check that the system collects only the minimal amount of information needed, like alarm-signal events, but not the exact times the door was opened or closed. The assessment can be made by inspecting database schema and SQL queries for the presence of legitimate data elements only, according to the privacy requirements defined in the conceptual model. For example, let's say that privacy requirements specify that the attribute *alarmTimestamp* is part of the Relation *PatientRR,* whose *minimisation* attribute is set to true. The plugin will then analyse the schema of the database table Patient in the system under assessment, detect and warn about any fields beyond the above-mentioned permitted minimal set of attributes, such as *doorOpenTimestamp* or *doorClosedTimestamp*.

Yet another metric assessor plugin can deal with detection of obligation events and give an assessment of their compliance with the privacy requirements. For example, the plugin will analyse AAL log files, look for and correlate between records related to glycemia alarm events and events of establishing phone calls. The plugin will then verify that the user being contacted by phone has a ConditionalRole *doctor* and that this role is among the roles to which this particular patient gave the consent for accessing his *bloodSugarMeasurement* data.

The last module of the assessment tooling, the **Administration** module, provides management and configuration capabilities. For example, the Plugin Manager component will be responsible for the management of the Metrics Assessor Plugins repository, including viewing, editing and capabilities of adding new assessors or removing others. The Scheduler component allows the scheduling of automatic runs of pre-configured audits. The Tool Configuration component will enable adjustments of any other tool configuration, e.g., user interface options, general reporting options and any other settings.

We do not address the implementation details of any of the plugins at this stage of our work, Any discussion of algorithms for the privacy preserving computation of the metrics is beyond the scope of this conceptual design.

5 Related Work

In this section we describe relevant related work concerning privacy metrics and assessment tools.

5.1 Privacy Assessment

The systematic development of security and privacy assessment techniques is recognized today as a paramount requirement to assess the quality of any system with respect to its security and privacy guarantees [Jaquith2007]. However, most of the efforts developed so far focus on security, rather than on privacy [SA2009]. Savola [Savola2006] provides some high-level guidelines for the development of a framework for security evaluation based on security behaviour modelling and security evidence collection. The use of an ontology-based approach in support of run-time security monitoring is presented in [EOS09], and [HSHJ08] presents a security metrics framework, in which security metrics are associated with security patterns as a way to facilitate the interpretation of measurements. Information assurance metrics are described in [SPMNLH04], in which a review of existing metrics is performed, along with the proposal for a new taxonomy for information assurance metrics. A logic-based approach for reasoning about system security properties has been presented in [DFGK09] and applied to trusted computing.

Research on privacy assessment, however, is still in early stages, due mainly to the fact that despite recent advances in the field, privacy is still not a clearly defined concept. A discussion of measurements of compliance with security and privacy regulations and standards is presented in [Herrmann07]. Additionally, the preliminary study

reported in [Savola2010] presented a high level risk-driven methodology for privacy metrics development.

Much of the research efforts in the field have been instead devoted to anonymity metrics for privacy-preserving microdata releasing. Examples of such metrics are k-anonymity [Sweeney2002], l-diversity [MGKV06], t-closeness [LTV07], and differential privacy [Dwork2008]. These metrics capture different aspects of the disclosure risk, imposing some requirements on the association of an individual with the released sensitive private attributes, by making different assumptions on the attacker's background knowledge. Other works [Bezzi2010, RFD09] attempt to define an aggregated anonymity metric, based on information theory.

The need for a formal approach to privacy preservation was recognized by [Datta2011]. In this work, a logic-based model was defined with the aim of facilitating privacy polices specifications, and enforcement and compliance analysis. That model has been complemented with algorithms to check audit logs for compliance with privacy policies. It was also applied to several US privacy laws and resulted in the first complete logical specification and audit of all disclosure-related clauses of the HIPAA Privacy Rule [DGLKD10].

5.2 Existing Assessment Tools

Today privacy compliance analysis is still mainly done by specially trained experts, using reference documents, templates, forms and guidelines about how the audit should be conducted, rather than by applying automatic or semi-automatic analysis tools. Several commercial resources exist, for example the Privacy Management Toolkit [InfoShield] providing templates, forms, regulations library and expert commentary, and the Compliance Meter [CompMeter] which assigns privacy compliance scores based on the expert review of the templates filled by the customers. Following the European Committee for Standardisation (CEN) workshop on data protection & privacy (WS/DPP) in 2005, several workshop agreement (CWA) reference documents have been developed, including those on “Personal Data Protection Audit Framework" defining standard practices, templates, questionnaires and processes for audits.

The research on privacy metrics is still in its incipient stage and therefore not many tools exist. One of the more researched areas is the anonymity metrics of datasets, where some metrics are already available and thus naturally more tools exist to measure them. For example, the Privacy Analytics Risk Assessment Tool [PARAT] measures the risk of re-identification in different scenarios. There are tools providing capabilities both for anonymised dataset creation and for evaluation of the anonymisation status, like [CAT] and UTD Anonymisation ToolBox [UTDToolBox].

Another area where some metrics exist is on access control policy. Although access control is usually associated with security, it has also privacy aspects, mainly related to the policy adequateness and conflicts. For example, if the policy allows all the users to access all the data, there is probably a privacy violation. There are works that analyse the access control policy for finding conflicts and dominance [Vaniea08, Martin07], but little exists in terms of assessment of the quality and quantity of the privileges given to the various roles for accessing various data objects.

Protecting individual privacy on the web draws a lot of research attention. Web site privacy seal solutions, such as TRUSTe [TRUSTe], provide certain web site assessment capabilities. For example, TRUSTe is able to verify the site against its privacy policy but not compliance, and to scan the site for potential threats but not towards compliance with legal regulations. However, there is not enough transparency in terms of how exactly these capabilities are achieved, what particular assessment steps are performed and what techniques are used. Moreover, the TRUSTe seal does not address EU regulations, especially in terms of data collection.

Protecting individual privacy on a web site requires first of all that the site itself is secured from any type of hacking. Several commercial tools for testing web site vulnerability exist, for example, Acunetix Web Vulnerability Scanner [Acunetix] and IBM Rational AppScan [AppScan]. While these tools assess web site ability to resist various types of known attacks, it is not enough from the privacy preservation perspective. Web sites should be also examined in terms of the privacy policy existence and relevance, the limitation of the private data that the users are required to supply for clearly specified purposes, the processing that this data undergoes, the level of data protection within the data store, open sessions separation and more. There are no automatic tools with such wide assessment capabilities.

6 Conclusions

We have presented a conceptual framework for privacy auditing based on the legal concept of data protection goals, supported by a formal definition of technical privacy metrics. We described the privacy compliance assessment tools architecture, based on transparency and extensibility principles.

Our goal is to define a set of technical privacy metrics, by using a sound and formal approach to privacy quantification. This will represent a significant advance to the state-of-the-art for many reasons. First, the majority of previous proposals focused mainly on security. Those addressing privacy only considered data sharing by proposing a set of metrics to quantify the degree of anonymity of the released data. In contrast, we plan to develop a more general framework, in which data sharing is only one of the considered dimensions. Moreover, our ambitious goal is to combine both a sound and theoretical foundation of the developed metrics with an easy way of computing them and presenting the results to users.

We have designed a framework to allow easy plugging of privacy compliance metrics assessment components, each providing its specific user interface, analysis and reporting capabilities for the particular technical metrics. This approach enables gradual development, addition of more assessment techniques and areas in the future by other users and certification bodies, and the creation of a customised toolset packaging depending on particular target customer needs and audit type. Use of the framework will provide assessment transparency, by clearly showing how the decisions have been made, and what evidence has been collected during the analysis.

In this paper we have discussed the initial results we have achieved with our privacy preserving framework. The work is still in its early phases and a lot of work

remains to be done. In the near future we plan to work both on theoretical and implementation aspects. We plan to investigate the completeness and effectiveness of the conceptual model. We plan to identify new privacy metrics that can provide a measure of robustness to inference and statistical privacy attacks. We also plan to assess the capabilities and the scalability of our framework with case studies of realistic complexity. We also plan to work with users with the aim of getting feedback and suggestions on how it can be enhanced.

References

[AAL] Unabhaengiges Landeszentrum fuer Datenschutz (ULD). Juristische Fragen im Bereich Altersgerechter Assistenzsysteme, pre-study on behalf of VDI/VDE-IT, funded by the German Bundesministerium fuer Bildung und Forschung, https://www.datenschutzzentrum.de/projekte/aal/

[Acunetix] Acunetix Web Vulnerability Scanner, http://www.acunetix.com/vulnerability-scanner/

[AppScan] IBM Rational AppScan, http://www-01.ibm.com/software/awdtools/appscan/

[Article29] The Article 29 Data Protection Working Party was set up under Article 29 of Directive 95/46/EC, http://ec.europa.eu/justice/policies/privacy/index_en.htm

[Bezzi2010] Bezzi, M.: Expressing privacy metrics as one-symbol information. In: Proc. of the 2010 EDBT/ICDT Workshops (2010)

[BL08] Byun, J.-W., Li, N.: Purpose based access control for privacy protection in relational database systems. VLDB J. 17(4), 603–619 (2008)

[BM2012] Bock, K., Meissner, S.: Datenschutz-Schutzziele im Recht. DuD – Datenschutz und Datensicherheit 36(6), 425–431 (2012)

[BSI] German Federal Office for Information Security, http://www.bsi.bund.de

[CAT] Xiao, X., Wang, G., Gehrke, J.: Interactive Anonymization of Sensitive Data. In: SIGMOD 2009 (2009)

[COBIT] ISACA: COBIT Framework for IT Governance and Control, http://www.isaca.org/Knowledge-Center/COBIT/Pages/Overview.aspx

[CompMeter] The Compliance Meter, http://www.compliancehelper.com/compliance-meter/

[CF2012] Colombo, P., Ferrari, E.: Towards a modeling and analysis framework for privacy aware systems. Technical report, University of Insubria (2012) (submitted for publication)

[Datta2011] Datta, A., et al.: Understanding and Protecting Privacy: Formal Semantics and Principled Audit Mechanisms. In: Proc. of the International Conference on Information Systems Security (2011)

[DFGK09] Datta, A., Franklin, J., Garg, D., Kaynar, D.K.: A Logic of Secure Systems and its Application to Trusted Computing. In: Proc. of the IEEE Symposium on Security and Privacy (2009)

[DGLKD10] DeYoung, H., Garg, D., Jia, L., Kaynar, D., Datta, A.: Experiences in the Logical Specification of the HIPAA and GLBA Privacy Laws. In: Proc. of 9th ACM Workshop on Privacy in the Electronic Society (October 2010)

[DSK] Ein modernes Datenschutzrecht fuer das 21. Jahrhundert, Eckpunkte; Konferenz der Datenschutzbeauftragten des Bundes und der Laender, http://www.lfd.m-v.de/dschutz/beschlue/Eckpunkte.pdf (presented on March 18, 2010)

[Dwork2008] Dwork, C.: Differential Privacy: A Survey of Results. In: Agrawal, M., Du, D.-Z., Duan, Z., Li, A. (eds.) TAMC 2008. LNCS, vol. 4978, pp. 1–19. Springer, Heidelberg (2008)

[EOS09] Evesti, A., Ovaska, E., Savola, R.: From Security Modelling to Run-time Security Monitoring. In: Proc. of the Fifth European Conference on Model-driven Architecture Foundations and Applications, Enchede, The Netherlands (June 2009)

[EuroPriSe] EuroPriSe, the European Privacy Seal for IT Products and IT-Based Services, http://www.european-privacy-seal.eu

[GB2012] Geisberger, E., Broy, M. (eds.): AgendaCPS, Integrierte Forschungsagenda Cyber-Physical Systems, acatech Studie, Deutsche Akademie der Technikwissenschaften (2012)

[HDB] IBM Hippocratic Database (HDB) Technology Projects, http://www.almaden.ibm.com/cs/projects/iis/hdb/hdb_projects.shtml

[Herrmann07] Herrmann, D.S.: Complete guide to security and privacy metrics – measuring regulatory compliance, operational resilience and ROI. Auerbach Publications (2007)

[HSHJ08] Heyman, T., Scandariato, R., Huygens, C., Joosen, W.: Using security patterns to combine security metrics. In: Proc. of the 3rd Int. Conf. on Availability, Reliability and Security (ARES) (2008)

[InfoShield] The Privacy Management Toolkit, http://www.informationshield.com/privacy_main.html

[ITIL] Arraj, V.: ITIL - IT Infrastructure Library, The Basics, White Paper, http://www.itil-officialsite.com/AboutITIL/WhatisITIL.aspx (downloaded January 1, 2012)

[Jaquith2007] Jaquith, A.: Security metrics: replacing fear, uncertainty and doubt. Addison-Wesley (2007)

[JABK2008] Jouault, F., Allilaire, F., Bézivin, J., Kurtev, I.: Atl: A model transformation tool. Science of Computer Programming 72(1-2) (2008)

[LDSG-SH] Schleswig-Holstein Act on the Protection of Personal Information of February 9, 2000 last amended by Article 1 of the Act to amend the State Data Protection Act (January 11, 2012) (GVOBl. Schl.-H. p. 78)

[LTV07] Li, N., Li, T., Venkatasubramanian, S.: t-closeness: Privacy beyond k-anonymity and l-diversity. In: Proc. of the 23rd IEEE International Conference on Data Engineering (ICDE 2007). IEEE Computer Society (April 2007)

[Martin07] Martin, E.: Testing and Analysis of Access Control Policies. In: ICSE 2007 (2007)

[MASTER] Managing Assurance, Security and Trust for Services, European research project, http://cordis.europa.eu/fetch?CALLER=PROJ_ICT&ACTION=D&CAT=PROJ&RCN=85559

[MGKV06] Machanavajjhala, A., Gehrke, J., Kifer, D., Venkitasubramaniam, M.: l-diversity: Privacy beyond k-anonymity. In: Proc. of the 22nd IEEE International Conference on Data Engineering (ICDE 2006). IEEE Computer Society, Washington, DC (2006)

[OCL] OMG, Object Constraint Language (OCL) (2012), http://www.omg.org/spec/OCL/2.3.1

[PARAT] PARAT, http://www.privacyanalytics.ca/products.asp

[PIA] European Commission (EC): The Privacy Impact Assessment Framework for RFID Applications: PIA Framework (January 2011), http://ec.europa.eu/information_society/policy/rfid/pia/index_en.htm

[PICOS] Privacy and Identity Management for Community Services, European research project, http://cordis.europa.eu/fetch?CALLER=PROJ_ICT&ACTION=D&CAT=PROJ&RCN=85533

[PRBAC] Ni, Q., Bertino, E., Lobo, J., Brodie, C., Karat, C.-M., Karat, J., Trombeta, A.: Privacy-aware role-based access control. ACM Trans. Inf. Syst. Secur. 13(3), Article 24 (July 2010)

[Probst 2012] Generische Schutzmassnahmen für Datenschutz-Schutzziele. DuD – Datenschutz und Datensicherheit 36(6), 439–444 (2012), https://www.european-privac-seal.eu/results/articles/201206-DuD-Probst.pdf

[QVT] OMG, Meta Object Facility (MOF) 2.0 Query/View/Transformation (QVT) (2011), http://www.omg.org/spec/QVT/1.1

[RFD09] Rebollo-Monedero, D., Forne, J., Domingo-Ferrer, J.: From t-closeness-like privacy to postrandomization via information theory. IEEE Transactions on Knowledge and Data Engineering 99(1) (2009)

[RP2009] Rost, M., Pfitzmann, A.: Datenschutz-Schutzziele – revisited. DuD – Datenschutz und Datensicherheit 33(6), 353–358 (2009)

[Rost2011] Rost, M.: Datenschutz in 3D. DuD – Datenschutz und Datensicherheit 35(5), 351–353 (2011)

[RB2011] Rost, M., Bock, K.: Privacy by Design und die neuen Schutzziele. DuD – Datenschutz und Datensicherheit 35(1), 30–35 (2011)

[SA2009] Savola, R., Abie, H.: Development of Measurable Security for a Distributed Messaging System. International Journal on Advances in Security 2(4), 358–380 (2010) ISSN 1942-2636

[Savola2006] Savola, R.: A Requirement Centric Framework for Information Security Evaluation. In: Yoshiura, H., Sakurai, K., Rannenberg, K., Murayama, Y., Kawamura, S.-i. (eds.) IWSEC 2006. LNCS, vol. 4266, pp. 48–59. Springer, Heidelberg (2006)

[Savola2010] Savola, R.: Towards a Risk-Driven Methodology for Privacy Metrics Development. In: Proc. of the Symposium on Privacy and Security Applications (PSA 2010) (August 2010)

[Schmidt2006] Schmidt, D.C.: Model-Driven Engineering. IEEE Computer 39(2) (2006)

[SPMNLH04] Seddigh, N., Pieda, P., Matrawy, A., Nandy, B., Lambadaris, J., Hatfield, A.: Current trends and advances in information assurance metrics. In: Proc. of the 2nd Annual Conference on Privacy Security and Trust (2004)

[Sweeney2002] Sweeney, L.: k-anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems 10(5), 557–570 (2002)

[TRUSTe] TRUSTe, http://www.truste.com/privacy_seals_and_services/enterprise_privacy/web_privacy_seal

[UML] OMG, Unified Modeling Language, v2.4.1 (2011), http://www.omg.org/spec/UML/2.4.1/

[UTDToolBox] UTD Anonymization ToolBox, http://cs.utdallas.edu/dspl/cgi-bin/toolbox/index.php

[Vaniea08] Vaniea, K., Ni, Q., Cranor, L., Bertino, E.: Access control policy analysis and visualization tools for security professionals. In: USM 2008: Workshop on Usable IT Security Management (2008)

[XACML] OASIS eXtensible Access Control Markup Language (XACML), http://www.oasis-open.org/committees/xacml/

[ZH2012] Zwingelberg, H., Hansen, M.: Privacy Protection Goals and Their Implications for eID Systems. In: Camenisch, J., Crispo, B., Fischer-Hübner, S., Leenes, R., Russello, G. (eds.) Privacy and Identity Management for Life – 7th IFIP WG 9.2, 9.6/11.7, 11.4, 11.6 International Summer School Trento, Italy (September 2011); Revised Selected Papers. Springer, Boston (2012) (to appear)

Privacy-Preserving Computation
(Position Paper)

Florian Kerschbaum

SAP Research
Karlsruhe, Germany
florian.kerschbaum@sap.com

Abstract. Private data is commonly revealed to the party performing the computation on it. This poses a problem, particularly when outsourcing storage and computation, e.g., to the cloud. In this paper we present a review of security mechanisms and a research agenda for privacy-preserving computation. We begin by reviewing current application scenarios where computation faces privacy requirements. We then review existing cryptographic techniques for privacy-preserving computation. And last, we outline research problems that need to be solved for implementing privacy-preserving computations. Once addressed, privacy-preserving computations can quickly become a reality enhancing the privacy protection of citizens.

1 Introduction

There are very strong privacy-enhancing techniques for communication and authentication available to citizens, but when it comes to actually processing the data few choices are available. Projects such as ANON and JAP [10] or TOR [20] enable anonymous communication. Tools such as IDEMIX [15] even enable anonymity-preserving authentication over these communication channels. Nevertheless, the data sent over anonymous (and authenticated) channels is rarely protected from the recipient.

This problem becomes prevalent when outsourcing storage and computation, e.g., to the cloud. Service providers such as Facebook or Google have made a business out of exploiting this data for advertising purposes. Big data is collected about habits and preferences of customers.

This problem is not due to a lack of available security mechanisms. Cryptography provides tools for privacy-preserving computations, such as secure multi-party computation, homomorphic encryption, order-preserving encryption or zero-knowledge proofs. In contrast to anonymous communication and authentication these mechanisms often lack a user-friendly implementation. This paper proposes a research agenda for implementing privacy-preserving computation bringing it closer to users' (and service providers') adoption. This paper does not argue that no further research in cryptography is necessary – quite the contrary, but existing mechanism should already be made available to system implementers. Furthermore, this paper focuses on the computational aspect,

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 41–54, 2014.

i.e., securing data while computing, and not data privacy aspects dealt with k-anonymity [52] or differential privacy [21].

In the next section we present two exemplary scenarios where computation on private data is strictly necessary, but not desired or even illegal. In criminal investigations police institutions often need to exchange data about subjects, places or objects, but are (rightfully) restricted to necessary and proportionate cases. This may hinder investigation success. Privacy-preserving computation enables performing this data exchange, such that only data which is clearly linked to a related case is revealed, but other data is kept private. Innocent suspects are protected while criminal investigations are fostered.

In the future smart electricity grid household collect fine-grained consumption data and transmit it to the grid and utility providers. This data enables inferences about the household inhabitants and their preferences and represents a significant privacy invasion. In the Netherlands smart meter roll-outs have been stopped for this reason. Privacy-preserving computation enables smart meter data processing, such as billing, without revealing one's consumption data to the providers. It thereby reconciles the need for a smart grid with the privacy concerns of the citizens.

In Section 3 we review and compare a number of mechanisms for privacy-preserving computation. Secure multi-party computation [54] has been invented thirty years ago, but implementations are only beginning to emerge. Recently fully homomorphic encryption [22] has arisen as an alternative, but theoretical solutions are still to inefficient to be put to practice. Outsourced storage [11,45] can be efficiently processed if it is encrypted using searchable [14] or order-preserving encryption [3,13]. Data collected by cyber-physical sensors can be securely (and privately) processed using zero-knowledge proofs [26]. We will only gives a coarse characterization of techniques in order to judge their applicability in different scenarios.

Finally, in Section 4 we outline an agenda of tools that will help implement privacy-preserving computations and thereby increase user adoption. Currently, designing and implementing privacy-preserving computations is difficult. The programmer has to have problem-specific know-how for its application, programming experience to select the best implementation method and security awareness in order to ensure privacy. In a lot of cases the programmer has to make a choice between secure or privacy guarantees and efficiency. If he chooses to be efficient security needs to verified manually. Furthermore, every new computation (even if done at run-time) requires the same level of effort. We present a proposal for tools that should help remedy these problems.

2 Scenarios

We consider two exemplary scenarios where privacy needs to be protected during computation, i.e. they require privacy-preserving computation. These are criminal investigations and smart meter billing. Although both scenarios deals with computation on private data their most prominent solutions use different

techniques. Whereas criminal investigations are implemented using secure computation, smart meter billing is implemented using zero-knowledge proofs. This also highlights the applicability of different scenarios to different technologies. We will try to devise the characteristics of the scenarios lending them to specific technologies.

2.1 Criminal Investigation

In federated states or organization of states, such as the European Union or the United States, a common approach to organized crime is necessary. For this purpose, federal law enforcement agencies, such as Europol or the FBI, have been established. Nevertheless, data privacy laws (rightfully) restrict supplying institutions from sharing their data, unless there is a hard corroborating evidence on a case and subject under investigation.

A common tool for the criminal investigator is data mining using social network analysis of the data stored in their warehouses. It graphically depicts the suspects and their connections to other people or artifacts, such as telephone numbers or bank accounts, and allows the computation of certain metrics. Not all the facts composing the entire picture of a case may be known to one investigator. In particular, in pan-European organized crime, local police forces may only be aware of a partial view of the picture.

This necessitates data exchange between the institutions, but European data privacy laws restricts data exchange to necessary and proportionate cases. Therefore we propose a solution where the local investigator, or an investigator at the superordinate institution has access to all information, but without revealing sensitive or private details. This allows the investigator to still use SNA and profit from its achievements without breaking individual privacy rights or guidelines of other institutions.

Therefore privacy-preserving SNA – a special form of privacy-preserving data mining – has been suggested in the literature [39]. Each party inputs their view of the social network to a secure computation and the result is the anonymized combined view. No additional information, e.g., about unrelated suspects, is revealed, i.e., their privacy is being preserved.

Note that in order to compare for identical entries in the social network – a computation on joint data is necessary. No party can by itself decide whether the data of the other party matches its own. It is therefore important to note that in this scenario both parties provide (private) input.

For the practical adoption of privacy-preserving computation in criminal investigations several legal, political and social considerations have to be taken into account. We argue that the ready-to-use availability of the technology will spur the public discussion on these topics.

2.2 Smart Meter Data

Smart metering refers to the collection of consumption profiles at customer's households with the help of so called smart meters (SM). Smart meters measure

electricity consumption in households and communicate their readings at regular intervals to the back-end system. Alternatively, the back-end system can also query the smart meter for its data (pull). A Trusted Platform Module in the smart meter holds key material and creates signatures over the data to ensure authenticity and integrity until it arrives at the back-end system. There the consumption profile and the tariff data from the respective customer's contract are used to calculate the price the customer has to pay for the time period covered by the profile.

Smart metering has encountered massive privacy concerns from media [35], data privacy experts [16] and consumers [30]. The fact that whole consumption profiles of households are transmitted to and stored by suppliers is troubling w.r.t. customer privacy. Data confidentiality can be easily protected in transit between smart meter and back-end system. However, their storage at the suppliers' IT-systems still endangers customer privacy. Depending on resolution and the availability of different services' profiles (e.g. water, heat, electricity) one can read the profile and "see" more or less clearly what happens in the household: For instance, when family members wake up (light switched on), whether they shower in the morning (water, heat, and electricity for water heater), whether they drink hot beverages with their breakfast and when or if they leave for work or school. Furthermore, the frequency of washing and drying clothes, cooking or the amount of time the TV is turned on can be inferred. For further research on what electricity consumption profiles tell about the inhabitants see [7,29,40,51].

These inferences make consumption profiles very privacy-sensitive data and these profiles might even have value in the advertising market, for instance. On one hand, disgruntled employees or external attackers might attempt to steal it for profit or out of malice. On the other hand, the supplier could seek subsidiary revenues by selling this data himself. Depending on the local jurisdiction, this might even be legal.

The important point is, that currently there are no reliable, technical measures in place to prevent abuse of consumption profiles. Merely organizational measures, policies or laws sanction the abuse of privacy related data but require a trace or proof of abuse and do not prevent it in the first place.

First [36] and later independently [46] introduce a privacy component into the standardized smart meter / meter data management (MDM) reporting communication link. This component hides the actual consumption profile from the MDM and therefore also from the supplier. This only requires small changes compared to current smart meter reporting. The privacy component intercepts smart meter readings, then uses tariff information provided externally (over the Internet or by the MDM) to calculate the billing amount and sends only the resulting billing amount to the MDM.

The technical solution in this scenario is a Zero-Knowledge Proof (ZKP). The user proves to the service provider that it truthfully computed the bill. This works for two reasons: First, the user only computes on his private data, i.e., there is no second party (private) input. Second, the integrity and authenticity

of the data is ensured by trusted components. The smart meters can digitally sign their measurements and are implemented with trusted hardware.

We envision this scenario setup to be quite common in cyber-physical systems. It has already been described by Danezis and Livshits in a related position paper for cloud computing [19]. Sensors collect data about people. Obviously, this data may be privacy-sensitive and protected by privacy legislation. These sensors can ensure the integrity of the collected data, but should not reveal it. Using a ZKP the user can perform the computation itself in a privacy-preserving fashion and prove correct computation. Other examples of this scenario setup include, e.g., road toll pricing [6] or e-ticketing in public transportation [32].

For the practical adoption of this technology a user device (plug-in component) needs to be part of the protocol [36]. While the business implications of this design are not field-tested, we argue that increased privacy of the consumer might make an interesting business case.

3 Security Mechanisms

In this section we briefly and mostly non-technically review the mechanisms of secure computation, homomorphic encryption, order-preserving encryption and zero-knowledge proofs. We address the properties of each in terms of security, performance and functionality. Balancing these three objectives is the challenge of privacy-preserving computation. Particularly, we highlight their suitability for certain application scenarios.

3.1 Secure Multi-party Computation

Yao introduced secure two-party computation in [54]. Secure (two-party) computation allows two parties to compute a function f over their joint, private inputs x and y, respectively. No party can infer anything about the other party's input (e.g. y) except what can be inferred from one's own input (e.g. x) and output (e.g. $f(x, y)$).

Yao's initial protocol uses a technique called garbled circuits. Alice prepares a circuit for the function to be computed and encrypts (and garbles) this circuit. The encrypted circuit is transferred to Bob who obtains keys for his input using oblivious transfer. Bob then decrypts (part of) the circuit obtaining the function result. For a detailed, technical description of circuit garbling and its implementation see [42].

Now, there are many protocols for secure computation. They can be classified into generic and special protocols. Generic protocols can implement any functionality whereas special protocols implement one specific function. Generic protocols contain a translation step for the function – similar to Yao's garbled circuit construction. Their security proof, however, is independent of the function. Special protocols are usually more efficient, since they use problem insight to optimize the protocol. They need to proven secure manually for each protocol instance.

For a very long time generic protocols exist for multi-party computations in the computational [25] and information-theoretic setting [9]. Also they exist for many different security models. Most notable are the semi-honest and malicious security models [24]. In the semi-honest model the parties are assumed to follow the protocol. In the malicious model they may deviate arbitrarily. The policy implications of these models need to be discussed in the context of the concrete use case and its actors, e.g., criminal investigations. A detailed investigations is subject to further research and out of scope of this position paper.

Security of secure computations is often defined by comparison to an ideal model. In the ideal model there is a trusted third party. All parties send their inputs to the trusted third party which computes the function and returns the result to the parties. The function implemented by the third party is also ideal functionality. Each attack feasible in the real protocol execution must also be feasible in this ideal model.

It is important to note that neither model prevents inferences about the input from the result. This may be particularly sensitive if one party may influence the function to be computed. For example, assume that one may party may privately issue a query about the other's party private database. Then it is hard to preserve the privacy of the database without additional measures such as differential privacy [21]. Furthermore, no security model prevents the substitution of inputs. Therefore an interest in the correct computation of the result needs to be assumed. This is subject to economic security models, such as non-cooperative computation [50]. Even when implementing a non-cooperative computation secure computation may be difficult to implement [2,27,28]. Nevertheless using specialized protocols it can even be implemented for mixed security models [41]. Implementing a proxy, e.g., using cloud computing, may help solve the problem [37]. In theory, it is possible to implement any computation using rational players [34], but it requires physical assumptions.

There exist a number of domain-specific programming languages for implementing secure computations [8,12,18,31,33,42,48]. They can be classified into those tied to a generic protocol [8,12,31,42] or those based on generic programming languages [18,33,48]. The second kind can implement a wider variety of protocols, but also enables implementing insecure protocols. The ones tied to a specific protocol may be proven secure independent of the functionality. Such a proof extends to all protocols implemented in this language, but the language prevents implementing many special, possibly more efficient protocols.

We classify them into systems specifying the ideal functionality and systems specifying the protocol description. Just FairPlay [42] and FairPlayMP [8] are instances of systems which only describe the ideal functionality of a secure computation, i.e. *what* is to be implemented by the protocol. All of the other languages, compilers or frameworks are instances of systems where the programmer can – at least partially – specify *how* the protocol is implemented. This approach leads to significantly more efficient protocols, but puts an additional burden on the programmer.

3.2 Homomorphic Encryption

Homomorphic encryption supports a homomorphism of (at least) one arithmetic operation on the ciphertexts to an arithmetic operation on the plaintexts. Additively homomorphic operation supports addition as the homomorphic operation on the plaintexts. Let $E(x, r)$ denote the encryption of plaintext x with randomization parameter r. Then the following addition properties hold

$$D(E(x)E(y)) = x + y$$

The most popular additively homomorphic encryption system is Paillier's [44]. It is public key and satisfies modern security definitions (semantic security). Its performance is comparable to other public key encryption systems.

Gentry recently developed a fully homomorphic encryption scheme [22]. It supports both, addition and multiplication, and therefore allows the computation of arbitrary functions on the ciphertext. Nevertheless, its performance is severely restricted in practice [23] and more efficient schemes are still subject to research.

As middle-ground there are somewhat homomorphic encryption schemes. They support a limited number of multiplications and thereby enable to compute a wider class of functions than purely additively homomorphic encryption. Their performance is comparable to additively homomorphic encryption [43].

A fundamental advantage of homomorphic encryption compared to secure computation is its non-interactivity. The client submits input and a server can perform the computation without learning anything but ciphertexts. The result is a ciphertext as well. Therefore computations on homomorphic encryption can be performed off-line and do not depend on the network performance. The client only has to communicate data linear in its input length whereas in secure computation it is linear in the function's complexity.

Nevertheless, besides its performance homomorphic encryption also has a number of limitations. First, the result and each intermediate result is encrypted. Therefore the server cannot make any decision based on its value. Similar to secure computation the server therefore needs to compute over all choices, i.e., in every conditional branch both branches need to be evaluated. This increases the complexity of each function to its worst case complexity and further increases the performance penalty.

More severely, computations on homomorphic encryption must be performed under the same key. It has been proven in [53] that no fully homomorphic encryption scheme with multiple keys can exist. This implies a collusion attack for collaborative computations. When two parties submit input, one party only holds the public key. The input data of this party may be decrypted by the private key holder, e.g., by a collusion with the server. Homomorphic encryption therefore seems less suitable for collaborative computations.

Additively homomorphic encryption can also be used to implement secure computation [17]. This makes the computation interactive again, but prevents the collusion attack. Usually secure computation based on homomorphic encryption is less efficient than other generic protocols [47].

3.3 Order-Preserving Encryption

Order-preserving encryption [3,13] ensures that the order (greater-than relation) of the ciphertexts is the same as the order of the corresponding plaintexts. This allows a server to efficiently search on the ciphertexts using binary search or perform range queries. In turn this capability enables performing most database queries on encrypted data [11,45].

Efficient encrypted database – where the key is stored outside of the database – have many applications in privacy-preserving computing. Cloud computing using a database-as-a-server model can be secured against the service provider such that insider or targeted attacks are significantly complicated. As such real-world applications of prototypical system have been already reported [11].

The efficiency gain of order-preserving encryption mainly stems from its main difference to the homomorphic encryption. The result of the computation on the ciphertexts is publicly available to the server performing the computation. This enables implementing significantly more efficient algorithms.

Another instance of such an encryption scheme is searchable encryption [14]. Searchable encryption allows to compare for equality of plaintexts using a token issued by the private key holder. Differently from order-preserving encryption it can be proven secure in more standard security models. The best security proof for order-preserving encryption is that it is as secure as possible under the order-preserving constraints [13]. How secure this level of security is still subject to research. Searchable encryption has been proven secure against chosen plaintext attacks.

Searchable encryption requires a linear scan over the data, i.e., an index is useless. It therefore has not found the same acceptance in the database community as order-preserving encryption, although it also allows range queries [49].

Clearly, since the result of the comparison is revealed, this type of encryption is limited to specific functions. First, the size of the ciphertexts needs to remain manageable. Second, certain functions may allow breaking the encryption scheme. For example, a public-key order preserving encryption scheme would allow binary search for arbitrary plaintexts. It cannot be secure. Therefore all order-preserving encryption schemes are symmetric.

Due to the limitation of the functionality, order-preserving encryption has limited applications. A secure, encrypted database is certainly a great achievement, but more complicated applications seem to be difficult to design. They may not be securely implementable and if they are, then their construction can be very complicated.

3.4 Zero-Knowledge Proofs

Zero-Knowledge Proofs (ZKP) have been introduced in cryptography a long time ago [26]. They allow the proof of knowledge of some data that satisfies a certain function. In the example of smart meter data, this function is the signatures by the smart meters and the billing amount. The household then proofs knowledge of consumption data that is signed and amount to the bill.

ZKPs in the first place ensure integrity, i.e., the function has been computed truthfully. As such their application to privacy-preserving computation is not obvious, but ZKPs allow the outsourcing of the computation while verifying the integrity of the computation. Therefore they allow the outsourcing the originator of the (private) data, although it might not be trusted. At least it can ensure the confidentiality (privacy) of the data. The ZKP ensures that it cannot cheat.

Consequently, ZKPs are applicable in settings where data is collected about a person, such as smart metering or road-toll pricing. ZKPs are similarly limited as homomorphic encryption when it comes to input of multiple parties. Then secure computation is the method of choice, but ZKPs can also be implemented non-interactively. Nevertheless due to the emergence of more and more cyber-physical systems where sensors collect data about person ZKPs will probably enjoy wider adoption.

The initial ZKPs were generic allowing to proof for any function in NP. Then some special ZKPs were designed, e.g., for discrete logarithms or range proofs. Recently, compilers translating function descriptions into ZKPs have been presented [4,5].

ZKPs are also the basis for privacy-preserving authentication. A user can proof the possession of an attribute - such as age or driver's license – without revealing any private information. Furthermore, all transaction may remain unlinkable. Therefore the future adoption of ZKPs to privacy-preserving computation seems even more likely. We recommend to also consider the position paper [19] by Danezis and Livshits which already outlines similar ideas on ZKPs.

4 Research Agenda

In order to foster the adoption of privacy-preserving computation for many more applications there need to be tools. These tools need to ease the design and development of privacy-preserving computing applications. Currently, there is a lot of manual effort in developing a privacy-preserving application. This hinders adoption in practice due to a lack of skills and available resources. Development tools can significantly lower the barrier of adoption. Other characteristics such as security (privacy) and even performance can be made to fit, if the system is designed cleverly.

4.1 Compiler

As mentioned before there already exist some compilers for secure computation and zero-knowledge proofs. These compilers unfortunately currently still fall short of the requirements of the developers. We will highlight some design principles, challenges and ideas for achieving the full potential. These design principles should be seen as visionary, ambitious objectives and even partial, but rigorous achievement can be viewed as scientific success.

I. *Only the ideal functionality should be specified.*

The principle of FairPlay shows the right direction. It is already difficult enough for the programmer to acquire problem domain-specific knowledge in order to design a successful application. If the programmer also has to care about security (and complex performance aspects), he will be overburdened. The compiler has to take care of it.

A necessary additional specification is the data origin, i.e., which party provides which input. There may be a need for some security policies, but largely extending current access control. Privacy-preserving computation allows to enforce policies, such as security against a service provider. There is no longer a grant or deny decision.

In the system architecture there may also be a need to specify trust relationships. Nevertheless, there should not be a restriction to or reliance on specific trust assumptions. Instead, privacy-preserving can reduce or even remove most of the commonly necessary trust assumptions.

II. *Security should be guaranteed.*

Except for the specification of policies and trust assumption security must be guaranteed. Every successfully compiled program must be secure. The definition of security must depend on the policy and the security model (which may also be specified as part of the policy). Ideally, the compiler generates a proof or is certified that the compiled program is secure. We have seen for secure computations that some domain-specific languages violate this principle for security reasons.

There are several security models available in the cryptographic literature, but sometimes they may fall short in capturing the system's dependencies. For example, it is easy to design a protocol that is secure in the semi-honest protocol, if all inputs can be revealed by the result of the function. Therefore additional tools analyzing the functionality for admissible information flow in the entire system are necessary. Only, if the programmer is not capable of "shooting himself in the foot", a system can be considered simple to implement and secure. Therefore, there is a clear need for security models and tools augmenting the available cryptographic security models.

Language annotations, such as type systems, may significantly simplify this problem. E.g. type systems for information flow [1] allow only specifying programs that do not contain non-admissible information flows. Type systems can encode any type of security proof, but the research challenge remains to analyze the admissible information flows.

III. *Performance optimization should be automatic.*

Out of the set of admissible protocols according to the security model the compiler should select the best performing one. This selection should be automatic, i.e., without the need of specification by the programmer. There are several challenges and approaches to this problem.

First, the compiler may use many of the algorithms available to the compiler community. Algorithms like data flow or program analysis need to be augmented for additional criterion – security. We see a number of results in this area, e.g., in secure computation optimization [38].

Second, there is a need for a clear performance model. The typical complexity measure of algorithms do not fit any more, since there is this additional criterion of security. Let n be the input length, then we also have the security parameter k. Now, what is faster $O(n^2k)$ or $O(nk^2)$? Performance models can make decision for this and are already often used in other compilers, e.g., SQL query optimization.

Given a comprehensive performance model, the programmer should no longer need to specify the type of mechanism to be used. Instead, based on analysis of the data model and the function the compiler should select the optimally suitable one, e.g., secure computation or homomorphic encryption. Admittingly, some methods, such as order-preserving encryption, can be hard to fit into this model, particularly due to their unclear security model.

Based on the problem of complicated complexities the programmer may actually implement a sub-optimal algorithms, because he is not aware of all information for the decision. Therefore the compiler should be equipped with capabilities to rewrite programs, such that they perform better. Some rewriting techniques, e.g., common sub-expression elimination, have been developed in compiler design. Again, these need to be augmented by a security criterion. Furthermore, since there is no programming language established yet, we can also adapt the design of the language.

5 Conclusions

In this paper we investigated the status and future challenges of privacy-preserving computation. We have the security mechanisms available, but face the challenge of implementing them. Performance has already been proven in several applications not to be the prohibiting factor anymore and available computational resources continue to increase.

The problem will shift to the development of the applications. Current tools do not scale to the expected increase in privacy-preserving computation. We outlined some exemplary application which can serve as blue prints for others waiting to be implemented.

We then outlined the research challenges for a compiler for privacy-preserving computation. Based on three principles relating to the three objectives of privacy-preserving computation we described some research challenges and approaches. The purposes of this agenda is entice discussion and interest among interdisciplinary stakeholders in order to foster the adoption of privacy-preserving computation. Seeing what will be possible may lift some of the prejudices privacy-preserving computation currently faces. Ultimately only the uptake of technology will lead to a better protection of the citizen's privacy.

References

1. Abadi, M., Morrisett, G., Sabelfeld, A.: Language-based security. Journal of Functional Programming 15(2), 129 (2005)
2. Abraham, I., Dolev, D., Gonen, R., Halpern, J.: Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation. In: Proceedings of the 25th ACM Symposium on Principles of Distributed Computing, PODC 2006 (2006)
3. Agrawal, R., Kiernan, J., Srikant, R., Xu, Y.: Order preserving encryption for numeric data. In: Proceedings of the ACM International Conference on Management of Data, SIGMOD 2004 (2004)
4. Almeida, J.B., Bangerter, E., Barbosa, M., Krenn, S., Sadeghi, A.-R., Schneider, T.: A certifying compiler for zero-knowledge proofs of knowledge based on σ-protocols. In: Gritzalis, D., Preneel, B., Theoharidou, M. (eds.) ESORICS 2010. LNCS, vol. 6345, pp. 151–167. Springer, Heidelberg (2010)
5. Backes, M., Maffei, M., Pecina, K.: Automated synthesis of privacy-preserving distributed applications. In: Proceedings of 19th Network and Distributed System Security Symposium, NDSS 2012 (2012)
6. Balasch, J., Rial, A., Troncoso, C., Preneel, B., Verbauwhede, I., Geuens, C.: Pretp: privacy-preserving electronic toll pricing. In: Proceedings of the 19th USENIX Conference on Security, USENIX Security 2010 (2010)
7. Bauer, G., Stockinger, K., Lukowicz, P.: Recognizing the use-mode of kitchen appliances from their current consumption. In: Barnaghi, P., Moessner, K., Presser, M., Meissner, S. (eds.) EuroSSC 2009. LNCS, vol. 5741, pp. 163–176. Springer, Heidelberg (2009)
8. Ben-David, A., Nisan, N., Pinkas, B.: Fairplaymp: a system for secure multi-party computation. In: Proceedings of the 15th ACM Conference on Computer and Communications Security, CCS 2008 (2008)
9. Ben-Or, M., Goldwasser, S., Wigderson, A.: Completeness theorems for non-cryptographic fault-tolerant distributed computation. In: Proceedings of the 20th ACM Symposium on Theory of computing, STOC 1988 (1988)
10. Berthold, O., Federrath, H., Köhntopp, M.: Project "anonymity and unobservability in the internet". In: Proceedings of the 10th Conference on Computers, Freedom and Privacy: Challenging the Assumptions, CFP 2000 (2000)
11. Binnig, C., Hildenbrand, S., Färber, F.: Dictionary-based order-preserving string compression for main memory column stores. In: Proceedings of the ACM International Conference on Management of Data, SIGMOD 2009 (2009)
12. Bogdanov, D., Laur, S., Willemson, J.: Sharemind: A framework for fast privacy-preserving computations. In: Jajodia, S., Lopez, J. (eds.) ESORICS 2008. LNCS, vol. 5283, pp. 192–206. Springer, Heidelberg (2008)
13. Boldyreva, A., Chenette, N., Lee, Y., O'Neill, A.: Order-preserving symmetric encryption. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 224–241. Springer, Heidelberg (2009)
14. Boneh, D., Di Crescenzo, G., Ostrovsky, R., Persiano, G.: Public key encryption with keyword search. In: Cachin, C., Camenisch, J.L. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 506–522. Springer, Heidelberg (2004)
15. Camenisch, J., Van Herreweghen, E.: Design and implementation of the idemix anonymous credential system. In: Proceedings of the 9th ACM Conference on Computer and Communications Security, CCS 2002 (2002)
16. Cavoukian, A., Polonetskyand, J., Wolf, C.: Smart privacy for the smart grid: embedding privacy into the design of electricity conservation. Identity in the Information Society 3(2), 275–294 (2010)

17. Cramer, R., Damgård, I.B., Nielsen, J.B.: Multiparty computation from threshold homomorphic encryption. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 280–300. Springer, Heidelberg (2001)
18. Damgård, I., Geisler, M., Krøigaard, M., Nielsen, J.B.: Asynchronous multiparty computation: theory and implementation. In: Jarecki, S., Tsudik, G. (eds.) PKC 2009. LNCS, vol. 5443, pp. 160–179. Springer, Heidelberg (2009)
19. Danezis, G., Livshits, B.: Towards ensuring client-side computational integrity (position paper). In: Proceedings of the ACM Cloud Computing Security Workshop, CCSW 2011 (2011)
20. Dingledine, R., Mathewson, N., Syverson, P.: Tor: the second-generation onion router. In: Proceedings of the 13th USENIX Conference on Security, USENIX Security 2004 (2004)
21. Dwork, C.: Differential privacy. In: Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (2006)
22. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: Proceedings of the 41st ACM Symposium on Theory of Computing, STOC 2009 (2009)
23. Gentry, C., Halevi, S.: Implementing gentry's fully-homomorphic encryption scheme. In: Paterson, K.G. (ed.) EUROCRYPT 2011. LNCS, vol. 6632, pp. 129–148. Springer, Heidelberg (2011)
24. Goldreich, O.: The Foundations of Cryptography. Basic Applications, vol. 2. Cambridge University Press (2004)
25. Goldreich, O., Micali, S., Wigderson, A.: How to play any mental game. In: Proceedings of the 19th ACM Symposium on Theory of Computing, STOC 1987 (1987)
26. Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof systems. SIAM Journal of Computing 18(1), 186–208 (1989)
27. Gordon, S.D., Katz, J.: Rational secret sharing, revisited. In: De Prisco, R., Yung, M. (eds.) SCN 2006. LNCS, vol. 4116, pp. 229–241. Springer, Heidelberg (2006)
28. Halpern, J., Teague, V.: Rational secret sharing and multiparty computation: Extended abstract. In: Proceedings of the 36th ACM Symposium on Theory of Computing, STOC 2004 (2004)
29. Hart, G.W.: Nonintrusive appliance load monitoring. Proceedings of the IEEE 80(12), 1870–1891 (1992)
30. Heck, W.: Smart energy meter will not be compulsory. NRC Handelsblad (April 2009), http://www.nrc.nl/international/article2207260.ece/Smart_energy_meter_will_not_be_compulsory
31. Henecka, W., Kögl, S., Sadeghi, A.-R., Schneider, T., Wehrenberg, I.: Tasty: tool for automating secure two-party computations. In: Proceedings of the 17th ACM Conference on Computer and Communications Security, CCS 2010 (2010)
32. Heydt-Benjamin, T.S., Chae, H.-J., Defend, B., Fu, K.: Privacy for public transportation. In: Danezis, G., Golle, P. (eds.) PET 2006. LNCS, vol. 4258, pp. 1–19. Springer, Heidelberg (2006)
33. Huang, Y., Evans, D., Katz, J., Malka, L.: Faster secure two-party computation using garbled circuits. In: Proceedings of the 20th USENIX Conference on Security, USENIX Security 2011 (2011)
34. Izmalkov, S., Micali, S., Lepinski, M.: Rational secure computation and ideal mechanism design. In: Proceedings of the 46th IEEE Symposium on Foundations of Computer Science, FOCS 2005 (2005)
35. Jamieson, A.: Smart meters could be 'spy in the home'. Telegraph (UK) (October 2009), http://www.telegraph.co.uk/finance/newsbysector/energy/6292809/Smart-meters-could-be-spy-in-the-home.html

36. Jawurek, M., Johns, M., Kerschbaum, F.: Plug-in privacy for smart metering billing. In: Fischer-Hübner, S., Hopper, N. (eds.) PETS 2011. LNCS, vol. 6794, pp. 192–210. Springer, Heidelberg (2011)
37. Kerschbaum, F.: Adapting privacy-preserving computation to the service provider model. In: Proceedings of the International Conference on Privacy, Security, Risk and Trust, PASSAT 2009 (2009)
38. Kerschbaum, F.: Automatically optimizing secure computation. In: Proceedings of the 18th ACM Conference on Computer and Communications Security, CCS 2011 (2011)
39. Kerschbaum, F., Schaad, A.: Privacy-preserving social network analysis for criminal investigations. In: Proceedings of the 7th ACM Workshop on Privacy in the Electronic Society, WPES 2008 (2008)
40. Lisovich, M.A., Mulligan, D.K., Wicker, S.B.: Inferring personal information from demand-response systems. IEEE Security and Privacy 8(1), 11–20 (2010)
41. Lysyanskaya, A., Triandopoulos, N.: Rationality and adversarial behavior in multi-party computation. In: Dwork, C. (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 180–197. Springer, Heidelberg (2006)
42. Malkhi, D., Nisan, N., Pinkas, B., Sella, Y.: Fairplay – a secure two-party computation system. In: Proceedings of the 13th USENIX Conference on Security, USENIX Security 2004 (2004)
43. Naehrig, M., Lauter, K., Vaikuntanathan, V.: Can homomorphic encryption be practical? In. In: Proceedings of the 3rd ACM Cloud Computing Security Workshop, CCSW 2011 (2011)
44. Paillier, P.: Public-key cryptosystems based on composite degree residuosity classes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 223–238. Springer, Heidelberg (1999)
45. Popa, R.A., Redfield, C.M.S., Zeldovich, N., Balakrishnan, H.: Cryptdb: protecting confidentiality with encrypted query processing. In: Proceedings of the 23rd ACM Symposium on Operating Systems Principles, SOSP 2011 (2011)
46. Rial, A., Danezis, G.: Privacy-preserving smart metering. In: Proceedings of the 10th ACM Workshop on Privacy in the Electronic Society, WPES 2011 (2011)
47. Schröpfer, A., Kerschbaum, F.: Forecasting run-times of secure two-party computation. In: Proceedings of the 8th International Conference on Quantitative Evaluation of Systems, QEST 2011 (2011)
48. Schröpfer, A., Kerschbaum, F., Müller, G.: L1 – an intermediate language for mixed-protocol secure computation. In: Proceedings of the 35th IEEE Computer Software and Applications Conference, COMPSAC 2011 (2011)
49. Shi, E., Bethencourt, J., Chan, T.-H.H., Song, D., Perrig, A.: Multi-dimensional range query over encrypted data. In: Proceedings of the IEEE Symposium on Security and Privacy, SP 2007 (2007)
50. Shoham, Y., Tennenholtz, M.: Non-cooperative computation: boolean functions with correctness and exclusivity. Theoretical Computer Science 343(1-2), 97–113 (2005)
51. Sultanem, F.: Using appliance signatures for monitoring residential loads at meter panel level. IEEE Transactions on Power Delivery 6(4), 1380–1385 (1991)
52. Sweeney, L.: k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10(5) (2002)
53. Van Dijk, M., Juels, A.: On the impossibility of cryptography alone for privacy-preserving cloud computing. In: Proceedings of the 5th USENIX Workshop on Hot Topics in Security, HotSec 2010 (2010)
54. Yao, A.C.: Protocols for secure computations. In: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science, FOCS 1982 (1982)

Designing Privacy-by-Design

Jeroen van Rest, Daniel Boonstra, Maarten Everts,
Martin van Rijn, and Ron van Paassen

TNO, Delft / The Hague, The Netherlands
{Jeroen.vanrest,Daniel.boonstra,Maarten.everts,
Martin.vanRijn,Ron.vanPaassen}@tno.nl

Abstract. The proposal for a new privacy regulation d.d. January 25th 2012 introduces sanctions of up to 2% of the annual turnover of enterprises. This elevates the importance of mitigation of privacy risks. This paper makes Privacy by Design more concrete, and positions it as the mechanism to mitigate these privacy risks.

In this vision paper, we describe how design patterns may be used to make the principle of Privacy by Design specific for relevant application domains. We identify a number of privacy design patterns as examples and we argue that the art is in finding the right level of abstraction to describe a privacy design pattern: the level where the data holder, data subject and privacy risks are described.

We give an extended definition of Privacy by Design and, taking Solove's model for privacy invasions as structuring principle, we describe a tool and method to use that tool to generate trust in systems by citizens.

Keywords: privacy, privacy design pattern, privacy-by-design, system engineering, trust, tooling.

1 Introduction

The European Commission is preparing a reform of the current European data protection directive [1]. In their proposal for a new Data Protection Regulation [2], the Commission stated that it will

> *"… examine [...] the possibilities for the concrete implementation of the concept of 'privacy by design' [...] to enhance data controllers' responsibility*." [3]

An explanation of the principle of Privacy by Design (PbD) is given in a footnote:

> "*The principle of 'Privacy by Design' means that privacy and data protection are embedded throughout the entire life cycle of technologies, from the early design stage to their deployment, use and ultimate disposal. This principle features inter alia in the Commission Communication on 'A Digital Agenda for Europe' - COM(2010) 245.*"

The Digital Agenda for Europe embraces the same definition of Privacy by Design with no further explanation or reference [4]. The trail for defining the concept stops there.

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 55–72, 2014.

Legislation makes extensive references to "privacy by design", but fails to specify what it means exactly, as has also been pointed out by Van Lieshout [5], who argues that a holistic approach to Privacy by Design indeed offers surplus value but that actual implementation is confronted with difficulties such as lack of economic incentives, legacy systems, and lack of adoption of trust of end-users and consumers in PbD.

The omission of a clear definition of PbD entails that for European citizens, policy makers, authorities and industry it is currently unclear what a request for PbD practically means. Still, many are confronted with such requests, not only in communications from the Commission, but also by local authorities[1] and in calls for the 7th Framework Research Programme. PbD currently is only an apparent solution, not a real one.

1.1 Privacy by Design: History of a Vague Concept

Although the term Privacy by Design was not yet used, a joint paper by the Dutch Data Protection Authority and TNO FEL[2] is one of the first papers that take a generic look at privacy from a designer's perspective [7]. The first explicit public reference to the term Privacy by Design can be found in the title of a workshop held at the conference "Computers, Freedom & Privacy 2000: Workshop on Freedom and Privacy by design" [8]. Around the same time, the EU FP5 project PISA focussed on Privacy Incorporated Software Agents under the name of PbD and Privacy Enhancing Technologies [9]. In North America, PbD was cultivated mainly by Cavoukian [10]. Cavoukian stated that PbD is based on seven "foundational principles" [11], but recently other principles, such as *data minimisation* are also emerging [12]. The translation of these principles to actual designs of systems is done by example. Therefore, everybody is free to postulate a particular design (process) as "Privacy by Design", and we see companies as Microsoft and IBM doing exactly that [13, 14].

1.2 Roles and Responsibilities

Cavoukian hypothesises an evolutionary approach [15] where industry and consumers find out together what works and what not. The advantage of this approach is that it leaves room for innovation whereas a purely government regulated approach would merely leave room to comply. However, a disadvantage of a pure evolutionary approach is the individual nature of it. If each (commercial) party can decide for itself what PbD means in their application domain, then citizens, consumers, local authorities, end users and policy makers will have to understand the differences between how

[1] Charter for a Democratic Use of Video Surveillance [6]: "*It is important to include in these protocols the "Privacy by design" method, which encourages personal data protection to be considered at the early stages of the system design.*" No further explanation of the concept is given.

[2] With support from the Information & Privacy Commissioner of Ontario of that time, Tom Wright.

each different party implements PbD. This asks a lot of those citizens, end users, etc., and is therefore at least a missed opportunity for transparency with regard to respecting citizen's rights, but also for market transparency.

In every market domain also operational governmental organisations are active in the role of data holder and designer, collecting data and designing and building information systems such as police, defence, healthcare institutes, education, etc. Whether these parts of government also can afford to make the mistakes[3] that are inherent to an evolutionary approach remains to be seen, because liability and responsibility may weigh differently on public services than they do on commercial businesses. This is a second disadvantage of a pure evolutionary approach.

An open question is what citizens and consumers actually need to know about PbD in order for the concept to be useful. Perhaps the concept of PbD has the biggest value between industry and government, not between industry and consumers?

PbD emphasises the role of the designer and integrator in preventing privacy breaches. This does not discharge a owner from taking his responsibility, as, in EU laws, the owner of the system (which makes him the data controller/holder) is responsible, not the party that designed or built it. These parties are not even "known" to data protection laws. The designer does have an influence on the use of technology because technology is not neutral as has been stated in the first of Kranzberg's truisms [16]. It has a function, which through its form –its design- is communicated to its users.

Privacy certification (PC) [17] and privacy impact assessment (PIA) are in itself not PbD. One could argue that a system that was designed according to the principles of PbD should have a good PIA result, and should very quickly and easily be certified by a PC.

1.3 (Behavioural) Economic Perspective

If individual and collective privacy interests could be aligned with economic interests then our economic interests would also stimulate privacy. However, a study on the benefits of Privacy Enhancing Technologies [18] found that individual citizens in general do not flock to products and services that protect their privacy better. This should be investigated further, but this first study undermines the potential market mechanism. Borking theorizes [19] that the application of models for customer adoption [20] should be further studied to address this issue.

If we cannot leave it to behavioural economics or a strict evolutionary approach to protect data subjects' privacy, then we may need a joint approach with industry and government together. Already, the proposal for a (new) EU privacy regulation [2] includes sanctions of up to 2% of the annual turnover (in case of international enterprises) or 1.000.000 EUR. Whether this will actually have an impact remains to be seen, but it creates the first economic incentive on an organisational level to address privacy risks. As legislation does not specify *how to mitigate* those risks, that void could be filled by making Privacy by Design more concrete.

[3] Bad evolutionary changes die off because the phenotypes that carry them are being "punished" by their environment. Who punishes bad forms of PbD? Will this happen before or after people's privacy has been breached?

1.4 Problem Statement

The specific meaning in a particular application domain of the principle of 'Privacy by Design' is unknown. It is an open question what citizens and consumers actually need to know about PbD in order for the concept to "work".

This means that currently PbD as a concept is not usable to communicate trust (or a lack of trust) in particular systems to European citizens and end users. This is a problem for policy-makers, end users, system integrators, system designers, researchers and, last but not least, for citizens themselves, because it keeps alive a sense of opaqueness and general distrust with regard to information systems[4].

There are no collective resources available, other than by example and/or by industry, as to what Privacy by Design actually means in a particular application domain. Nor is it known what relevant best practices are, what their consequences are, or which methodologies and tools are available. Innovation in protecting privacy is also hampered, because it is unclear to policy makers, citizens and purchasers how to compare innovations (Privacy Enhancing Technologies, PET) on their ability to mitigate privacy risks.

This leaves some questions, which we will attempt to address from a European perspective:

- How can PbD be made a useful concept? This also leads to the next question:
- Do we let each designing party (industry and technical parts of governments) decide per casus or product line what PbD means, or is there a need for some government involvement, for example to guard some definitions and to help define PbD per application domain?

This paper describes a vision for the second approach of the second question: how the general concept of privacy by design can be translated to specific application domains in a way that is transparent and practical. In this effort, progress is the goal, not perfection. The steps that are needed are

(1) Common understanding of the key concepts involved;
(2) Clear, workable definition of privacy by design;
(3) Set of tools and / or methods to give substance to privacy by design.

This vision paper follows this structure. The intended audience is quite broad: policy makers, system designers, legal and sociological scientists. This implies that we cannot go too deep into each topic, and that each type of audience may feel that too many basics have been included. This approach is however necessary to obtain common understanding through this medium.

2 Common Understanding of Key Concepts

The presence of common understanding in the key concepts is a requirement to build trust upon: privacy, trust, design and the life cycle of a system.

[4] We have accepted the risks that come with the heavy use of IT systems. That is not the same as trusting them.

2.1 Privacy

From the point of view of a designer, a sort of checklist to verify whether his system potentially violates the privacy of data subjects is desirable. However, privacy is a broad, abstract, and subjective concept and its meaning depends on context scope, and culture. As such, it is hard to define. These are some examples in recent literature and reference works [21, 22, 23, 24, 25], one of which we repeat here: privacy is the ability to control and limit physical, social, psychological and informational access to the self or one's group [26]. However, from these definitions we learn that privacy is considered a right, a freedom, a capacity, a claim and an ability. Apparently it is a concept that is hard to capture in a single complete definition. Langheinrich gives a short history of the concept [27], and illustrates the origination of five specific categories of privacy that together appear to encapsulate all previous definitions:

- Privacy of personal behaviour (media privacy);
- Privacy of territory (territorial privacy);
- Privacy of the person (bodily privacy);
- Privacy of personal communications (interception privacy), and
- Privacy of personal data (data or information privacy).

2.2 Privacy Invading Activities

From a (US) law perspective, Solove [28] describes a taxonomy of invasions of privacy, with a collection of activities that potentially interfere with one's privacy, grouped into four categories. To do this, he illustrates the concept of privacy with a very basic system design, see Fig. 1. Solove mentions the data subject and the data holder, and describes all encountered kinds of privacy invasions in those terms.

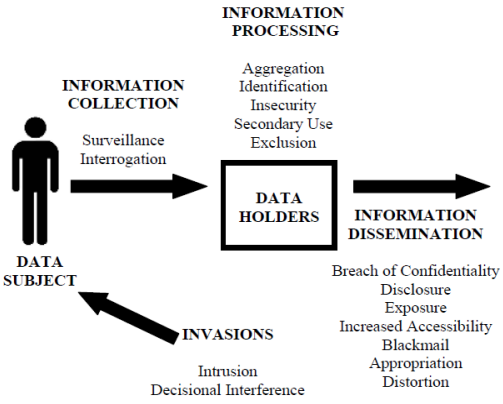


Fig. 1. - 16 potential privacy invasions (Solove, 2006)

There may be other models that are more extensive, specific or appropriate for the EU perspective. The expression of this model in information flows between entities and the completeness makes it accessible from the point of view of a designer. A model that separates all possible concerns involved helps as a checklist, which

makes it very practical and readily useful in a design process. The explicit mapping on the data subject and data holder of the potential privacy invasions can easily be used to map those invasions also into other views in system design such as behavioural (use case diagrams, sequence diagrams) and structural (composition, deployment) as can be done with UML [29].

The five categories that Langheinrich described can be mapped onto Solove's system design according to Table 1.

Table 1. -Mapping of Langhenrich's categories on Solove's privacy invasions

Solove	Langheinrich
Information collection: surveillance	Privacy of personal behaviour (media privacy); Privacy of personal communications (interception privacy)
Invasions: intrusion	Privacy of territory (territorial privacy); Privacy of the person (bodily privacy)
Information collection, Information processing and Information dissemination	Privacy of personal data

This short mental exercise suggests that a formal and complete definition for privacy may be possible.

2.3 Trust

The impact of applying privacy by design to an engineering process, should be a minimum "amount" of trust from data subjects in the new system with regard to the protection of their privacy. As Solove's model clearly illustrates, information flows one way: from the data subject to the data holder. Information is power, so there is fundamentally an imbalance of power in the relation between the data subject and the data holder which erodes trust. The process of privacy by design should be such that the interaction between all these parties leads to trust on all six points described below for the relation between data subject and data holder. This situation is complicated because in the design of systems there are more parties involved than just the data subject and the data holder: the user, designer, installer, maintainer, etc. The relation between the data subject and the data holder is the primary relation where trust should flow, but this depends on the (perceived) trust between the designer and data holder.

Trust may have even more definitions than privacy [30]. Just like in the virtual communities of Ridings et al [32], the data subjects, data holders and system designers involved in the design and use of systems generally do not converse directly with each other, so trust is at the generalized, collective level. Ridings et al asserts that trust consists of three factors: ability, benevolence and integrity. Ability is skills or competencies that enable an individual to have influence in a certain area. Benevolence is the expectation that others (i.e. trusted parties) will have a positive orientation or a desire to do good to the trustee. Integrity is the expectation that another will act in accordance with socially accepted standards of honesty or a set of principles that the trustor accepts, such as not telling a lie and providing reasonably verified information. In the context of PbD, these could be explained as described in Table 2.

Table 2. - Three factors of trust in the relations between data subject, data holder and designer

	Ability	Benevolence	Integrity
Data-subject➔data holder	The data holder has the skills to protect my privacy.	The data holder is concerned about my privacy.	The data holder acts in accordance to the written and unwritten rules w.r.t. my privacy.
Data-holder➔designer	The designer has the skills to design a system according to our needs.	The designer is concerned about the privacy impact of his design.	The designer acts in accordance to the design rules and best practices in this domain.

Ridings et al describe objective and measurable criteria to assess these different factors of trust in a particular community [32]. Let's focus on benevolence of the designer w.r.t. the privacy impact of his design. One criterion of trust might be "*who initiates the discussion of potential privacy risks of a particular future technology?*" Another criterion might be "*in which phase of the design are privacy concerns taken into account?*" In the public discussion around the use of UAV's for domestic surveillance, a privacy advocate might ask "*if even such an invasive tool can be made 'privacy by design' after it has been designed in the first place, what then, is the value of PbD?*" This is a clear sign of a lack of trust of the part of the privacy advocate in the designers of such systems. System designers and (future) data holders can address this issue by pro-actively communicating their worries about particular technologies, as head of Google Eric Schmidt recently did about drones [31].

2.4 The Law

Underlying these factors of trust, there is the legal basis, the way that a society has agreed upon to interact with each other. In Europe, the ECHR, Art.8 [32] addresses privacy. The data protection directive of 1995 is about personal data protection, which is a subset of privacy (e.g. excluding Solove's *invasions*). A directive must be integrated in the national legislative body of all EU Member States. Hence, organisations operating on the European markets face a diverse set of implementations of the European Data Protection Directive [34]. This has some disadvantages. It makes it more difficult for citizens and consumers to understand how their privacy is protected in other member states. It requires more investments by industry to tailor products and services for each member state, and multinational corporations are forced to incorporate multiple, perhaps conflicting, privacy policies within one organisation. Addressing these issues, the new Data Protection Regulation becomes immediately enforceable as law in all member states simultaneously without the necessity to be transposed into national law. This ends the numerous interpretations of the Directive in the member states.

In the domain of security, there will probably still be a Directive [35], not a Regulation. So in that domain, the EU KP7 SMART project [36] investigates some form of template structure for laws that deal with privacy-by-design. This might be helpful in keeping and building recognisability from the point of view of citizens, and therefore trust.

2.5 Design

Of the four definitions that the dictionary holds for "design", three are relevant for the notion of PbD[5]. The first is design as noun, i.e. *a plan or blueprint*. The second is design as a verb, i.e. *creating* the plan or blueprint and refers to the early phases of the life cycle of a system. The third is also design as a verb, but with the meaning *to intend*. [37] All three definitions are relevant because an intentional design process is necessary to keep and earn trust in the resulting blueprint and its application.

When the design process becomes complex or large in some aspect, it is often called *systems engineering*. This introduces the concept of the *system*. According to the International Council on Systems Engineering (INCOSE),

> *"a system is a construct or collection of different elements that together produce results not obtainable by the elements alone. The elements, or parts, can include people, hardware, software, facilities, policies, and documents; that is, all things required to produce systems-level results. The results include system level qualities, properties, characteristics, functions, behaviour and performance. The value added by the system as a whole, beyond that contributed independently by the parts, is primarily created by the relationship among the parts; that is, how they are interconnected."* [38]

The notion of a system also sets boundaries around it – its scope. This implies that PbD is limited to the boundaries of a system, and that these boundaries of the system should be described in order to be able to understand what the scope of PbD is in that particular instance. This hints that if a particular subsystem is designed according to the principles of PbD, whichever they may be, this subsystem can still be violating privacy laws when it is part of a broader system or connected to other systems[6]. The consequence is that a system that is designed according to the principles of PbD can still violate privacy laws when used improperly. Not only do we need to describe the limits of the system, we also need to describe and regulate its use.

2.6 Privacy and Process Models to Describe the Life Cycle of Systems

A process model for the life cycle of systems is a useful simplification of the inherently stochastic and sometimes unpredictable life cycle phases of a system. Such models range in detail, scope, track record and in flexibility, among other aspects. A good model for a design process covers all relevant phases of a system's life cycle,

[5] The fourth definition is *a pattern used to decorate something*. This could be useful for reducing the sense of invasiveness of a design [50].

[6] E.g., the apps on mobile phones and on social network sites can drastically alter the privacy impact of those systems.

and lets the designer and owner keep track of result, time and costs, in relation to requirements and constraints, even, or especially, in adverse circumstances, during those phases. A process model also serves as a template in the sense that it helps to methodologically follow each prescribed step. This enhances traceability, accountability and transparency in the design process. From the point of view of accountability, it is good practice to choose and apply a specific design process: you can verify whether the necessary steps have been taken, whether the transitions between the steps have been done under the right circumstances, and you can predict the next steps. The output of a phase is the start point for the next phase. So, if privacy requirements have been set at the beginning, then later in the process there should be a design where these privacy requirements have been taken into account. However, an actual design process is never flawless, so there may be mistakes in the resulting design.

There are many models for system life cycles, most of them originating in system engineering processes. Depending on the nature of a particular design challenge, a specific process-model is selected. For example, the Waterfall model is a simple and widely taught process model for designing systems. It consists of the phases Specification, Implementation, Integration, Test and Maintenance. However, the definition for PbD from the Commission itself references other phases in the life cycle of systems: the deployment, the use and the disposal phases. This disqualifies the Waterfall model for PbD because it does not acknowledge the importance of how the system is actually used, or of data disposal at the end of the life cycle of a system. Another issue with the Waterfall model is that it assumes that all stakeholders know what the problem is and on top of that, wastes no time deliberating different solution directions. The first phase of the Waterfall model is directly the requirements phase. The motivation for a particular solution direction stays implicit, and the Waterfall model is therefore not transparent. This is a risk for keeping and gaining trust. Another issue is that of repurposing a system. In a sense, scope creep is recycling, which can be beneficial from environmental and economic perspective. From the point of view of trust and transparency, we need methodological approaches to changing purposes of existing systems. The Waterfall model also lacks these.

There are other methods such as SIMILAR [39] and TOGAF [40] that do not ignore these phases of the life cycle of existing systems. The S of SIMILAR stands for State the Problem, and both SIMILAR and TOGAF have a cyclical structure, which addresses repurposing of systems. It goes too far to specify a particular design process model for PbD, but it is important to be clear about the process that is being followed, and to have a process that addresses the full life cycle of a system. A process model however, is an empty shell without the design patterns to apply it to.

2.7 Privacy and Design Patterns

With regard to PbD, the concept of design patterns is quite elegant. It does not mention implementation details, while at the same time describing the relevant aspects: problem, solution and consequences. A further, quite interesting property, is that it is for design patterns not necessary for an implementation to *exist*. This makes it

possible to also describe and register *future* technologies as design patterns. A design pattern can be considered good, or even *best practice* if it is agreed to have a particularly good track record.

Design patterns were first introduced in the domain of building architecture [41], but they became known to the information processing communities when they were introduced in IT and Object Oriented software design [42]. A design pattern is an abstraction of a design, in the sense that it is not concerned with implementation details. However, from the point of view of system design, design patterns exist at different levels of abstraction of the system. This can be illustrated with four abstractions of the same part of a house of which one function is to deliver privacy: an outer wall. A well placed wall can shield the data subject from (the feeling of) invasion, and can also actually prohibit information collection, as visualised in Fig. 2.

- A wall is a solid structure that separates outside from inside, on one floor-level; (physical view)
- A wall can be built with masonry; (design view)
- A wall can be used to shield against information collection; (use view)
- The strongest variety of this brickwork pattern to resist winds blowing straight into the wall, is *Flemish Bond* with *headers* every 5th row. (performance view)

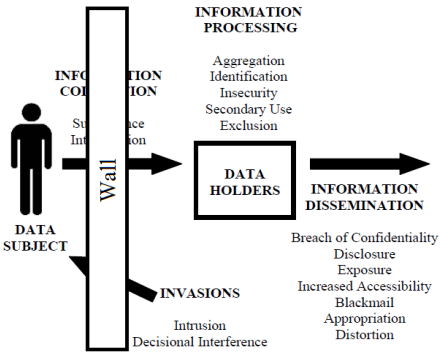


Fig. 2. A wall shields the data subject against invasions and information collection

As shown in this example, a design pattern can be described in different system aspects: the physical layout, the use case, the functional (as in information flow) design and all other system aspects or system views. Whether they are documented and intended as such is not relevant.

An example of a privacy design pattern that is about information processing is illustrated in Fig. 3. In this functional decomposition (not a process flow!) we point out that information and information processing that works at levels 0 (signal) and 1 (entity), by definition could contain personal data, while other layers should not, because they are by definition not concerned with individuals. Processing at level 2 (situation) determines "what" is happening. And finally, only processing at level 3 could make the decisions with regard to proportionality[7], because that is where the (potential) *impact* of the situation, and therefore also of (not) doing something is taken into account.

[7] Note that this design pattern does not say anything about whether processing is being done by an automated, or a human agent.

JDL Level	Example
Est Impact	Not secure
Impact assessment	Impact assessment
Est Situation	Known robber present
Situation assessment	Blacklisting
Est Entities	Faces, identity
Entity assessment	Face recognition
Est Signals	Images
Signal assessment	Video camera

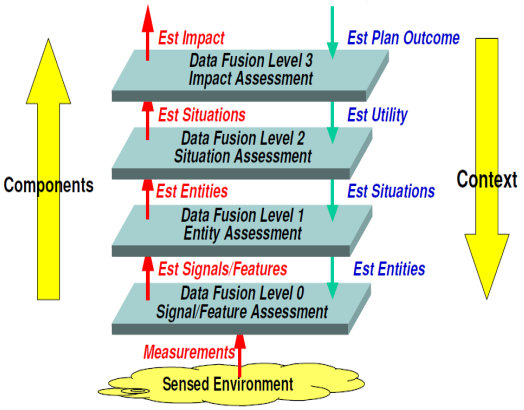


Fig. 3. -Data abstraction layers [43]- personal data resides at levels 0 (signal) and 1 (entity). The proportionality equation is done at level 3: impact assessment.

The level of abstraction of a design pattern determines whether a design pattern is relevant for privacy. This level of abstraction should be the same as the elements that are relevant for privacy: the data subject, the data holder and the privacy risks to be mitigated. In the example of the outer wall, the design view (how to build it) is not relevant from the point of view of privacy. The physical view as described is still incomplete because it still allows transparent walls, which would feel invasive and allow surveillance. So, to make the design pattern a *privacy design pattern*, it would have be changed as follows: the "privacy-wall" is a solid opaque structure that separates the data subject from a data holder, *specifically in terms of information collection*. *How* this particular type of wall implements the shielding against information collection, is not mentioned. A different design pattern could be envisioned that puts the wall around the data subject *and* the data holder instead, thereby allowing data collection, but preventing *information dissemination*.

3 Extended Definition of Privacy by Design

Based on the discussion of related concepts above, we propose the following extended definition of PbD:

> *The principle of 'Privacy by Design' envisions that privacy and data protective measures are operative throughout the entire life cycle of technologies: from the early design stage to their deployment, use and ultimate disposal. This is done by applying a design process that covers all life cycle stages and by applying privacy and data protection design patterns which are well understood and are the known best-practice for the particular purpose they are used for, and domain they are used in. The resulting design documents and systems should limit all the privacy invading activities to the minimum according to the foundational principles of privacy by design.*

A system designed according to this approach would score good on a PIA, and would be very easy to certify on any privacy norms. This definition links the definition as used in EU legislation with the foundational principles of PbD. It makes explicit what is expected from designers, and how traceability is organised through design methods.

3.1 Privacy Design Patterns

This definition requires agreement in a domain on what the respective best practices are for design patterns. Most design patterns that have been published so far are applicable in the domain of software engineering, and many security design patterns have also been published. Examples of a few very specific privacy design patterns can be found in [44]. A large collection –and structuring- of ICT Security design patterns is described in [45]. This paragraph introduces a more generic set of Privacy Design patterns that may be the start of a more complete set:

- privacy requirements patterns;
- anonimization and pseudonymization;
- hiding of personal data;
- data minimization;
- transparence, auditing and accounting patterns;
- informed consent.

Since for every engineering effort the system requirements are the starting point, we argue that *privacy requirements* should form an integral part of such system requirements. And since the expression of user needs in requirements in itself can be seen as a pattern, we propose the following first set of patterns that help describe the privacy requirements and select the appropriate set of controls that ensure (or make it likely) that the privacy requirements are met: *privacy needs identification*, *vulnerability assessment*, *privacy threat assessment*, *private information valuation*, *determination of privacy risks* and *selection of privacy controls*. The first five of these six patterns would combine to a Privacy Impact Assessment (PIA).

We continue with the description of a second set of privacy design patterns related to *Anonimization and Pseudonymization*. These are important and often recurring patterns in the privacy domain: they relate to the removal of the identity of the data subject in the data, or the replacement of the data subject's identity by an anonymous identifier: the pseudonym. These patterns can be recognized in many cases, e.g. the use of a pseudonym by a user of an internet forum, or the anonymization of data as is done in e.g. the voting process which is by law required to be completely anonymous. Patterns in this category are: *aggregation*, *k-anonymity*, *pseudonymous email*, *revocable privacy* [47], *blur (part of) image* [48, 49] and *decrease time resolution*. A third set of privacy design patterns can be found around the *hiding of personal data*, such as different types of *encrypt*ion, *batched routing* and *morphed representation*. The fourth set of design patterns we propose, revolve around *data minimization*. The idea here is to limit risks for disclosure of PII by limiting either the contents or the size of the data collection. The fifth set of design patterns has to do with methods that improve transparence and show – or even better: prove – to the data subject or to parties acting on his behalf, that due care has been taken: *logging*, *publication*, *peer review*

and *right of inspection by data subject*. The last set of privacy design patterns we discuss here, is a set of patterns around informed consent. These describe patterns that may be used to inform the data subject, and to get his or her consent for the particular purpose and method of private data processing and/or storage: *opt-in* and *opt-out*, *notification sign* and *privacy statement*.

3.2 Use Cases

In the article "Engineering Privacy by Design" [12] Gürses et al also identify the problem of vagueness of the definition of Privacy by Design, and they posit *data minimization* as a "*foundational first step to engineer systems in line with the principles of privacy by design*". From this perspective two case studies are described that show how privacy preserving techniques may be applied to design privacy friendly systems for two applications: (1) an e-petition system, where the identity of the data subject needs to be concealed but the transaction content needs to be disclosed, and (2) an electronic toll pricing (ETP) system where the identity is disclosed but the transaction data (being data subject locations) remain concealed.

Looking at these use cases, we can easily discern which of the privacy design patterns described above, have been applied here. For both of the use cases, we recognize the following design patterns: *identification of privacy needs, privacy threat assessment, private information valuation, vulnerability assessment, and determination of privacy risks*. These design patterns emerge in the analysis the authors have done in order to derive the privacy requirements.

For the first use case, that of the privacy-friendly e-petition system, we recognize the use of *encryption* in order to provide anonymity. Furthermore, *claims* are used ("verified credentials") instead of ordinary names or identifiers. The *publication* patterns is used to improve transparence. And finally, the authors employ *Layered Encryption* (via the use of the TOR network) in order to achieve anonymous communication.

The second use case, that of the privacy-friendly electronic toll pricing system, naturally uses a different set of design patterns because it has different privacy goals. In this use case, we see: *subsidiarity* (by choosing to not store all of the location data in one single database, but rather to leave it in the on-board units) and *cryptography,* through the cryptographic commitments.

We argue that, in addition to these design patterns, a more clear choice of the design methodology, and of the application of the other seven foundational principles of PbD would benefit the privacy of data subjects in these use cases. For example, the end-of-life phase is also ignored by Gürses et al.

3.3 Supporting Tool and Methods

In the areas of software engineering and of security, rich collections of design patterns already exist. Large online communities contribute to the descriptions of existing and new design patterns. Such communities also arise for privacy design patterns [46]. Published and reviewed collections of design patterns may help a system designer

determine which privacy design patterns (PDP) are best practice, and it may help data subjects to recognize the signs of good implementations of such design patterns. A good knowledge base contains a body of reference of PDP's and facilitates searching from different entries:

- privacy invading activities and their associated risks which are to be mitigated by a PDP, along the categories of Solove and Langheinrich, or those introduced by Van Lieshout [5], or by Cvrček and Matyás [51];
- real world instances (implementations) of the design pattern. This information would typically come from a Privacy Impact Assessment. Citizens and designers can inspect those implementations;
- consequences of applying the PDP. Citizens can recognize those consequences;
- types of systems that would benefit from applying PDP's, e.g. road pricing system, surveillance system, hospital information management system;
- manual's, documentation, publications and intellectual property that helps building the PDP's, while staying technology-, and therefore vendor-neutral;
- maturity of PDP's. It may have an indication of the *Technology Readiness Level*. Designers and procurers can choose the amount of risk they want to take with regard to new PDP's.

The tool could be implemented in the form of a website with a database behind it. This tool and methods will increase trust because it directly influences the six factors influencing trust that were introduced earlier:

- the *benevolence* of the data holder and the system designer is illustrated by how well they maintain their respective systems in the knowledge base;
- the *ability* of the data holder and the system designer is illustrated by their choice for a suitable design process and design choices w.r.t. PDP's;
- the *integrity* of the data holder and the system designer is illustrated by their participation in the knowledge base, and by following the best practices and design methodologies that are recommended for their respective domains.

4 Conclusions

Legislation makes extensive references to "privacy by design", but does not specify what it exactly means. This omission of a clear definition of PbD in the European landscape entails that from the point of view of European citizens, policy makers, local authorities and industry there is currently an unclear situation with regard to what the implications of a request for PbD actually entails. For example, it should be investigated whether the value of PbD can be improved by including the roles of designer or builder of a system in Data Protection Laws.

It would be elegant if individual and collective privacy interests could be aligned with economic interests. However, citizens currently do not flock to products and services that better protect their privacy, so there is no market mechanism favouring privacy protecting design. If we cannot leave it to economics to protect data subject's privacy, then we may need a more regulated approach. Already, the proposal for a

(new) privacy regulation creates a clear economic motivation, at least for enterprises, to address privacy risks. Privacy by Design could be the label of the method that describes how to mitigate such risks. At the same time, technology advances, the context changes and our knowledge improves. So, in addition to regulation, designing *privacy by design* should be an on-going, transparent dialogue between representatives of data subjects, data holders and system designers.

In order to facilitate this dialogue, some terminology must be defined more clear. The five categories of privacy that we encountered in literature can be mapped on Solove's system design which suggests that Solove's taxonomy is inclusive enough. This helps build trust in any methodologies and tools that are based on such a taxonomy.

Applying a design process model enhances traceability, accountability and transparency in the design process. An essential element of PbD should be, that the *privacy requirements* must be clearly and early stated as part of the functional requirements. Apart from serving as guide in the implementation /construction process, they also give guidance when testing the resulting system for compliance. The design process model should further cover the entire life cycle from problem statement to system disposal, including also the use-phase: a system that is designed in line with the principles of PbD can still violate privacy laws when used improperly.

In addition to common understanding of key concepts and the application of the right process model, we propose a more coherent approach to privacy design patterns (PDP). PDP's describe technology on an abstract level, and as such can also describe both existing, and not-yet-existing technology that enhances privacy of the data subject. We suggest that, when describing a privacy design pattern, the privacy invading activities (of Solove) that have to be mitigated are also mentioned, in other words: which privacy problem is solved. The community of PbD-designers could be supported with a knowledge base of such patterns.

There may be more foundational principles than the seven of Cavoukian, e.g. data minimisation could be another. An extended definition of PbD has been given that links the definition as used in EU legislation with the foundational principles of PbD. It makes explicit what is, in a certain domain, expected from designers, and how traceability is organised through design methods. This approach can be applied to all domains, as long as the PDP's are considered best practice in the respective domains. This implies that intrinsically intrusive domains such as surveillance can also benefit from a PbD-approach.

A tool and a set of use cases for that tool are envisioned to approach PbD with the help of privacy design patterns and design process models. Different types of stakeholders can address their interests with the tool in different design phases. It can be used by designers, citizens, DPAs and policy- and decision makers, among others. This tool would expand the level of transparency from *what data is being collected*, to *how are we protecting your privacy?*

The new privacy regulation introduces sanctions up to 2% of the annual turnover of enterprises. This elevates the importance of mitigation of privacy risks. This vision paper positions Privacy by Design as the mechanism to mitigate these privacy risks, and gives practical guidelines to *design Privacy by Design*.

Acknowledgements. The authors wish to thank Dr. John Borking, Dr. Jaap-Henk Hoepman, Sander van Oort and Johanneke Siljee for their thorough reviews.

References

1. EC, Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (1995)
2. EC, COM(2012) 11 (final) 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 (General Data Protection Regulation) (January 25, 2012)
3. EC, COM(2010) 609 (final), A comprehensive approach on personal data protection in the European Union (November 4, 2010)
4. EC, COM(2010) 245 (final)/2, A Digital Agenda for Europe (August 26, 2010)
5. van Lieshout, M., Kool, L., van Schoonhoven, B., de Jonge, M.: Privacy by Design: an alternative to existing practice in safeguarding privacy. Info. 13(6), 55–68 (2011)
6. European Forum for Urban Security, Charter for a Democratic Use of Video Surveillance (2011)
7. Hes, R., Borking, J.: Privacy Enhancing Technologies: the path to anonymity (Revised Edition) Registratiekamer, Achtergrondstudies en Verkenningen 11 (first edition 1995)
8. CFP2000, Conference on Computers, Freedom & Privacy (2000), http://www.cfp2000.org/
9. EC / TNO et al, FP5, PISA project (2003), http://cordis.europa.eu/projects/rcn/53640_en.html (accessed June 2, 2012)
10. Cavoukian, Origins of Privacy by Design, http://privacybydesign.ca/publications/pbd-origin-and-evolution/ (accessed August 3, 2011)
11. Cavoukian, Privacy by Design – The 7 foundational principles (August 2009) (revised January 2011)
12. Gürses, Troncose, Diaz: Engineering Privacy by Design. In: Conference on Computers, Privacy & Data protection, CPDP (2011)
13. Jean-Philippe Courtois, Privacy by Design at Microsoft (November 29, 2010)
14. Winterfield, K. (2009), http://ibmresearchnews.blogspot.com/2009/10/inventors-corner-innovations-enable.html
15. Cavoukian, Privacy by Design – The answer to overcoming negative externalities arising from poor management of personal data, Trust Economics Workshop (June 23, 2009)
16. Kranzberg, M.: Technology and History: Kranzberg's Laws. Technology and Culture 27(3), 544–560 (1986)
17. EuroPrise - the European Privacy Seal for IT Products and IT-Based Services (2007), https://www.european-privacy-seal.eu/ (accessed June 2, 2012)
18. London Economics, Study on the economic benefits of privacy-enhancing technologies (PETs) (July 2010)
19. Borking, J.: Privacy law is code (2010)
20. Rogers, E.M.: Diffusion of Innovations (1962)

21. Warren and Brandeis, Harvard Law Review. The right to privacy, vol. IV(5) (December 15, 1890), http://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html
22. Agre & Rottenberg, Technology and privacy: the new landscape (1997)
23. Clarke, R.: Roger Clarke's 'What's Privacy?, http://www.rogerclarke.com/DV/Privacy.html (accessed May 12, 2011)
24. Cambridge Essential English Dictionary, lemma Privacy (accessed August 6, 2011)
25. Westin, A.: Privacy and Freedom. Atheneum, New York (1967)
26. Burgoon, K., Parrott, R., Le Poire, B.A., Kelley, D.L., Walther, J.B., Perry, D.: Maintaining and Restoring Privacy through Communication in Different Types of Relationships. Journal of Social and Personal Relationships 6(2), 131–158 (1989)
27. Langheinrich, M.: Privacy by design - principles of privacy-aware ubiquitous systems. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) UbiComp 2001. LNCS, vol. 2201, pp. 273–291. Springer, Heidelberg (2001)
28. Solove, D.J.: A Taxonomy of Privacy. University of Pennsylvania Law Review 154(3), 477–564 (2006)
29. UML 2.4.1 Specification, http://www.omg.org/spec/UML/2.4.1/ (accessed December 2011)
30. Harrison McKnight, D., Chervany, N.L.: The Meanings of Trust, University of Minnesota (1996), http://www.misrc.umn.edu/wpaper/wp96-04.htm
31. BBC, Eric Schmidt, Google (April 13, 2013), http://www.bbc.co.uk/news/technology-22134898
32. Ridings, C.M., Gefen, D., Arinze, B.: Some antecedents and effects of trust in virtual communities. The Journal of Strategic Information Systems 11(3-4), 271–295 (2002) ISSN 0963-8687, 10.1016/S0963-8687(02)00021-5
33. Article 8 of the European Convention on Human Rights (1950)
34. EC, undated, Status of implementation of Directive 95/46 on the Protection of Individuals with regard to the Processing of Personal Data (2011), http://ec.europa.eu/justice/policies/privacy/law/implementation_en.htm (accessed August 3, 2011)
35. EC, COM/2012/010 final - 2012/0010 (COD), Proposal for a Directive of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data (January 2012)
36. EU KP7 SMART project, http://www.smartsurveillance.eu/ (accessed May 13, 2012)
37. Cambridge Essential English Dictionary, lemma Design (accessed August 28, 2011)
38. INCOSE, A Consensus of the INCOSE Fellows, http://www.incose.org/practice/fellowsconsensus.aspx (accessed June 2012)
39. Bahill, A.T., Gissing, B.: Re-evaluating systems engineering concepts using systems thinking. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 28(4), 516–527 (1998)
40. The Open Group, "The Open Group Architecture Framework, TOGAF", http://www.opengroup.org/togaf/ (last accessed April 2, 2012)
41. Alexander, C.: A Pattern Language: Towns, Buildings, Construction (1977)

42. Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P.: Pattern-Oriented Software Architecture. A System of Patterns, vol. 1. John Wiley & Sons (1996)
43. Steinberg, A., Bowman, C.: Rethinking the JDL Data Fusion Levels, NSSDF JHAPL, June, 04 2. In: Bowman, C.L. (ed.) The Dual Node Network (DNN) Data Fusion & Resource Management (DF&RM) Architecture, AIAA Intelligent Systems Conference, Chicago, September 20-22 (2004)
44. Hafiz, M.: A collection of Privacy Design Patterns. In: Proceedings of the 13th Pattern Languages of Programs. Allerton, Illinois (2006)
45. Security Patterns – Integrating Security and Systems Engineering, Schumacher, Fernandez-Buglioni, Hybertson, Buschmann, Sommerlead. John Wiley & Sons (2006)
46. UC Berkeley School of Information (2013), http://privacypatterns.org/ (last visited May 2013)
47. Revocable Privacy, Jaap-Henk Hoepman. Privacy & Informatie 11(3), 114–118 (June 2008)
48. BSIA, Privacy Masking Guide (2011)
49. Roelofsen, Patent WO 03/010728/A1 Method and System and Data Source for Processing of Image Data (February 2003)
50. WeArePerspective (2007), http://www.weareperspective.com/project/ns-camera (accessed December 2011)
51. Cvrček, D., Matyáš, V.: D13.1: Identity and impact of privacy enhancing technology. FIDIS (2007), http://fidis-wp13-del13.1.final.pdf (accessed February 16, 2011)

Enhancing Privacy by Design from a Developer's Perspective

Christoph Bier[1], Pascal Birnstill[1], Erik Krempel[1], Hauke Vagts[1,2], and Jürgen Beyerer[1,2]

[1] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe, Germany
[2] Vision and Fusion Laboratory, Karlsruhe Institute of Technology, Germany
{christoph.bier,pascal.birnstill,erik.krempel,hauke.vagts,juergen.beyerer}@iosb.fraunhofer.de

Abstract. This work proposes a pragmatic approach towards refining as well as complementing Ann Cavoukian's seven principles of Privacy by Design. In an analysis of the principles' definitions, practical handicaps as well as essential complementary claims are pointed out. Based on these insights the authors come up with a more consistent and pragmatic definition of Privacy by Design governed by seven requirements.

A practical application of this new definition of Privacy by Design is demonstrated by means of an example scenario in the context of video surveillance. It is shown that by applying the principles of the new definition to the redesign process of a conventional surveillance system, a significantly more privacy-aware intelligent surveillance system can be be obtained.

1 Introduction

Ann Cavoukian's seven principles of *Privacy by Design (PbD)* are as well recognized as appreciated in the privacy community. Nevertheless, according to their rather abstract nature, these principles are often hard to apply to practical problems. Incorporating their practical experiences, the authors come up with an in some aspects slightly more relaxed, but in total more comprehensive and practically applicable refinement of PbD.

This work is organized as follows. Section 2 discusses Ann Cavoukian's principles of PbD. Section 3 summarizes related work on privacy paradigms. The authors come up with a new definition of the seven principles of PbD in Sect. 4. In Sect. 5 the new principles are applied to video surveillance, before concluding in Sect. 6.

2 Discussion of Privacy by Design Principles

In the following paragraphs, this work analyzes the seven principles of PbD by Ann Cavoukian [1] with regard to consistency and practical applicability.

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 73–85, 2014.

2.1 Proactive Not Reactive – Preventive Not Remedial

The first principle of PbD describes the process character of designing privacy-preserving systems, aiming at proactively eliminating privacy risks.

> "The Privacy by Design approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy invasive events before they happen. PbD does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred – it aims to prevent them from occurring. In short, Privacy by Design comes before-the-fact, not after."

The authors completely agree to the basic rationale behind this definition, i.e., dealing with threats to privacy at the earliest possible stage of a design process. However, technical progress often introduces new threats to formerly secure and privacy-preserving systems, i.e., by highly creative inventions that could not have been anticipated at the time the system was designed. Advances in cryptanalysis can significantly weaken a previous state-of-the-art encryption algorithm. Therefore, even though the principle is **consistent**, the authors claim that it is not sufficient.

Due to the process nature of this principle, it is notoriously difficult to measure to which extent it is fulfilled by a given system design. Therefore, only direct comparison is possible but unsatisfying. Given two functionally similar system designs, a qualitative head-to-head comparison can be carried out. However, only if two designs provide exactly the same functionality and system design A uses a real subset of personal information of the competing system design B, the result will be instructive.

The proactive approach is easy to understand but hard to apply. Although a privacy impact assessment (PIA) is a powerful method to identify potential privacy risks, expert know-how is needed to resolve the problems.

2.2 Privacy as the Default Setting

> "We can all be certain of one thing – the default rules! Privacy by Design seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy still remains intact. No action is required on the part of the individual to protect their privacy – it is built into the system, by default."

Ann Cavoukian describes *Privacy as the Default Setting* as a property of a system or process, which ensures maximum privacy as long as not requested otherwise by the individual. Non-mandatory system functionality that is in need of additional collecting or additional processing of personal data must be disabled by default.

From the authors' point of view the above definition is not satisfying regarding its **consistency**.

Users of real world systems are not likely to sustainably restrict themselves to a stripped-down default operational mode of a given system. Therefore PbD-compliant systems have to be designed in a way that privacy is not "knocked out" as soon as a user invokes a non-mandatory system function.

It needs to be stressed that the system must fulfill this requirement independent of a user's informed consent. Consequently, a PbD-compliant system must not provide inappropriately privacy-invasive functionality in any operational mode. In other words, each subsystem or functionality must be designed according to the principles of PbD, regardless of whether or not it is available in the default operational mode of the overall system.

In terms of measurability, operational modes of systems of comparable functionality can be qualitatively compared to each other. However, approaches on evaluating real world systems against abstract system models seems more promising, yet to this day is a topic of basic research.

Finally, the principle of privacy as the default setting can be operationalized by means of incorporating requirements such as enforcement mechanisms for purpose binding of collection and processing of personal data as well as the regular storage time of collected personal data at an early stage.

2.3 Privacy Embedded into Design

> "Privacy by Design is embedded into the design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality."

With the principle of *Privacy Embedded into Design* Ann Cavoukian postulates that privacy aspects and functional aspects should be addressed with equal priority when designing a new system or process from scratch. Moreover, Cavoukian claims that privacy measures must not be integrated on the cost of functionality.

The authors question the **consistency** of this definition as it is not satisfiable in any case. The idea of a system or of any innovation is usually born out of some kind of demand, be it missing functionality or poor usability of an existing system or process, or be it the general need of a method or technology that facilitates some recurring task. Due to the nature of innovation, functional requirements usually evolve prior to privacy aspects being considered at all. Thus, privacy can be integral to a system, but not necessarily without diminishing its functionality.

The extent to which the principle of privacy embedded into design is fulfilled by a given design can be evaluated qualitatively by means of carrying out a PIA. However, according to the process character of this principle, formal metrics for quantitative analysis are not conceivable.

Eventually, the principle of privacy embedded into design can be operationalized by integrating PIAs into each stage of the design and development process. The outcome of a PIA is always fed back into a further iteration or the next stage of the process.

2.4 Full Functionality – Positive-Sum, Not Zero-Sum

The fourth principle addresses the compatibility of privacy and functionality.

> "Privacy by Design seeks to accommodate all legitimate interests and objectives in a positive-sum 'win-win' manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. Privacy by Design avoids the pretence of false dichotomies, such as privacy vs. security, demonstrating that it is possible, and far more desirable, to have both."

Figure 1 illustrates the three different concepts in game theory needed for understanding Cavoukian's definitions as well as the revised version in Sect. 4. Depending on a starting point (dots in Fig. 1), functionality F and privacy P can evolve. Changes are denoted with ΔP and ΔF respectively.

- zero-sum: A concept in the field of game theory in which the sum of the outcomes is equal to zero (see Fig. 1a), i.e., a positive ΔP results in a negative ΔF with the same quantity and vice versa.
- positive-sum: A concept in the field of game theory in which the sum of the outcomes is greater zero (see Fig. 1b), i.e., either ΔP or ΔF can be negative, but the sum is positive.
- win-win: A special case of a positive-sum game where it is necessary that every participant has an outcome greater zero, i.e., privacy and functionality increases (see Fig. 1c).

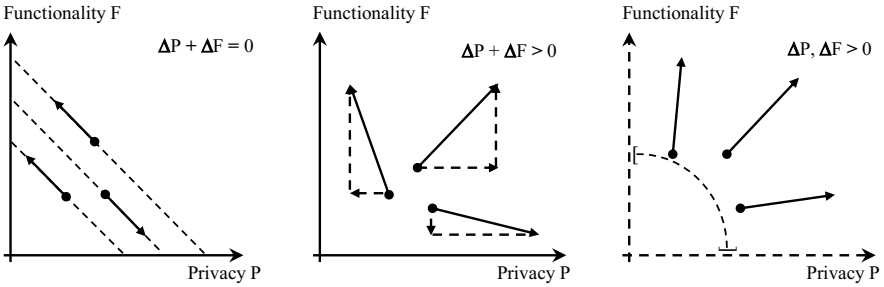
In other words, fulfilling this principle of PbD requires that a given system's functionality must only be extended if at the same time the systems privacy-awareness is improved. This requirement inhibits any kind of pragmatic trade-offs, i.e. neither tolerating a minor cut of functionality for significantly improved privacy nor sacrificing a bit of privacy for undoubtedly beneficial functionality is possible.

The principle is also **inconsistent** from a theoretical point of view. Starting with a situation of full privacy and zero functionality, adding functionality that requires personal information necessarily reduces privacy. Generally speaking, there have to be trade-offs between functionality and privacy in some cases.

In order to assess whether a new design results in a win-win, positive-sum or zero-sum situation, the degrees of privacy and functionality have to be measured. For this, privacy as well as functionality requirements have to be prioritized and weighed against each other. After determining to which fraction the requirements are actually fulfilled by a given design, the weighted sums over the fractions of privacy and functionality fulfillment can be calculated. From an operational point of view, however, weighting of requirements is a controversial issue and developing a method for objectively resolving conflicting requirements is an open research question.

2.5 End-to-End Security – Full Lifecycle Protection

> "Privacy by Design, having been embedded into the system prior to the first element of information being collected, extends securely throughout



(a) Zero-Sum (b) Positive-Sum (c) Win-Win

Fig. 1. Comparing zero-sum, positive-sum and win-win

> the entire lifecycle of the data involved – strong security measures are essential to privacy, from start to finish. This ensures that all data are securely retained, and then securely destroyed at the end of the process, in a timely fashion. Thus, Privacy by Design ensures cradle to grave, secure lifecycle management of information, end-to-end."

Adequate security mechanisms are essential to protect personal information and have to be incorporated into the design process from the beginning. As this also applies for privacy, both privacy and security measures have to go hand in hand during the whole lifecycle. Only then, **consistency** can be maintained and over-regulation can be avoided.

Formal models exist to prove the security of low-level building blocks. However, measuring the security of complex system is a challenging task. Using wrong attacker models or using faulty implementations of protocols can put personal information at risk, even when the system seems to be secure.

Standards for certification, e.g., Common Criteria, can be used to specify security requirements, verify their fulfillment and thus make security operational.

Security has to be considered as a process, thus technical specification is not sufficient. People's security awareness must be brought forward; roles and responsibilities must be defined and deployed.

2.6 Visibility and Transparency – Keep It Open

Knowing what is going on with one's personal data is a fundamental precondition for using one's legal right to informational self-determination. From a user's perspective, without transparency any other privacy measures do not take effect.

> "Privacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to both users and providers alike. Remember, trust but verify!"

Transparency can be achieved via audits, notifications and information. A basic requirement for trust and verification is the definition of consistent privacy compliance rules. A precondition for verifiable information is the integrity of the information chain. However, when considering transparency isolated from other privacy requirements, one neglects that privacy requirements are not without mutual conflicts, e.g., transparency has to be weighed against unlinkability.

How to measure transparency depends on the considered aspect, e.g., a transparency mechanism for a data storage can be judged by comparing a user's ability to find out where personal data is located prior to and after implementation.

Only a broad management approach allows to align transparency with the other data protection targets. Audits can support organizational measures like appointing a data protection officer.

2.7 Respect for User Privacy – Keep It Individual and User-Centric

The last principle focuses on the relationship between the user and the system.

> "Above all, Privacy by Design requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. Keep it user-centric!"

The last element of PbD is not an autonomous principle, rather it is a reminder that the whole privacy process is in fact about the user and that a PbD compliant system should be designed in this maxim. Only if privacy features are easy to use and user centric, they can fulfill their task.

This also means that the needs of different users have to be considered. While consent and configuration empower the user, general data avoidance is a paternalistic, illiberal concept. It is fair to offer data avoidance as an option for the user. But to set data avoidance as a default and unchangeable setting, refuses the user the right to information self-determination. Hence, data minimization (with respect to the specified purpose) and not data avoidance should be the guideline for privacy by design.

Multiple approaches allow to measure user acceptance of a system. Doll et al. [2] and Tullis et al. [3] show how to measure user satisfaction and experience.

As User-Centric Privacy is a transversal principle, activities like user surveys and usability walkthoughts are embedded into the design of all privacy-relevant components.

3 Related Privacy Principles

Besides the discussed PbD Principles by Ann Cavoukian [1], who introduced the term over ten years ago [4], Langheinrich [5] was one of the first researchers who focussed on the application of a PbD framework. His six principles for guiding system design are based on the Fair Information Principles (FIP) by the OECD [6]. His playground for a first case study was the domain of ubiquitous computing.

A couple of other examples for important application areas for PbD are given by Schaar [7] in 2010.

Spiekermann and Cranor [8] take a more user-centric approach. Based on a three-layer model of user privacy concerns they developed a "privacy-by-policy" and "privacy-by-architecture" design method. Both approaches are limited to specific privacy requirements, namely notice, choice and data minimization. Guerses et al. [9] criticise the missing distinctiveness of the PbD principles. They propose to start from data avoidance as the best and first step towards PbD. Their worry is that starting from a specified purpose and considering different privacy goals afterwards will lead to unlimited processing of personal data by stepwise extension of the purpose.

Based on the foregoing national and international discussions [10], Rost and Bock [11] propose six data protection targets. As they do not describe an overall design meta-process, the authors incorporate them as a mandatory but not sufficient building block of the PbD approach presented in this work.

4 A New Definition for Privacy by Design

Ann Cavoukian phrases PbD as an attitude rather than a meta-process and thus does not provide a strict definition and concrete requirements. As other PbD definitions and standards are either selective or incomplete, the authors redraft PbD in a more concrete and applicable fashion and add missing requirements.

Definition 1 (Privacy by Design). *PbD is a property of a meta-process for designing information technology, physical technology, business processes or organizational structures (target of design). Such a meta-process fulfills PbD if the requirements 1 to 7 hold. The final target of design is compliant to PbD if it is developed under permanent application of a design meta-process fulfilling PbD.*

PbD cannot assure that a target of design (ToD) is not privacy invasive. However, if PbD is applied properly the implementation of the ToD is as little privacy invasive as possible.

As a basis for the following requirements and as a precondition for legal considerations, the person or committee in charge has to define the purpose of the ToD explicitly. This is expressed in the first requirement:

Requirement 1 (Purpose Specification). *For each target of design, its concrete purpose has to be specified prior to the design process. The purpose must not be changed during the design process or the operational phase. When the purpose of a targed of design needs to be extended or changed, an entirely new design process has to be initiated. Functionality that is specified during the design phase must appropriate to the specified purpose.*

The second requirement rephrases the "Proactive not Reactive"-Principle:

Requirement 2 (Preventive Privacy Protecting Action). *Preventive Privacy Protecting Action holds, if there is no privacy invasive event in the target*

of design without a preventive privacy protecting action. A preventive privacy protecting action is triggered whenever a privacy invasive event occurs. In case the event is not mandatory for the functionality of the target of design, the preventive action has to inhibit the privacy invasive event. Else, the action has to limit the impact of the event to a minimum.

The third requirement incorporates the common principle of limiting collection, use, retention and disclosure of personal information into the design process.

Requirement 3 (Privacy as the Default Setting). *For the default operational mode of a target of design the following principles shall apply:*

- *Non-mandatory but privacy-relevant features or sub-processes must be excluded in a way that the system is reduced to its purely functional purpose plus non privacy-relevant features or subprocesses.*
- *Only the minimal necessary amount of personal data shall be collected by a system or during a process.*
- *The usage of this data shall be bound to the purely functional purpose of the system or process.*
- *The storage time of the data shall be well-defined beforehand.*
- *In addition to this, the latter three of the previously described principles shall also be applied to each additional feature or subprocess that can be activated by and with the informed consent of a user.*

The fourth requirement stresses that Privacy Embedded into Design requires that privacy is hierarchically nested into each component of the ToD.

The consideration of a more pragmatic regulation for trade-offs between functionality and privacy during a concrete design process requires a new definition of the Positive-Sum requirement. Designing a new ToD does not always start from scratch, i.e., privacy invasive features are often created by adding new functionality to an existing system. Thus, the new definition also has to be applicable as a guideline for the evolution of a system.

Requirement 4 (Positive-Sum Privacy). *Positive-Sum Privacy consists of a starting point (see Fig. 1), an evolutionary step and an assessment of the method:*

- *Starting point of a comparative evaluation of a target of design is an outdated predecessor with less than full functionality and less than full privacy, both greater than zero.*
- *In an evolutionary step, a trade-off between privacy and functionality is allowed if and only if it results in a positive-sum of functionality and privacy.*
- *The positive-sum has to be clear and not based on biased evaluation methods. If there is reasonable doubt, Positive-Sum Privacy is not fulfilled.*

In order to account different aspects of privacy and as a requirement to assess Positive-Sum Privacy, general data protection targets are needed. Such targets also assure a broader view on privacy issues.

Requirement 5 (Consider the Data Protection Targets). *Data protection targets are the main guideline for all design decisions regarding privacy and privacy management. For each major design decision, the consideration of the data protection targets: Confidentiality, Integrity, Availability, Transparency, Unlinkability, Intervenability has to be documented.*

Confer ISO 27001 and related standards for the definitios of Confidentiality, Integrity, and Availability. Integrity additionally incorporates accuracy of personal information.

Transparency represents the traceability of collection, usage and dissemination of personal data. Unlinkability refers to the limitation of use to well-defined circumstances, i.e., the specified purpose. Intervenability addresses the ability of the person responsible to control a system or process. Additionally, the person affected itself should be enabled to interact with the ToD.

Requirement 6 (Full Life Cycle Protection). *All security and privacy measures have to be in place before any personal information is processed by the target of design. The measures have to stay active until the last piece of personal information is deleted or has left the target of design.*

As privacy is an individual matter, it is difficult to standardize. Nevertheless PbD must account for the needs of persons affected by the ToD.

Requirement 7 (Individual Centric Privacy). *The target of design should be configurable as far as possible. Potential alternatives for implementing a specific functionality or privacy feature shall be interchangeable in a modular way. The configuration of the target of design shall be adaptable for each individual affected.*

For ease of use and practical reasons, the ToD should offer configuration templates. An individual configuration might also contain a more privacy invasive functionality. By this means, consent is a subset of Individual Centric Privacy.

5 Example: Intelligent Video Surveillance

The subsequent paragraphs illustrate the rephrased definition of PbD in the context of an intelligent privacy-aware video surveillance system.

As a starting point assume an airport being monitored using a conventional video surveillance system, which is supposed to be replaced for efficiency reasons. The purpose of the video surveillance system is to observe critical infrastructures of the airport, i.e., regions of the airport that must not be accessed by unauthorized persons. The system is also used for manual tracking of intruders, thus its cameras do already cover the airport to a great extent. The security personnel is faced by a large number of screens that show the live video streams. Furthermore, video data of all cameras is stored in an archive for manual investigations.

Video surveillance is often criticized as an unselective measure putting people under general suspicion. Modernizing video surveillance due to efficiency needs is an opportunity for carrying out a PbD compliant redesign.

5.1 Purpose Specification

The overall purpose of the considered surveillance measure is enabling security personnel to observe critical regions of the airport more efficiently, i.e., intrusions into critical regions have to be detected autonomously, so that an operator can concentrate on handling incidents. With regard to incident handling, the purpose of the system is also to assist at tracking intruders across several cameras. Finally, offline analysis of reported incidents must be facilitated, i.e., for hindsight investigations of the police the video archive must provide a search function for specified persons. This purpose specification must be adhered during the whole design phase of the system as well as at any instant in time of its usage. As a consequence, the final system must prevent reuse in illegitimate ways, e.g., observing employees by configuring their workplaces as critical regions.

5.2 Preventive Privacy Protecting Action

The baseline requirement of detecing intrusions into critical regions can be fulfilled in a rather noninvasive fashion. Person detector algorithms are only running on specific cameras that cover critical regions and neither the cameras' live streams nor the video archive is accessible in the default operational mode. A person detector virtually transforms certain events into alarms. Thus no personal data is stored and no *preventive privacy protecting action* regarding the system design itself is required.

Tracking intruders across various cameras is clearly more privacy invasive. As a *preventive privacy protecting action*, enabling this function requires the system to be put into an *alert mode* with extended logging of the operator's interactions.

Regarding the video search function, a first preventive privacy protecting action can be similar, i.e., the function is inhibited in the default operational mode and is only available in a dedicated *hindsight investigation mode* with extended logging and additional security measures.

As soon as the search function delivers results, *preventive privacy protecting actions* must be in place to ensure that these items of personal data do not leave the system and must not be used for any other than the intended purpose. This can be implemented by a usage controlled infrastructure as well as integrating policy enforcement mechanisms into system components that are capable of inhibiting and modifying system actions [12].

5.3 Privacy as the Default Setting

Privacy as the default setting can be applied to the workflow of intelligent video surveillance systems in a relatively straight forward fashion. The default operational mode does not offer any privacy-invasive functionality, i.e., neither person tracking nor searching in the video archive. Regarding the video archive, there must be a default period of time after which video data is regularly deleted (e.g., after 72 hours) if not assigned to criminal investigation proceedings. Only when an intrusion into a critical area has been detected, the operator is asked to assess the situation and to decide whether to put the system into the alert state or not.

5.4 Positive-Sum Privacy

Paragraph 5.2 already explained how system functionality is being separated into a rather unselective, but less privacy-invasive default operational mode, i.e., intrusion detection for critical regions, and a highly invasive yet highly selective alert mode, i.e., tracking or locating intruders throughout the airport.

This design is compliant to *Positive-Sum Privacy* as in total the system is less privacy invasive and only in highly selective situations trades privacy for valuable functionality.

By adding signs which announce the presence of a video surveillance system, the overall transparency and thereby the privacy protection of the system is improved. This change improves the privacy of the system without improving its functionality. Thus, it does not comply with the rather strict PbD Win-Win principle. In the *Positive-Sum Privacy* requirement such changes are possible and highly desirable.

5.5 Data Protection Targets

The first subset of the data protection targets, namely *confidentiality*, *integrity*, and *availability*, can be subsumed under the heading of *data security targets*. Although challenging, these are well investigated and concepts from the areas of cryptography, access control and redundancy can be applied. This work focusses on the second subset of the data protection targets, which can be considered as *data privacy targets*: *transparency*, *unlinkability*, and *intervenability*.

In [13], Vagts proposes a more *transparent* system where in addition to the mandatory notification signs, people can interact with the video surveillance system using their smart phones. By this means people can easily access more detailed information such as who is the operator in charge for the system, what kinds of data are being collected and processed by the system, for which purpose, and how long is data regularly stored.

Regarding an intelligent video surveillance system, it is reasonable to claim that any piece of collected personal data must be tied to the purpose of the surveillance measure. By incorporating a usage control infrastructure [12], *unlinkability* can be achieved by means of inhibiting the association of collected data with external databases.

Regarding the data protection target of *intervenability*, video surveillance systems constitute a special case of data processing systems that generally collect and process personal data without an individual's explicit consent. Stored data also has to be deleted after a relatively short period of time, e.g., 72 hours, if it has not been tagged as subject to criminal investigation proceedings. Hence, from the point of view of individuals affected by a video surveillance measure, intervenability in the sense of individuals' legal rights to request deletion or correction of their personal data is neither realistic nor necessary.

However, when conceiving the security personnel operating the video surveillance system as users, then intervenability can be seen as enabling operators to be in control of the system. An intelligent video surveillance system must by

no means be enabled to take automated decisions. Referring to the underlying scenario, if the intelligent video surveillance system detects a person entering a critical region, the system must not automatically start tracking this person, but rather incorporate the operator into the assessment and handling of the situation.

5.6 Full Life Cycle Protection

When applying the paradigm of *full life cycle protection* to the context of intelligent video surveillance systems, it is sufficient to claim that the system must not be run if only a single privacy or security measure is faulty or inactive. According to section 4, each privacy or security measure is a preventive privacy protecting action of some other privacy-invasive system functionality, otherwise a measure would be unnecessary.

5.7 Individual Centric Privacy

Intelligent video surveillance systems constitute a special case where the system's user is its operator rather than the affected individuals whose personal data are at risk. However in contrast to the data protection target of intervenability, the paradigm of *individual centric privacy* should not be projected to the operator, who has to be considered as a potential attacker at the same time. Providing the operator with too many options for configuring the system will most likely result in a system that is vulnerable to abuse.

6 Conclusion

Based on existing notions of PbD, in particular Ann Cavoukian's original formulation, this work proposed a more consistent, applicable and comprehensive definition of PbD. This definition should be understood as a meta-process with explicit requirements. This meta-process shall serve as a guideline for system or process designers tackling the controversial relationship between innovative functionality and state-of-the-art privacy protection. By means of applying the new definition of PbD to an example in the context of video surveillance the authors demonstrate their perspective on privacy engineering in modern information systems.

Acknowledgment. This work was partially funded by Fraunhofer Gesellschaft Internal Programs, Attract 692166, the KASTEL project by the Federal Ministry of Education and Research, BMBF 01BY1172 and the SURVEILLE project in the 7th Framework Programme by the European Commision (Project reference: 284725). The views expressed are those of the authors alone and not intended to reflect those of the Commision.

References

1. Cavoukian, A.: Privacy by Design - The 7 Foundational Principles (2011)
2. Doll, W., Torkzadeh, G.: The measurement of end-user computing satisfaction. MIS Quarterly, 259–274 (1988)
3. Tullis, T., Albert, W.: Measuring the user experience: collecting, analyzing, and presenting usability metrics. Morgan Kaufmann (2008)
4. Cavoukian, A.: Privacy by Design (2009)
5. Langheinrich, M.: Privacy by Design - Principles of Privacy-Aware Ubiquitous Systems. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) UbiComp 2001. LNCS, vol. 2201, pp. 273–291. Springer, Heidelberg (2001)
6. OECD: Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Organisation for Economic Cooperation and Development (1980)
7. Schaar, P.: Privacy by Design. Identity in the Information Society 3(2), 267–274 (2010)
8. Spiekermann, S., Cranor, L.F.: Engineering Privacy. IEEE Transactions on Software Engineering 35(1), 67–82 (2009)
9. Gürses, S., Troncoso, C., Diaz, C.: Engineering Privacy by Design. In: Proceedings of the 4th International Conference on Computers, Privacy & Data Protection, Brüssel (August 2011)
10. The public voice: The Madrid Privacy Declaration – Global Privacy Standards for a Global World (2010)
11. Rost, M., Bock, K.: Privacy By Design und die Neuen Schutzziele. DuD 35(1), 30–35 (2011)
12. Pretschner, A., Hilty, M., Basin, D.: Distributed Usage Control. Communications of the ACM 49(9), 39–44 (2006)
13. Vagts, H., Beyerer, J.: Enhancing the acceptance of technology for civil security and surveillance by using privacy enhancing technology. In: Elsner, P. (ed.) Future Security: 6th Security Research Conference. Fraunhofer Verlag (2011)

A Solution, But Not a Panacea for Defending Privacy: The Challenges, Criticism and Limitations of Privacy by Design

Demetrius Klitou

Leiden University, eLaw@Leiden, Centre for Law in the Information Society
Leiden, The Netherlands
d.g.klitou@umail.leidenuniv.nl

Abstract. This paper briefly explains the concept of *Privacy by Design* (PBD); provides an overview of the benefits and value of PBD for regulating and minimizing the threats to privacy posed by the increasing development, deployment and use of *Privacy-Invading Technologies* (PITs); outlines some of the limitations, constraints and potential criticism of PBD and the practical challenges of achieving, implementing and enforcing PBD legislation; and provides some potential counter-arguments to the criticism and some ways to help overcome the main challenges and constraints.

Keywords: Privacy by Design, Code as law, Privacy-Enhancing Technologies, Privacy-Invading Technologies, Legislation, Privacy principles, Data protection.

1 Introduction

The benefits and value of PBD[1] are now increasingly recognized or apparent for regulating and minimizing the threats to privacy posed by PITs.[2] But, PBD is not a

[1] Coined by Peter Hope-Tindall, during 1999, and promoted in a significant way by Ann Cavoukian, PBD can be described as *both* a form of 'Value-Sensitive Design' (VSD) and Lessig's "code as law". PBD is the realization of values, in this case the principles of privacy and corresponding rules/regulations, via the physical design, technical specifications, architecture and/or computer code of the device, system or technology concerned, where applicable. The aim of PBD is to design and develop a system or device (i.e. software and/or hardware) in a way that supports and materializes those privacy principles, values and rules as goals and functions, whereby that system or device then becomes 'privacy-aware' or 'privacy-friendly'. In other words, PBD can be defined as practical measures, in the form of technological and/or design-based solutions, aimed at bolstering privacy/data protection laws, better ensuring or almost guaranteeing compliance, and minimizing the privacy-intrusive capabilities of the technologies (i.e. PITs) concerned.

[2] For the purposes of this paper, Privacy-Invading Technologies (or Privacy-Intrusive Technologies) (PITs) are generally/broadly defined as and encompass:

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 86–110, 2014.

panacea for defending privacy. The concept is certainly not immune to criticism and the significant challenges and difficulties of legislating for PBD, implementing/enforcing PBD, and monitoring, measuring or assessing its effectiveness cannot be overlooked.

Section 2 briefly explains the concept of PBD and outlines the benefits and value of PBD legislation. Section 3 summarizes the practical challenges of achieving PBD legislation. Section 4 outlines the challenges of enforcing and implementing PBD legislation. Section 5 provides an overview of some of the limitations and constraints of PBD. Section 6 explains some of the additional and potential criticism of PBD, while Section 7 offers counter-arguments to the criticism. Section 8 proposes ways to help overcome these challenges. Section 9 concludes the paper with some concluding remarks.

2 The Growing Need for PBD Legislation

Although the words "privacy by design" (or "data protection by design") are not yet specifically found in the current legal framework in the US and EU, and the current data protection legal framework and privacy policies certainly do not seek to influence and/or alter the basic architecture of computer systems/information technology [Agre and Rotenberg, 1997],[3] some provisions that require the implementation of technical measures can be found in US and EU legislation.

In the US, for example, the Privacy Act 1974 requires government agencies to: "establish appropriate administrative, technical, and physical safeguards to insure the security and confidentiality of records".[4] In addition, Section 1173 of the Health Insurance Portability and Accountability Act 1996 similarly requires healthcare providers to implement technical and physical safeguards, in order to "ensure the integrity and confidentiality of the [health] information". In the EU, Directive 95/46/EC

Any form or type of technology, whether hardware or software, product or service, which poses a particular threat to privacy and/or is capable of being used to substantially violate an individual's right to privacy and/or data protection rights.

PITs mainly consist of technologies with (blatant or latent) surveillance capabilities or other technologies that disallow techniques/approaches for reducing privacy risks/threats. To some extent, however, nearly all information and communication technologies (ICTs) could be potentially regarded as PITs, including, for example, the Internet, digital services, mobile phones, cameras, credit cards, electronic voting machines and even contemporary photocopiers. Moreover, all technologies that enhance and/or replace human senses, particularly sight and hearing, are PITs. Therefore, PITs include not just ICTs, but especially other types of technologies, such as DNA technology, neurotechnology, identification technologies, nanotechnologies, advanced imaging technologies and mass surveillance technologies.

[3] This issue may be about to change. Article 23 of the EC's draft proposal for a General Data Protection Regulation (COM(2012) 11/4 draft) is dedicated to "data protection by design" requirements. However, as later explained, these requirements are only applicable to data controllers.

[4] Title 5, U.S.C. Part I, Chapter 5, Subchapter II, § 552a (e) (10).

(Article 17(1)) requires that data controllers "must implement appropriate technical and organizational measures to protect personal data". Article 4.1 of the ePrivacy Directive (Directive 2002/58/EC) requires that a "provider of a publicly available electronic communications service must take appropriate technical and organisational measures to safeguard security of its services", and Article 14(3) requires the adoption of measures "to ensure that terminal equipment is constructed in a way that is compatible with the right of users to protect and control the use of their personal data". Article 3.3(c) of Directive 1999/5/EC[5] also delineates that certain apparatuses may be required to incorporate "safeguards to ensure that the personal data and privacy of the user and of the subscriber are protected".

However, while there is already a very limited legal basis for PBD in the US and EU (i.e. legal provisions that require technical measures), the legal provisions place an emphasis on Privacy-Enhancing Technologies (PETs).[6] The existing legal

[5] Directive 1999/5/EC of the European Parliament and of the Council of 9 March 1999 on radio equipment and telecommunications terminal equipment and the mutual recognition of their conformity.

[6] The PISA (Privacy Incorporated Software Agent) project consortium defines PETs as "a system of ICT measures protecting informational privacy by eliminating or minimizing personal data thereby preventing unnecessary or unwanted processing of personal data, without the loss of the functionality of the information system" [van Blarkom, G.W. et al., 2003, p. 33]. PETs mainly comprise of cryptographic techniques, encryption, pseudonymization and anonymization software, firewalls, onion-routing communications, and other privacy protection tools developed primarily to better ensure the security, confidentiality and integrity of personal data. Indeed, as initially conceptualized, PETs, like PBD, are equally intended to be built into the architecture or 'fabric' of an information system (or technology, device, etc.) at the very outset [Hes and Borking, 2000]. However, while the concept of and premise behind PETs are similar to PBD, PETs and PBD are not the same. PETs are effective (software) technologies or ICT measures, but they can still be potentially circumvented or penetrated. On the other hand, the circumvention of PBD solutions is essentially meant to be (practically) impossible or exceptionally difficult, since it would mean attempting to force the device/system concerned to perform an act it is not designed or engineered to do or is not capable of doing (in its present form). The privacy risk or threat of technologies and/or the potential for abuse of the privacy-intrusive capabilities of PITs by the controllers/operators of these technologies is permanently removed, for the most part, through the regulation and minimization (or elimination) of those risks, threats and capabilities. Furthermore, PBD goes beyond PETs. Whereas PETs are *mainly* technical/technological or software-based solutions/ICT measures for protecting privacy and maintaining data security, PBD, as Cavoukian [2011] advocates, also includes 'privacy-friendly' physical design/architectural solutions and technological/software-based solutions, *and* business practices/processes, organization measures and modes of operation [see Cavoukian, 2011]. However, it is important to note that these other non-technical measures should not serve to substitute 'privacy engineering'. In addition, PBD emphasizes the need to implement PETs, but also requires 'privacy by default' settings and the necessary tools to allow users to participate in the protection and management of their personal data (e.g. access controls, user participation tools, etc.). Moreover, while PETs are often mainly focused on ICT privacy/security issues, PBD, as opposed to PETs, can potentially address the privacy threats of not just ICTs, but also the threats posed by other types of PITs (e.g. imaging technologies). Hence, PBD does not just apply to computer or information systems, but may apply to just about any type of device, system, service or technology, albeit to a certain extent.

provisions are also only applicable to data controllers/service providers, and primarily do not apply to technology manufacturers/developers. Moreover, the technical emphasis, found both in law and industry standards, is all too often focused on data security. As a result, there is a lack of guidance, binding rules and established industry standards on the technical solutions for ensuring the principles of privacy overall.[7] While data security is an important element of privacy protection, it is just one principle of privacy and not the whole picture. All of the privacy principles, which are similarly enshrined in the OECD Privacy Guidelines (1980), EU Directive 95/46/EC and the US Privacy Act 1974, must be realized through PBD.

An emphasis on data security is especially not sufficient to address the type of threats posed by the latest PITs. Many of the latest PITs pose a threat to privacy beyond the consequences of unauthorized access to personal information/data. For instance, the ability to see through clothes (using body scanners) or walls (using LEXID® devices), listen to and record public conversations (using, e.g., public CCTV microphones), scold people from afar (using, e.g., public CCTV loudspeakers), conduct wide-area aerial surveillance (using UAVs with advanced cameras), potentially read people's minds (using neurotechnology), constantly track people's movements (using, e.g., mobile phones, RFID technology, GPS tracking devices, automatic license plate recognition technology, etc.) and predict the health, intelligence and attributes of individuals (using DNA technology) are just a few examples of the incredible privacy threats that data security nor informational privacy/data protection alone cannot adequately address. Essentially, given the legal requirements for safeguarding privacy and the diverse privacy risks, the law must significantly go beyond legal provisions that only mandate technical solutions/measures for data security [Borking, 2010].

As opposed to only focusing on the technical measures for data security, technical and design solutions should, for instance, also help to control what personal data may be collected or accessed, when and how it may be collected, accessed or shared, for how long it may be stored, and provide data subjects the means to easily access their stored personal data. Accordingly, a holistic approach should be taken, whereby all the privacy principles are incorporated into the design of the system or device concerned, where applicable.

The physical or practical implementation/realization of the privacy principles through PBD can potentially address the threats to privacy at the earliest possible stage of a PIT's lifecycle – i.e. during the research, design and development stages. However, if we do not take into consideration the other privacy principles or norms (i.e. beyond data security), when designing and/or developing the functions of the relevant system or device, a diminishing potential for the realization or viability of those principles/norms will likely result. For example, with regards to the access/participation principle of privacy, while a data subject's right to request access

[7] See, for further discussion, the Online consultation comments on the European Commission staff paper, "Early Challenges to the Internet of Things", comments submitted by CA, Inc., p. 6, http://ec.europa.eu/information_society/policy/rfid/library/index_en.htm

to review the personal information stored on them by a data controller is provided for, e.g. within Directive 95/46/EC, the implementation of this right will likely be too difficult, impractical or costly, if the relevant information system has not been designed/developed in the first place to execute this request efficiently and/or cost-effectively. PBD will also especially be imperative in a 'ubiquitous information society', where it will likely prove difficult to determine all the responsible entities and to enforce privacy/data protection laws in the traditional way, and will also be evermore important as ICT becomes increasingly pervasive within everyday life and as cross-border data flows continue to increase. The latter is especially a problem, since different legal jurisdictions will likely continue to have different or even conflicting data protection laws/rules/standards. As Reidenberg points out, technological solutions can better ensure the consistent protection of personal data, to a certain extent, regardless of geographic location, legal jurisdiction or the adequacy of the legal framework, since "mechanisms that automate the implementation of data policies will facilitate uniformity across the areas of law and marketplace" [Reidenberg, 2000]. By working to address the privacy-intrusiveness of PITs at the design stage, PBD is also needed for helping to ensure that we are able to better stay apace with the ever-increasing and ever-evolving technological threats to privacy.

PBD may also be a pragmatic and integrated approach for safeguarding *both* privacy/liberty and security in the 21st Century. There are clear technological examples demonstrating this to be true. For instance, the automatic employment of privacy algorithms/software solutions during the generation of body scanner images, together with intelligent detection engines or automatic threat recognition capabilities, can help airport screeners/security officers to detect/locate threats by highlighting objects and helping to reduce human errors. At the same time, these measures can better protect the privacy of passengers by reducing the unnecessary level of graphic detail found in the images and/or potentially doing away with the need for human operators to directly view the images. Built-in and automatic limitations on storing, printing and transmitting the body scanner images can also better ensure the privacy principles are implemented. Strong encryption for RFID microchips, which can potentially help to prevent 'cloning' and the unauthorized access to the stored personal information, and protocol-level controls, which can potentially ensure that only authorized readers are able to read certain RFID microchips, also enables the security benefits of electronic identification and tracking to be realized, where and when appropriate. Designing/developing CCTV microphones, currently deployed in the UK and more recently in Canada, using artificial intelligence/software agents, can enable the permanent limitation of the activation of the recording capabilities of CCTV microphones when only certain sounds considered dangerous are detected. The potential sound detection capability of microphones attached to pan-tilt-zoom (PTZ) CCTV cameras can also enhance the ability of CCTV camera operators to aid in criminal investigations and support public security and potentially reduce the number of PTZ cameras needed to cover a larger

area [see Kim et al., 2007],[8] while at the same time can potentially facilitate a certain level of privacy out in public.[9] Designing/developing CCTV loudspeakers, also currently deployed in the UK, in a way that enables their use to be registered and prevents abuse or disproportionate use, for instance, may also allow the operators to assess where and how the loudspeakers can be more effectively deployed and used. As a result, PBD can also help to provide the potentially effective means for avoiding the false dichotomy of *privacy vs. security* [Cavoukian, 2009][10] and for proving that protecting privacy and maintaining public security is not necessarily a zero-sum game [*Ibid.*]. U-Prove cryptographic technology, the 'anonymous credential system' Identity Mixer, the Prime/PrimeLife FP7 research project and Cynthia Dwork's Differential Privacy scheme, for example, may bring to light or demonstrate the non-zero-sum properties of PBD and PETs.

However, neither law nor technology alone can effectively ensure privacy is maintained and both are not fully self-sufficient [Reidenberg, 2000].[11] As Reidenberg [2000] further argues, both *forms of regulation* "embody inherent limitations that preclude adequacy for effective protection of privacy". Therefore, a proper

[8] Sound is omni-directional as opposed to vision, which is directional, and, unlike vision, sound is not negatively affected by poor lighting or entirely obstructed by obstacles/objects. Microphones can provide CCTV systems and operators the ability to detect crime beyond a camera's field of view and can help them to work better in areas with insufficient light. If several microphones are installed at a certain distance from each other, the location of the sound source can automatically be determined, based on the time difference of the arrival from the sound source to the sensors [see Kim et al., 2007, p. 384]. A PTZ CCTV camera can be pointed in that direction and the operator can be both audibly and visibly alerted to contact the police immediately. CCTV microphones can, therefore, enhance the vigilance and effectiveness of CCTV camera operators and help them to observe more monitors or video streams, without having to hopelessly attempt to watch each simultaneously at all times.

[9] The potential PBD solutions, based on artificial intelligence/software agents, for CCTV microphones could also be applied to the video recording capabilities of public surveillance CCTV cameras. For example, a software algorithm/agent could be potentially used to process images in real-time and distinguish between 'suspicious behavior' or illegal activities and 'innocent behavior' or legal activities and then only begin to record video when a suspected crime or anti-social act is actually taking place, ignoring ordinary activities (see The Royal Academy of Engineering: Dilemmas of Privacy and Surveillance: Challenges of Technological Change. London, (2007), p. 42,
`http://www.raeng.org.uk/news/publications/list/reports/dilemmas_of_privacy_and_surveillance_report.pdf`

However, the current capabilities of artificial intelligence/software agents is not yet sophisticated enough to realize these potential PBD solutions. Moreover, the incorporation of artificial intelligence/software agents may also require separate legislation to address the associated ethical issues (see, for further discussion, Schermer, 2007).

[10] See also Ann Cavoukian's "7 Foundational Principles of *Privacy by Design*", January 2011,
`http://www.ipc.on.ca/images/Resources/7foundationalprinciples.pdf`

[11] The same is true for security, since security cannot be maintained with technology alone. Laws are also clearly still required.

combination or mixture of law and technology is required to effectively safeguard privacy. PBD is that critical combination of law and technology. Indeed, while PBD may significantly ease the dependence of privacy/data protection on user-level regulations and the compliance thereof, legislative instruments or other legal instruments will not simply become obsolete due to technological solutions. As Bruce Schneier, similarly points out, while technology is critical for protecting privacy, in the end, privacy boils down to the existence of laws and legal protections.[12] PBD solutions (and computer code) are not a substitute or replacement for law, but rather are complementary to law, and PBD is not an approach to replace lawmakers or lawyers with computer programmers or engineers. Computer code neither replaces lawmakers or lawyers. Moreover, computer code, when used to enforce privacy/data protection laws, is not or does not become law, but remains as the technical means for enforcing the laws [Dommering, 2006]. For instance, as Schwartz argues, a technical solution like P3P is necessary to provide the machine-to-machine protocol to enable a web browser and website to negotiate privacy standards, but laws are also necessary to require that those negotiations take place [Schwartz, 2000, p. 759].[13] Besides, PBD should be based on and enforced through law [Hildebrandt and Koops, 2010].

In order to realize the potential benefits of PBD, legislation, which mandates the implementation of PBD solutions/measures, should be adopted. Alternatively, explicit PBD provisions could be further incorporated into existing (privacy/data protection) legislation for different domains and technologies. In addition, PBD provisions/requirements could also be incorporated into existing legislation on defective products and the liability of manufacturers thereof. However, adding specific provisions to existing legislation may not be sufficient. In any case, PBD requirements should not be laid down in more voluntary codes, guidelines or self-regulations, but rather ought to be mandated by binding 'hard' laws.

Either way, the legal requirements to implement PBD should be applicable, where relevant, to *both* private and public entities and *both* manufacturers/developers of hardware and software (i.e. technology providers) *and* data controllers/service providers. As the Article 29 Working Party similarly points out, data controllers are often merely users of ICT and can hardly be considered in a position to take any relevant or effective security or data/privacy protection measures on their own, even if they wanted to.[14] Data controllers may also not be able to find or procure adequate and/or suitable PETs and the unfavorable economic incentives are not helping to address this technological deficiency. In addition, PBD legislation should mandate that the principles of privacy must be engineered into all PITs (with certain exceptions) (or initially just ICTs and digital services) manufactured/developed for private use and/or commercial sale *and* government use in the jurisdiction(s) concerned.

The implementation/enforcement of PBD legislation will require technology manufacturers/developers to be reasonably held accountable/liable for failing to incorporate

[12] Schneier, B. "Strong Laws, Smart Tech Can Stop Abusive 'Data Reuse'". Wired News, 28 June 2007, http://www.schneier.com/essay-175.html

[13] However, P3P was never fully realized and is not the best example of PBD.

[14] Article 29 Working Party, "The Future of Privacy", 1 December 2009, WP 168.

adequate and verifiable PBD solutions/technical measures.[15] Likewise, manufacturers/developers should be held accountable/liable for 'privacy defective' devices/products and services that result from demonstrated negligence/fault and cause damages to a person (or group of persons) as a consequence,[16] albeit subject to certain exceptions (e.g. the "state of the art defense" exception). The implementation/enforcement mechanisms could potentially consist of a mix of certification schemes, privacy audits (performed by neutral third parties), conformity declarations, product recalls, and sanctions. The individuals substantially affected may also be entitled to receive compensation.

Accordingly, Article 23 of the official draft EU General Data Protection Regulation,[17] which significantly proposes "data protection by design" (i.e. PBD) requirements, should additionally stipulate that these requirements also apply to the *manufacturers/developers* of the products/services in question. Besides, the additional application of Article 23 (paragraphs 1 and 2) to manufacturers/developers could bring greater legal clarity and purpose to paragraph 3 of Article 23, which empowers the EC to adopt delegated acts specifying appropriate technical measures/mechanisms (i.e. PBD solutions) for implementing PBD for products and services.[18]

3 Challenges of Achieving PBD Legislation

However, there are a number of considerable challenges that must be prevailed over before achieving the adoption of PBD legislation, not to mention the resistance and objections that will likely be raised by technology developers/manufacturers and powerful lobby organizations on their behalf.

To begin with, it will be difficult to ensure that PBD legislation (or PBD provisions) is both comprehensive and technology-independent, covering all threats to privacy posed by the latest technologies in existence, let alone those yet to be developed or imagined, *and* also specific/precise enough at the same time.

---

[15] In 2009, Senator Patrick Leahy introduced S.1490, titled "the Personal Data Privacy and Security Act of 2009", which aimed to hold software companies liable for security flaws or vulnerabilities and to mandate that business entities implement data privacy and security technical and physical safeguards in the system's design and impose civil penalties on entities that fail to do so. While the legislation essentially covered 'information privacy', as opposed to the protection of privacy overall, the legislative proposal has some similarities to the PBD legislation proposed here.

[16] A perfect and fairly recent example of a privacy defective device/service includes certain models of the Trendnet home security cameras that were discovered to have flawed firmware allowing anyone to access online live feed without requiring a password.

[17] Proposal for a Regulation of the European Parliament and the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), COM(2012) 11/4 draft.

[18] Article 23, para. 3 of the proposed General Data Protection Regulation could potentially have an indirect effect on technology manufacturers/developers, since the specification of appropriate measures/mechanisms for implementing PBD for products and services would likely put pressure on the manufacturers/developers of those services/products to conform.

The law, in terms of privacy protection, is often enhanced either with greater specificity through additional specific legislation or additional specific provisions/amendments to existing laws. Specificity helps to allow the law to be predictable and consistent and removes ambiguity. Moreover, in order to ensure that the necessary PBD solutions can be developed appropriately, the underlying PBD requirements mandated through PBD legislation will need to be clarified precisely and consistently [Pasic, 2011].[19]

Still, *both* greater precision and clarity and sufficient room for flexibility is needed. Flexibility allows for the adjustment to new circumstances or the emergence of new technologies, which is especially required in a world of constantly advancing PITs. Sometimes flexibility in law is effective, while at other times more specificity is required. For instance, the legal definition of personal data and the definitions of what constitute PITs, location information and tracking devices require flexibility, in order to ensure all applicable technologies, devices, etc. are broadly covered now and in the future. On the other hand, the definition of location information also requires a certain level of specificity, in order to remove any doubts or close any legal loopholes concerning the privacy of location information. Moreover, stipulating which particular sounds, for example, may appropriately activate CCTV microphones to begin recording clearly requires a certain degree of specificity. Essentially, the difficulty is to balance the need for specificity with the benefits and needs of flexibility.

PBD legislation, in particular, also requires flexibility, since it is nearly impossible to delineate every design and/or technical requirement and may also be unhelpful to overly prescribe the PBD solutions for each and every PIT, even though there may often be just one or two meaningful solutions for each type of PIT. The goal is for the potential PBD legislation to be as broad and comprehensive as possible when mandating the implementation of PBD solutions. Nevertheless, the PBD solutions will also need to consider the particular or unique characteristics and privacy threats and risks of the different devices, systems or technologies concerned.

The resulting practical challenges of implementing PBD and embedding the principles of privacy into the design/architecture of ICT and other categories of PITs, in reality, also cannot be ignored. Evidently, 'translating' written legal norms/principles into design solutions/computer code or bridging the significant differences between legal (natural) language and computer/machine language is a major challenge.[20] Creating devices or systems that are capable of effectively implementing laws and rules that we as humans understand in the form of legal natural language (LNL) and devices, systems, computers, etc. understand in the form of legal machine language

[19] For further discussion, see Pasic, A.: Privacy by Design: An industry perspective on the challenges and opportunities of privacy (2011), http://www.eurescom.eu/?id=531

[20] For further discussion on the conflict between laws/regulatory principles and privacy engineering approaches, see Guarda, P., Zannone, N.: Towards the development of privacy-aware systems. Information and Software Technology, vol. 51, no. 2, 337-350 (2009).

(LML)[21] is also a major challenge. Indeed, the extent to which legal protection can be programmed, engineered or automated is open to discussion, and there is currently no single widely accepted methodology or approach for translating privacy/data protection laws/principles/policies into technological/design solutions. Basically, "there is no well established and worldwide accepted view on the way privacy protection and the consolidation thereof can be built into software" [van Blarkom, G.W. et al., 2003, p. 2]. The same is still true overall for PBD.[22]

There are some limited benefits from the flexibility, as briefly explained above, often intrinsic in natural language, which will be mostly forgone due to the specificity and rigidity of machine/computer language. As Grimmelmann argues, "[b]ecause a computer, rather than a person, makes a program's decisions, rules encoded in software are free from ambiguity, discretion, and subversion" [Grimmelmann, 2005, p. 1723]. The challenges or difficulties generally concern the flexibility of human interpretations and understanding of natural language, and how this also differs or conflicts with the rigidity of machine/computer language interpretation [Grimmelmann, 2005]. For instance, lawmakers/legal practitioners may interpret or understand the legal norms/privacy principles differently, given the lack of overall consensus on what constitutes privacy, and these interpretations may also change over time and under different circumstances, causing the 'translation' to be further complex. And, evidently, "[a]s the complexity of particularized rules increases, their formal realizability decreases" [Grimmelmann, 2005, p. 1733]. Technology/design-based solutions work best for areas/matters where there is a consensus on meanings, but is certainly more challenging where there is significant disagreement [Yeung and Dixon-Woods,

[21] Using LML to implement LNL was also proposed/developed for enhancing the "safety intelligence" of Next Generation Robots (NGRs) and may also be applicable for better ensuring the privacy friendliness/privacy awareness of other technologies. see Yueh-Hsuan, W., Chien-Hsun, C., Cheun-Tsai, S.: Safety Intelligence and Legal Machine Language-Do we need the Three Laws of Robotics? In: Yoshihiko, T. (ed.) Service Robot Applications. InTech Education & Publishing (2008), http://works.bepress.com/weng_yueh_hsuan/3

[22] Accordingly, the consortium for the Privacy Incorporated Software Agents (PISA) project, funded under the European Framework Programme, set out to establish an accepted methodology for incorporating privacy protection into software and to address the technical challenges of data protection, formulating a process that included analyzing the legal requirements to determine the required human behavior, then translating privacy laws and data protection rules into design or technical solutions for each requirement, based on the "engineering psychology" approach, and subsequently conducting a privacy audit. The methodology, called "Design Embedded Privacy Risk Management" (DEPREM), was developed to realize "privacy knowledge engineering" (PYKE), in order to build the privacy principles and data protection rules, based on the formulation of ontologies, into an intelligent software agent [van Blarkom, G.W. et al., 2003, p. 169]. While the *Handbook of Privacy and Privacy-Enhancing Technologies* [van Blarkom, G.W. et al., 2003], created by the PISA project consortium, provides a methodology for designing for privacy, it is more relevant for developing PETs and middleware to protect privacy, than for PBD. Ontologies can help to provide the common language and understanding necessary for incorporating the principles of privacy into the design and architecture of technologies [van Blarkom, G.W. et al., 2003].

2010]. In addition, some of the legal norms/privacy principles may be too vague or may not be specific or detailed enough, which may call into question the ability of programmers/engineers to effectively develop/implement PBD solutions for realizing the privacy principles/legal norms in a methodical and consistent way [Pasic, 2011].[23] As Grimmelmann insightfully also explains, programmers require precision, since they must articulate their objectives within the computer program text as "a list of instructions, written in one of a number of artificial languages intelligible to a computer" and these languages are "highly constrained", relative to human languages, and the instructions have "a fixed and precise meaning" [2005, p. 1728]. Accordingly, the significant difficulty of identifying a widely accepted methodology or approach for 'translating' privacy/data protection laws/principles into technological/design solutions may be due to the fact that the laws and principles are somewhat too ambiguous.

Therefore, the PBD requirements will need to be detailed and precise enough, in order to ensure that developers/manufacturers are able to clearly identify or determine the specific requirements [Pasic, 2011]. Then, developers/manufacturers/engineers will also be better able to develop/implement concrete PBD-based solutions for complying with these specific requirements and norms, while still taking into consideration the specific characteristics and various privacy threats/risks of different devices, systems or technologies concerned. Equally, the principles of privacy and other legal privacy norms will also need to be as specific as possible, in order for computer programmers to effectively codify the privacy principles/norms through computer code. Nevertheless, as Grimmelmann points out, even the most precise rules could still provoke a certain degree of discretion or judgment and facts are still vulnerable to different preconceptions and other "non-legal sensibilities" [Grimmelmann, 2005, p. 1733].

But, while legal specificity/precision is required for practical purposes, at the same time, the concept of PBD legislation again may also need to be technologically-neutral, goal-orientated and general or flexible enough to ensure that all technologies, devices, systems, etc. and domains are covered. Also, the corresponding PBD requirements will need to be flexible enough to allow and encourage the development of innovative PBD solutions. But, in any case, all PBD solutions must strictly be based on the defined privacy principles, norms and relevant laws. Finding the right balance will be a considerable challenge.

Furthermore, the challenges pertain to the potential political and economic reservations. In order to induce lawmakers to take the necessary steps to adopt new comprehensive laws requiring the implementation of PBD, they will need to further recognize that the protection of privacy is an additional source of political legitimacy. In addition, before the adoption of PBD legislation can be achieved, lawmakers/policymakers will also need to be influenced and convinced, perhaps through concrete PBD solutions and validated real-life demonstrations, that privacy can be engineered into PITs. Indeed, by providing the actual ability to take concrete steps,

[23] Pasic, A.: Privacy by Design: An industry perspective on the challenges and opportunities of privacy (2011), http://www.eurescom.eu/?id=531

PBD could potentially offer the necessary preconditions for addressing privacy concerns on a political and economic level [see Agre and Rotenberg, 1997].

The economic reservations can come from the extra costs and burdens associated with PBD. Significant investment and resources from both the private and public sector will need to be allocated to carry out the necessary research and development and innovation, in order to realize effective PBD solutions, tools and methods for implementing and enforcing the relevant privacy principles and laws thereof. In view of that, there will certainly still be moments when companies and governments are inclined to violate privacy and design devices or systems that threaten privacy, whether deliberately or unintentionally, lawfully or unlawfully. The actual implementation of PBD depends, in part, on the willingness of technology developers/manufacturers to comply, which in turn is influenced by their need to design/develop profitable and marketable products. Technical solutions cost money and avoiding or delaying compliance may be the easy or more profitable way out. Since PBD solutions come at an additional cost, in order for PBD to be employed or implemented at an acceptable rate, the developers/manufacturers of PITs must also be convinced and fully aware of the value and financial justification or business benefits in complying and the financial costs, risks and liabilities of failing to comply. From a business perspective, as Borking points out, it makes no sense for a technology developer to invest in a privacy protecting solution (i.e. a PBD solution), if the actual costs of the solution are greater than the value it actually offers [Borking, 2010, p. 260]. The value will only increase as consumers increase their demand for privacy-friendly products/services. However, relatively "little work has been done to evaluate the economic impact of privacy policy" [Agre and Rotenberg, 1997] or to study the "economics of privacy"[24] or to explore how privacy can potentially be monetized.[25]

Essentially, as long as the business case and economic incentives are weak and the political will is absent, PBD legislation will not materialize. Reaping the benefits of PBD will equally require constructive political and economic choices, in addition to technical choices. Therefore, radically changing the way companies and governments design, develop and/or procure PITs will require, not just new technological and legal

[24] See, e.g., Posner, R.: The economics of privacy. American Economic Review, vol. 71, no. 2, 405-409 (1981); Posner, R.: An economic theory of privacy, Regulation, 19-26 (1978).

[25] The topic, however, since 1997, has grown into its own area of specialty. see, e.g. Taylor, C.R.: Private demands and demands for privacy: Dynamic pricing and the market for customer information. Technical report, Department of Economics, Duke University (2002); Acquisti. A.: Security of Personal Information and Privacy: Technological Solutions and Economic Incentives. In: Camp, J., Lewis, R., (eds.) The Economics of Information Security. Kluwer (2004); Jentzsch, N., Preibusch, S., Harasser, A., Ikonomou, D., Tirtea, R.: Study on monetising privacy. An economic model for pricing personal information. ENISA (2012). For additional examples of papers/books on the "economics of privacy", see, e.g., Acquisti's academic website/blog available at:
`http://www.heinz.cmu.edu/~acquisti/economics-privacy.htm`
John Borking has also conducted significant research on the costs of privacy risks for businesses and on quantifying the economic justifications for organizations to invest in privacy risk-reducing technical solutions, such as PETs.

solutions, but also the sound basis, legal support and incentives to overcome significant economic and political reservations.

4 Practical Challenges of Enforcing and Implementing PBD Legislation

Moreover, there are also noteworthy challenges of enforcing PBD legislation. As data controllers/processors and service providers increasingly use so many different technologies, devices, tools and systems, and as systems become increasingly complex, determining the specific technical problem or defect, identifying the responsible/liable party and establishing a link between the problem, defect or malfunction and the privacy damage is now less and less obvious.[26] For example, a RFID system could be composed of different types of RFID tags and readers, databases, fixed/mobile computing devices and software.[27] As a result of the (potential) lack of a clear understanding of responsibility or liability, the enforcement of PBD requirements will be problematic.

Again, the mixture of implementation/enforcement mechanisms could potentially consist of a mix of certification schemes, privacy audits, conformity declarations, product recalls and sanctions. There are established industry standards, implementing measures and audit mechanisms for ensuring data security, and, on top of that, comprehensive guidelines/checklists for conducting general and specific Privacy Impact Assessments (PIAs).[28] However, the difficult challenge is to develop additional standards, methodologies, criteria, indicators and mechanisms for auditing/judging the adequacy, performance and quality of the implementation of PBD. The challenges also pertain to the difficulty of measuring or quantifying privacy protection and, thus, the implementation of PBD requirements, in the normal or traditional sense. Indeed, often what you cannot measure, you cannot enforce effectively. Moreover, product recalls of software may be somewhat impossible or ineffective, and conformity

[26] See RISEPTIS: Trust in the Information Society: Research and Innovation on Security, Privacy and Trustworthiness in the Information Society. A Report of the Advisory Board RISEPTIS (2009).
(RISEPTIS was composed of more than 30 experts and was supported by an EC-financed 'Coordination Action' project, THINKTRUST, whose objective was to develop a research agenda for Trustworthy ICT).

[27] See Cannataci, J.A.: Recent developments in privacy and healthcare: Different paths for RFID in Europe and North America?. International Journal of RF Technologies, vol. 2, 173–187 (2010/2011).

[28] See, e.g., the ICO PIA Handbook for guidelines on conducting PIAs, available at: `http://www.ico.gov.uk/upload/documents/pia_handbook_html_v2/index.html`; Privacy and Data Protection Impact Assessment Framework for RFID Applications, 12 January 2011, available at: `http://ec.europa.eu/information_society/policy/rfid/documents/infso-2011-00068.pdf`

declarations are basically insignificant if the technical/design specifications utilized are still ambiguous or indefinite.

In addition, it will also be difficult to regulate the development of certain technologies/devices, such as surveillance technologies, strictly used by governments/public authorities for law enforcement and/or military purposes. These devices indeed may still need to be designed in way that more effectively violates privacy, while, nevertheless, still complies with the relevant laws and constitutional protections concerning their deployment and use. Even so, PBD requirements/obligations should overall still be applicable for technologies developed for law enforcement purposes.[29] While there may be some distinctions on how different actors (governed by different laws and needs) may be involved in using the same technology for different purposes, especially in light of creating PBD policies/requirements, the PBD approach is meant to be applicable regardless of the technology, legal framework or activity concerned. Once again, the concept behind the PBD approach, accordingly, must be technologically, entity and activity-neutral, as far as possible.

The success and effectiveness of the development and implementation of the required technical/design solutions is also dependent on the means, abilities, capacities and resources of the technology developers/manufacturers. Equally, the implementation of PBD depends on the availability of the required skills and know-how of programmers/engineers.[30] Undoubtedly, as a result, the development of certified-compliant PBD solutions and certified engineers/programmers will be a lengthy and complex process and will demand substantial investment, dedicated resources and training [Pasic, 2011].

Global data processing also poses significant challenges to the effectiveness of enforcement mechanisms [Reidenberg, 2000], and the responsible manufacturers/developers are not always located in the concerned/relevant legal jurisdiction. Indeed, most major digital/online services do not originate from the EU, and European PBD legislation would be pressed to regulate (or enforce technical measures for) any digital/online services, for example, that utilize servers based only in the US. Therefore, ideally, both the US and EU, and beyond, should adopt PBD legislation, given the global nature of the Internet and digital revolution and the privacy problems/threats at hand. Moreover, without common standards, between the US and the EU, for example, interoperability issues may also further emerge, if only the EU adopts and enforces PBD legislation. Although the EU is moving in the right

[29] Significantly, this is consistent with the EC's Proposal for a Directive on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data, COM(2012) 10 final, Brussels, 25.1.2012. Article 19 of the EC's proposal requires EU Member States to ensure that data controllers are complying with obligations arising from *data protection by design* and *privacy by default*.

[30] See Pasic, A.: Privacy by Design: An industry perspective on the challenges and opportunities of privacy (2011), `http://www.eurescom.eu/?id=531`

direction, with the proposed General Data Protection Regulation,[31] which significantly proposes "data protection by design" (i.e. PBD) requirements, it is highly unlikely that the US will adopt similarly comprehensive privacy legislation anytime soon.

5 Limitations and Constraints of PBD

While PBD may be critical for protecting privacy against the intrusive capabilities of the latest technologies, in practice the approach is *not* without additional limitations and constraints. Legally mandating that technical solutions be implemented at the earliest stage of development is, once again, not a panacea, not to mention the challenges of implementing/enforcing PBD and the potential criticism of PBD. There is simply no single, all-encompassing way to completely ensure that data controllers and operators/users of PITs consistently comply with all the privacy/data protection laws and principles.

Although PBD solutions aim to minimize the intrusive capabilities of the technologies concerned, PBD cannot address every privacy threat posed by every PIT, since likely not all privacy threats posed by the latest technologies can be designed or engineered away. As a case in point, PBD is understandably not an all-encompassing solution for dealing with the very complex and dynamic privacy issues surrounding the greater advancement and use of DNA analysis technology,[32] 'predictive technology' and neurotechnology. Similarly, as Grimmelmann [2005] points out, technology/software cannot implement every legal rule. Consequently, there are certainly some privacy threats/risks outside the scope of PBD solutions, at least for the time being. Therefore, where PBD (and PBD legislation) might not provide adequate safeguards for the most privacy-intrusive and disruptive technologies, further specific regulations/laws should also not be overlooked. Furthermore, maybe certain PITs should not be deployed or used at all, where and when the capabilities are particularly detrimental to the interests of a free and democratic society.

PBD solutions, in the end, are just as important as the laws, principles and norms that mandate these solutions be implemented, influence the end result of PBD, provide the legal mechanisms to intervene in the chain of production, specify the liability of not complying, punish those who illegally hacked or intentionally circumvented the PBD solution, ensure transparency, regulate overall how or where PITs are deployed/used and establish the enforcement and audit mechanisms. The law altogether must also be capable of ensuring that the inappropriate/unlawful development and use of PITs is not committed with impunity and that there are explicit penalties for

[31] Proposal for a Regulation of the European Parliament and the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), COM(2012) 11/4 draft.

[32] A simple potential example of a PBD solution for DNA technology includes limiting the creation and exchange of DNA profiles to chromosome zones containing no genetic expression (i.e. not known to provide information about specific hereditary characteristics), as recommended by the Council of the EU. see Council Resolution of 25 June 2001 on the exchange of DNA analysis results (2001/C187/01).

violations and available remedies for victims. In addition, PBD alone cannot implement all of the relevant legal requirements. For instance, administrative processes, such as the requirements of organizational accountability and notification requirements, cannot be implemented through PBD [van Blarkom, G.W. et al., 2003, p. 50]. Moreover, since certain PITs will again still need to be significantly intrusive for certain, lawful purposes (e.g., for law enforcement purposes or surveillance activities), constitutional and/or other legal instruments will, thus, still need to be considerably relied upon.

For all practical reasons, it will obviously also be difficult, for the most part, to apply PBD legislation retroactively, i.e. to existing (or already developed and deployed) devices/products/systems. PITs previously developed and deployed before the enactment of PBD legislation will certainly continue to exist in society, and will thus need to continue to be regulated primarily by user-level and *ex-post* regulations, where applicable. Thus, there will be a period of transition before achieving the new reality and specific objectives PBD promises. To address this limitation, the concept of "*Privacy by ReDesign"* was developed to apply PBD to existing systems by "*rethinking, redesigning* and *reviving*" these existing systems in a way that leads to the objectives of PBD.[33]

PBD neither can answer nor solve all of the critical legal questions. For example, while PBD can serve to develop location-based services and related products for consumer use in a privacy-friendly manner, it cannot help to determine or judge the legality of (warrantless) GPS tracking conducted in the US by law enforcement agencies or determine the privacy protections afforded to location information derived from mobiles phones, personal-locating devices, etc. or the overall level of privacy afforded to citizens/consumers out in public. Therefore, in summary, technical and/or PBD solutions *alone* cannot, in practice, guarantee privacy.

Furthermore, as Sollie and Duwell [2009] argue, an anticipatory outlook is required when addressing new technologies. However, given that the ability of the designers and engineers to imagine or anticipate all future scenarios is limited [Albrechtslund, 2007, p. 72], it is unlikely that all the intended and unintended eventual uses of a particular PIT, and the privacy threats thereof, can be foreseen at all times during the design and development stage or even after a PIA and privacy audit is conducted. For instance, predicting every privacy threat now and in the future will be particularly difficult in a 'ubiquitous information society'. Any uncertainty or unawareness of all the privacy threats and implications of the technology in question is equally a predicament for PBD, particularly if the technology, device, infrastructure, system or service has never been deployed and used yet. Therefore, since the development of new technologies regularly occurs under conditions of uncertainty, as Sollie and Düwell [2009] point out, then the effectiveness of PBD may equally be uncertain and limited at times.

[33] See the Seminar of the 33rd International Conference of Data Protection and Privacy Commissioners, *Privacy by ReDesign* Workshop, Mexico City, Mexico, November 1, 2011.

Despite the fact that PBD could help to better minimize the negative implications of the deployment and use of PITs, by controlling/minimizing the intrusive capability of the technology in the first place, care should be taken not to give the impression that technology developed under certain legal requirements is no longer susceptible to future ethical dilemmas or future technological advancements [Albrechtslund, 2007]. PBD is susceptible to the inclination that PITs, or any technology for that matter, are often never finished developing or advancing. As new capabilities are added, further unforeseen privacy implications may result. Though the goal is to design technology to be privacy-friendly in a way that transcends time, PBD, however, must be an ongoing process that requires continuous advancement, innovation and re-assessment as PITs continuously advance. If PBD is not as dynamic as technological advancement, then just like traditional laws, the PBD approach will also fall behind. Even with the methodical implementation of both PIAs and PBD solutions, unforeseen threats to privacy could still be encountered later on. Some of the PBD solutions themselves might later on result in unexpected privacy implications, as the technical solutions further advance.

In addition, not all PBD solutions will be effective at present or in the future. Some solutions, even those based on the BATs at the time and designed in a way to be future-proof as far as possible, could prove deficient or insufficient later on or end up being susceptible to circumvention or even end up failing. Some, if not most, PBD solutions will also be vulnerable to hackers. A number of PETs, for example, developed for ensuring privacy and data security on the Internet, have already been circumvented. Particularly during the initial phase, many of the new PBD solutions developed will likely fail or be circumvented. As experience has shown, there is no absolute guarantee that any system, software or device is completely free of vulnerabilities or privacy risks/threats, just as there is essentially no absolutely impenetrable security system or level of software encryption or error-free computer code.[34] Specifically, for instance, as Grimmelmann points out, "software is vulnerable to failure in three related ways: It is *buggy*, it is *hackable*, and it is *not robust*" [Grimmelmann, 2005, p. 1742]. Clearly, if a (PBD) software solution is hacked or somehow circumvented, the solution has not acted as an effective constraint [*Ibid*., p. 1731]. In other words, if the software component of a PBD solution is easy to bypass, for instance, then the physical design/architectural/hardware component is inconsequential.

In summary, PBD is, indeed, limited by the ability of designers and engineers to envision the privacy threats/risks posed by PITs, their ability to effectively design/engineer away all the various threats/risks to privacy and develop relevant robust solutions, and their ability to keep up with the ever growing threats/risks and intrusive capabilities of PITs.

[34] See, for further discussion, Grimmelmann, J.: Regulation by Software. Yale Law Journal, vol. 114, 1719-1758 (2005).

6 Other Potential Criticism of PBD

For various reasons, a number of legal authors/scholars have also criticized Lessig's "code as law",[35] which is an element/component of PBD. To begin with, Gutwirth et al. [2008] essentially argue that Lessig disregarded the politics, dynamics and complexity of lawmaking and how legal practitioners and courts operate in the real world. Gutwirth et al. [2008] also question the viability of achieving an "optimal mix" of Lessig's four dimensions/modalities of regulation.[36]

Schwartz [2000] also criticizes Lessig's concept of "privacy-control", clearly arguing that "privacy-control seeks to place the individual at the center of decision making about personal information use, but it can instead help us to accept smoke screens that disguise information privacy practices and lead to choices that are bad for individuals and for society".[37] Schwartz [2000] argues that Lessig's "technological solution, privacy-code, which relies on measures such as P3P, is likely to form such a smoke screen".[38] Schwartz [2000] further questions the effectiveness of over relying on individual control of personal data to achieve an elevated degree of privacy, as a result of 'market failures' and failures of private agreements,[39] further arguing that "due to the extent of the failure in the privacy market, the law at present should generally seek to minimize harms that flow from reliance on bargaining among consumers and data processors".[40] Moreover, Schwartz [2000] rightfully points out that Lessig's approach to 'individual privacy control' is mostly not relevant for law enforcement purposes, since law enforcement agencies are generally not required to obtain permission to carry out, for example, surveillance operations.[41]

As values, norms or rules are increasingly being built into technology, some authors, including, for example, Koops [2007], have also questioned the compatibility of "code as law" (and PBD) approaches with the democratic system of government, if not sanctioned by elected representatives, in accordance with the law.[42] Indeed, "code as law" may arguably be one way of bypassing regular democratic procedures of

---

[35] In *Code and Other Laws of Cyberspace*, Lessig (1999) outlined how regulating technology has four interacting and complimentary modalities or dimensions*: laws; norms; market*; and *physical architecture*, and how the effective regulation of technology can be achieved through the "optimal mix" of these dimensions.

[36] Gutwirth, S., De Hert, P., De Sutter, L.: The trouble with technology regulation from a legal perspective. Why Lessig's 'optimal mix' will not work. In: Brownsword, R. and Yeung, K. (eds.) Regulating Technologies, pp. 193-218. Hart Publishers (2008).

[37] Schwartz, P.M.: Beyond Lessig's code for Internet Privacy: Cyberspace Filters, Privacy-Control, and Fair Information Practices. Wisconsin Law Review, vol. 2000, no. 4, 743-787 (2000), at 760.

[38] *Ibid*.

[39] Schwartz, 2000, p. 782.

[40] *Ibid*.

[41] Schwartz, 2000, p. 784.

[42] See Koops, B-J.: Criteria for Normative Technology: An Essay on the Acceptability of 'Code as Law' in Light of Democratic and Constitutional Values. Tilburg University Legal Studies, Working Paper no. 007/2007 (2007).

lawmaking and/or law enforcement to regulate or restrict human behavior and activities. Computer programmers/engineers, in this sense, could indeed theoretically become the new lawmakers of the 21st Century, acting at the behest of either corporations or governments.

Moreover, PBD could also potentially have some unfavorable implications for defending privacy and liberty. For instance, PBD could be used by governments/corporations to cunningly ease privacy concerns and divert or minimize adverse reaction to the widespread deployment of certain PITs (e.g. UAVs/drones and RFID implants) that essentially conflict with the principles of privacy and the values/norms of a democratic society.

Additional potential criticism of PBD might come from those who raise the argument that *ex-ante* regulations on technological development (i.e. PBD legislation) could jeopardize or stifle innovation [Cave et al., 2009, p. 17] or hamper technology deployment and, therefore, could, in the long-run, also impede economic growth and competitiveness [*Ibid.*].

7 Countering Potential Criticism

In response to the potential criticism, PBD does not rely excessively on individual "privacy control". Indeed, PBD is an answer to Schwartz's [2000] criticism of "code as law", since PBD can serve as the means of *automatically* realizing the principles of privacy. In other words, PBD, in contrary, aims to implement the rules/principles of privacy protection primarily in a *self-executing* manner, i.e. without the constant, proactive or active involvement of individual choice or control or human intervention.

Moreover, while politics and market dynamics should obviously not be ignored, the effectiveness of PBD is not overly contingent on finding the "optimal mix" of the different modalities/dimensions of regulating technology development. After all, one of the main reasons in favor of mandating PBD, for example, is to overcome current market failures.

The PBD approach is also compatible with the way lawmakers, legal practitioners and courts operate in the real world. For instance, the concept behind PBD already has a limited legal basis in the US and EU, and the EC's draft proposal for a General Data Protection Regulation proposes PBD (or "data protection design") requirements. While the current (and proposed) privacy/data protection laws *primarily* apply to data controllers/processors, and not technology developers/manufacturers, there is, nevertheless, also a legal basis for this approach. For instance, patient safety, automobile safety and consumer and environmental protection laws already regulate how certain products are designed and developed/manufactured.

Given that PBD will still require traditional legal approaches and again is not a substitute for law or lawmakers, but is rather meant to implement/enforce existing laws, norms and principles; there is also little or no reason to assume that it is incompatible with democracy. As Schwartz similarly argues, the application of the privacy principles ensures the involvement of our democratic institutions, and since PBD is

based on the principles of privacy, lawmakers are already involved in the process of shaping the technological requirements and solutions [Schwartz, 2000, p. 787].

While the potential criticism of PBD (and PBD legislation), concerning its impact on innovation and business, is not without merit, this argument may overlook the potential benefits of PBD in promoting the deployment and innovation of future and emerging technologies through the increased trust and confidence of consumers/citizens. On the contrary, the hurdles to the substantial further deployment, innovation and mainstream take-up of technologies are partly due to the general perceptions, mistrust and concerns of citizens, consumers, privacy activists and civil society. The hurdles may also be due to the hesitation of both technology manufacturers/developers and service providers, which result from the uncertainties and ensuing investment risks. The growing lack of trust in companies to ensure privacy, data protection and data security is increasingly resulting in missed business opportunities and sluggish innovation [Williams, 2009, p. 78]. Lawmakers can alleviate the resistance and backlash to new technologies and facilitate their rollout and mass market take-up through the adoption of an appropriate and predictable legal framework. Specific and up-to-date privacy laws/regulations and PBD solutions will enable companies and governments to earn the trust and confidence of consumers/citizens over the use of PITs, thereby facilitating their widespread deployment and use, which in turn could further promote the necessary investments in advanced ICT research and innovation. Specific legal regulations on the design, development and manufacture of PITs could also enable the developers to design and manufacture these technologies with fewer concerns or uncertainties over the future legality and liability of their investment. Without specific, unambiguous, comprehensive and future-proof regulations/laws, the developers have no definitive standards to follow, which could further stifle innovation and lead to uncertainties and confusion for both industry players and consumers alike. As the RISEPTIS Advisory Board similarly points out, with regards to e-services, appropriate technical and legal infrastructures will remove barriers to innovation, as businesses will only invest in e-service solutions if the legal obligations are clear [RISEPTIS Report, 2009, p. 14].

Furthermore, Grimmelmann's warranted analysis that computer code/software is also *malleable* and *vulnerable*, and is not the same as physical architecture,[43] is somewhat offset by the fact that PBD includes *both* physical hardware/design/architectural solutions and software solutions. PBD does not and should not aim to equate the two types of solutions.

8 Overcoming the Challenges

To address or overcome some of the challenges of striking the right balance between specificity and flexibility for PBD legislation, lessons could be potentially learned from environmental law/regulation and the approaches to 'green by design'. As Hirsch notably argues, 'command-and-control regulation' applied in environmental

[43] Grimmelmann, J.: Regulation by Software. Yale Law Journal, vol. 114, 1719-1758 (2005).

law, is not necessarily suitable for privacy law/protecting privacy [2006, p. 33]. In environmental law, "regulators identify the best currently existing technology for controlling pollution in that industry (known as the "reference technology")" and "either direct all facilities in the industry to install the chosen technology (this is known as a "design standard")" or require that the regulated facilities do not exceed the rate of pollution they would emit if they had properly used the reference technology (this is known as a "rate-based standard") [Hirsch, 2006, p. 33]. As Hirsch [2006] further points out, with regards to environmental protection, "command-and-control also deters innovation in pollution prevention and locks in the current state of pollution control technology" [Hirsch, 2006, p.35]. The same may hold true, as Hirsch [2006] argues, for privacy protection technologies.

While the "rate-based standard" may make somewhat more sense for protecting privacy than the "design standard", since it may permit different methods or means for achieving the same goal, the "rate-based standard" still relies, in effect, on the reference technologies on which the rate is based, as Hirsch points out, and "almost all [companies] choose the reference technologies so as to avoid any misunderstanding about compliance" [Hirsch, 2006, p. 34]. As Hirsch further argues, "[b]y requiring firms to meet the best existing level of control technology, it gives them no incentive to exceed this level" and "the method is too slow for rapidly evolving industries" [2006, p. 35]. Therefore, as Hirsch [2006] argues, both standards are just different types of command-and-control regulation and, as a result, both would likely not hold up against the rapidly evolving technological means of privacy intrusion. The EDPS recommends that PBD could potentially adopt the 'Best Available Techniques' (BATs)[44] approach.[45] However, BATs, which are also based on command-and-control regulations, can impel companies to adopt technological solutions that are already available [Hirsch, 2006, p. 35], thereby diminishing the prospect for developing more innovative solutions that are not yet available. *Innovative* PBD solutions will be crucial for minimizing the threats to privacy posed by future and emerging technologies.

Overprescribing the technical/PBD solutions to address the privacy threats of PITs could also discourage the continuous development or enhancement of new solutions that could progressively achieve even better results. Unlike the EC's draft General

[44] Council Directive 96/61/EC of 24 September 1996 concerning integrated pollution prevention and control defines BATs as "the most effective and advanced stage in the development of activities and their methods of operation which indicate the practical suitability of particular techniques for providing in principle the basis for emission limit values designed to prevent and, where that is not practicable, generally to reduce emissions and the impact on the environment as a whole" (Art. 2.11). Techniques include the use of technology.

[45] See European Data Protection Supervisor Opinion on the Communication from the Commission on an Action Plan for the Deployment of Intelligent Transport Systems in Europe and the accompanying Proposal for a Directive of the European Parliament and of the Council laying down the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other transport modes, 22 July 2009, available at: http://www.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2009/09-07-22_Intelligent_Transport_Systems_EN.pdf

Data Protection Regulation, which gives the EC authority to mandate specific technical measures/solutions and standards, PBD legislation, as proposed here, should instead focus on ensuring the realization and implementation of the principles of privacy as a *policy objective* or outcome,[46] as the US Department of Commerce similarly argues.[47] If privacy laws are too prescriptive, as Hirsch argues, they could impede technological innovation for protecting privacy [2006, p. 36]. Likewise, as the US Department of Commerce also points out, "by requiring a particular technology, a regulator may preclude the implementation of better privacy solutions and stifle innovation that benefits consumers and the economy".[48] Essentially, "[e]nhanced privacy protection will depend on the development of new technologies" [Hirsch, 2006, p. 36] and the success of PBD (legislation) is dependent on the availability of the technology/solutions to bring about that success. The PBD method or approach to protecting privacy, therefore, benefits from the further development of technology and, as Hirsch emphasizes, "[t]his development requires regulatory methods that encourage innovation, not those that constrain it" [*Ibid*., p. 36]. Furthermore, as the US Department of Commerce points out, in response to the EC's draft General Data Protection Regulation, "granting the [European] Commission the power to specify technical mechanisms may have the significant unintended consequences because technology developments outpace government regulation".[49]

Therefore, the law should not overly prescribe these solutions, in order to prevent the drawbacks of overregulation. While the law should mandate that technology developers/manufacturers must take the necessary steps to implement technical/PBD solutions when designing and developing PITs, it would be advisable for lawmakers not to specifically determine or mandate which are those solutions, and let the responsible industry players and other stakeholders work that out. The decisions on the specific technical measures/solutions and standards should be left open to a multi-stakeholder process.[50] In addition, PBD could potentially benefit from open technical standards and open collaboration/open innovation. Technology developers/manufacturers and service providers should also be allowed to collectively and/or individually select and develop their own solution, as long as it is strictly based on the relevant fundamental privacy principles and applicable laws.

The smart implementation of privacy protection measures will require 'smart regulations'. If indeed written smartly, regulations need not slow or halt the innovation of even better PBD solutions. Accordingly, PBD, and the privacy laws thereof, should adopt the next-generation regulatory approach, as opposed to an overly prescriptive command-and-control approach [see Hirsch, 2006]. Next-generation standards, such as 'performance-based standards' for promoting innovation in environmental

[46] Of course, it will be important to clarify how the policy objectives or outcomes are specified and thus enforced objectively.

[47] See US Department of Commerce, Informal Comment on the Draft General Data Protection Regulation and Draft Directive on Data Protection in Law Enforcement Investigations (16 January, 2012).

[48] *Ibid.*

[49] *Ibid.*

[50] *Ibid*.

protection [Porter and van der Linde, 1995], which move away from both design standards and rate-based standards, are not based on reference technologies and may, therefore, potentially help to promote the innovation of PBD solutions for protecting privacy by encouraging companies to select/develop their own methods [Hirsch, 2006, pp. 38-40]. Environmental Management Systems (EMS) may also offer a helpful model for the protection of privacy and the implementation process of PBD, as argued by Hirsch, since EMS often entails continuous improvement practices [Hirsch, 2006, pp. 60-63].

Furthermore, in order to address some of the challenges of measuring the effectiveness of the implementation/enforcement of PBD legislation and to determine if the established objectives/targets have been met, measurable and viable indicators will need to be identified/established and data will need to be gathered to populate the selected indicators. Some of the indicators may include, for example, the number of developers/manufacturers fully aware of the new PBD requirements, the reduction in the number of reported/known privacy violations, the number of successful/unsuccessful PBD certification audits per sector and type/category of PIT, and the average/estimated level of investment in PBD solutions.

To address some of the other challenges regarding the necessary investment, research-funding programmes could fund studies that aim to identify and address the needs for developing concrete, specific and viable PBD solutions. In addition, the European Commission, for instance, should continue to fund projects that aim to facilitate the interplay between various stakeholders and actors (i.e. engineers, programmers, designers, lawmakers, regulators, policymakers, privacy commissioners, privacy officers, lawyers, certification bodies, certified PBD trainers, privacy certification auditors, research bodies, data controllers/processors, service providers, law enforcement agencies, privacy law scholars and social scientists). The projects should aim to preliminarily establish best practices, standards and a roadmap for promoting PBD. Companies, research centers and other actors could also receive public funding to develop and validate a variety of PBD solutions for the most threatening PITs, and then identify and exchange best practices and lessons learned for implementing PBD solutions, based on established facts/evidence and pilot demonstrations. This could also help to provide the required inspiration, driving force and knowledge/evidence for PBD lawmaking/policymaking. Subsequently, public funding could also be made available to establish dedicated PBD certified training programs to guide computer programmers/engineers and to communicate the identified best practices and lessons learned. Civil society and privacy commissioners, supported by grass roots activism, could also help to advocate for the necessary greater public and private investment and cooperation in the R&D of PBD solutions [see Cavoukian, 2009]. Furthermore, since governments are significant buyers of PITs, the adoption/implementation of policies in support of the public procurement/pre-commercial procurement of privacy-friendly devices and systems could set a good example and further influence the design and development of future PITs.

9 Concluding Remarks

PBD may be an important solution for ensuring the protection of the right to privacy, and arguably new laws should mandate that *both* technology developers/manufacturers and data controllers must implement PBD measures. But, as a result of the significant challenges and difficulties of implementing and enforcing PBD and the limitations and constraints of the approach, PBD is certainly not a panacea for defending privacy, in light of the increasing development and deployment of PITs. Laws or legal solutions do not perfectly regulate human behavior and neither do technologies or technical solutions. Some of these challenges, limitations and constraints of PBD, however, can be potentially addressed through the application/implementation of 'smart regulatory approaches' and the investment in necessary resources and training.

Nonetheless, the dire reality is that the serious threats/risks to privacy (and liberty) posed by the inertia of technological development is probably a dilemma simply too immense for PBD or any legal or technical solution alone. In the end, no matter how PITs are designed/developed, their widespread deployment and use will likely always be a serious cause for concern for the protection of privacy and liberty.

References

Agre, P.E., Rotenberg, M. (eds.): Technology and Privacy: The New Landscape. MIT Press (1997)

Albrechtslund, A.: Ethics and technology design. Ethics and Information Technology, 63–72 (2007)

Borking, J.: Assessing investments mitigating privacy risks. In: Mommers, L., Franken, H., van den Herik, J., van der Klaauw, F., Zwenne, G.-J. (eds.) Het Binnenste Buiten; Liber Amicorum Ter Gelegenheid van Het Emeritaat aan Prof. dr. Aernout, H.J.Schmidt, Hoogleraar Recht en Informatica te Leiden, eLaw@Leiden, pp. 255–273 (2010)

Cannataci, J.A.: Recent developments in privacy and healthcare: Different paths for RFID in Europe and North America? International Journal of RF Technologies 2, 173–187 (2011)

Cave, J., van Oranje, C., Schindler, R., Ahehabi, A., Brutscher, P.H.-B., Robinson, N.: Trends in connectivity technologies and their socio-economic impacts, Policy Options for the Ubiquitous Internet Society. Final Report. RAND Europe (July 2009)

Cavoukian, A.: Privacy by Design (2009)

Cavoukian, A.: 7 Foundational Principles of Privacy by Design (January 2011), http://www.ipc.on.ca/images/Resources/7foundationalprinciples.pdf

Dommering, E.: Regulating technology: Code is not law. In: Dommering, E.J., Asscher, L.F. (eds.) Coding Regulation: Essays on the Normative Role of Information Technology, pp. 1–17. T.M.C. Asser Press (2006)

Grimmelmann, J.: Regulation by Software. Yale Law Journal 114, 1719–1758 (2005)

Guarda, P., Zannone, N.: Towards the development of privacy-aware systems. Information and Software Technology 51(2), 337–350 (2009)

Gutwirth, S., De Hert, P., De Sutter, L.: The trouble with technology regulation from a legal perspective. Why Lessig's 'optimal mix' will not work. In: Brownsword, R., Yeung, K. (eds.) Regulating Technologies, pp. 193–218. Hart Publishers (2008)

Fischer-Hübner, S.: IT-Security and Privacy. LNCS, vol. 1958, pp. 107–165. Springer, Heidelberg (2001)

Hildebrandt, M., Koops, B.-J.: The Challenges of Ambient Law and Legal Protection in the Profiling Era. Modern Law Review 73(3), 428–460 (2010)

Hirsch, D.: Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law. Georgia Law Review 41(1), 1–64 (2006)

Lessig, L.: Code and Other Laws of Cyberspace. Basic Books (1999)

Kim, Y., Lee, S.W., Lee, D.H., Kim, J., Lee, M.W.: Sound Detection as an Aid to Increase Detectability of CCTV in Surveillance System. In: Aykin, N. (ed.) HCII 2007. LNCS, vol. 4560, pp. 382–389. Springer, Heidelberg (2007)

Koops, B.-J.: Criteria for Normative Technology: An Essay on the Acceptability of 'Code as Law' in Light of Democratic and Constitutional Values. Tilburg University Legal Studies, Working Paper no. 007/2007 (2007)

Pasic, A.: Privacy by Design: An industry perspective on the challenges and opportunities of privacy (2011), http://www.eurescom.eu/?id=531

Porter, M., van der Linde, C.: Green and Competitive, Harvard Business Review, 120–134 (September-October 1995)

Posner, R.: The economics of privacy. American Economic Review 71(2), 405–409 (1981)

Reidenberg, J.R.: Privacy Protection and the Interdependence of Law, Technology and Self-Regulation (2000), http://reidenberg.home.sprynet.com/Interdependence.htm

RISEPTIS: Trust in the Information Society: Research and Innovation on Security, Privacy and Trustworthiness in the Information Society. A Report of the Advisory Board RISEPTIS (2009)

The Royal Academy of Engineering: Dilemmas of Privacy and Surveillance: Challenges of Technological Change, London (2007), http://www.raeng.org.uk/news/publications/list/reports/dilemmas_of_privacy_and_surveillance_report.pdf

Schermer, B.W.: Software agents, surveillance, and the right to privacy: a legislative Framework for agent-enabled surveillance, PhD diss. Leiden University Press (2007)

Schneier, B.: Strong Laws, Smart Tech Can Stop Abusive 'Data Reuse', June 28. Wired News (2007), http://www.schneier.com/essay-175.html

Schwartz, P.M.: Beyond Lessig's code for Internet Privacy: Cyberspace Filters, Privacy-Control, and Fair Information Practices. Wisconsin Law Review 2000(4), 743–787 (2000)

Sollie, P., Düwell, M. (eds.): Evaluating New Technologies: Methodological Problems For The Ethical Assessment of Technology Developments. Springer (2009)

Williams, M.-A.: Privacy Management: The Law and Global Business Strategies: A Case for Privacy Driven Design. Innovation and Enterprise Research Laboratory, University of Technology, Sydney (2009)

van Blarkom, G.W., Borking, J.J., Olk, J.G.E. (eds.): The Handbook of Privacy and Privacy-Enhancing Technologies: The Case of Intelligent Software Agents (2003)

Yeung, K., Dixon-Woods, M.: Design-based regulation and patient safety: A regulatory studies perspective. Social Science & Medicine 71(3), 502–509 (2010)

Yueh-Hsuan, W., Chien-Hsun, C., Cheun-Tsai, S.: Safety Intelligence and Legal Machine Language-Do we need the Three Laws of Robotics? In: Yoshihiko, T. (ed.) Service Robot Applications. InTech Education & Publishing (2008), http://works.bepress.com/weng_yueh_hsuan/3

Integrating Anonymous Credentials with eIDs for Privacy-Respecting Online Authentication

Ronny Bjones[1], Ioannis Krontiris[2], Pascal Paillier[3], and Kai Rannenberg[2]

[1] Microsoft Corporate, Belgium
ronny.bjones@microsoft.com
[2] Goethe University Frankfurt, Chair of Mobile Business & Multilateral Security, Grueneburgplatz 1, 60323 Frankfurt, Germany
{ioannis.krontiris,kai.rannenberg}@m-chair.net
[3] CryptoExperts, Paris, France
pascal.paillier@cryptoexperts.com

Abstract. Electronic Identity (eID) cards are rapidly emerging in Europe and are gaining user acceptance. As an authentication token, an eID card is a gateway to personal information and as such it is subject to privacy risks. Several European countries have taken extra care to protect their citizens against these risks. A notable example is the German eID card, which we take as a case study in this paper. We first discuss important privacy and security threats that remain in the German eID system and elaborate on the advantages of using privacy attribute-based credentials (Privacy-ABCs) to address these threats. Then we study two approaches for integrating Privacy-ABCs with eID systems. In the first approach, we show that by introducing a new entity in the current German eID system, the citizen can get a lot of the Privacy-ABCs advantages, without further modifications. Then we concentrate on putting Privacy-ABCs directly on smart cards, and we present new results on performance, which demonstrate that it is now feasible for smart cards to support the required computations these mechanisms require.

1 Introduction

A number of countries have already introduced or are about to introduce electronic identity cards (eID) and drivers licenses. Electronic ticketing and toll systems are also widely used all over the world. As such, electronic devices become widespread for identification, authentication, and payment. Several European Union countries have already rolled out electronic ID cards and several others have committed to rolling out electronic ID cards and are in various stages of planning [1]. The increasing number of electronic identity management infrastructures are creating opportunities for pan-European initiatives of trustworthy services in e-government and e-commerce and set the basis to overcome fragmentation, closed solutions and lack of user control and transparency [2].

As an authentication token and personal data source, an eID card is a gateway to personal information. This implies a set of risks to the privacy of the citizen, via the unwanted disclosure of personal information and its subsequent misuse.

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 111–124, 2014.

These privacy risks could become even more prominent in the future, if citizens would be using their eIDs not only for e-government services, but also in e-commerce for shopping online, checking into hotels, renting cards online, opening bank accounts, etc.

A recent position paper issued by ENISA on "Privacy Features of European eID Card Specifications" [3] underlines this need for "privacy-respecting use of unique identifiers" in emerging European eID cards and mentions that countries such as Austria and Germany have taken some important steps in this direction. However, some important security and privacy threats still remain. In this paper we take as an example the German eID card, since many consider it to be the most advanced eID deployment [4] and we discuss three of these threats.

Technologies that can help to enhance existing eID card privacy functions are based on privacy-enhanced attribute-based credentials (Privacy-ABCs). In particular, Privacy-ABCs can help prevent monitoring and profiling of the citizens based on the usage of the eID cards, enforce minimal disclosure, offer the choice of complete anonymity for the user, but also help improve the scalability of the underlying infrastructure.

However, although these technologies have been available for a long time, there has not been much adoption in mainstream applications and eID card implementations [3]. We identify three reasons for this: first the available technologies based on Privacy-ABCs use different terminology for their features and even different cryptographic mechanisms to realize them, resulting in a difficulty for developers to understand, compare and use them. Second, the performance of Privacy-ABCs on smart cards (like eIDs) was poor and did not allow practical deployment. And third, Privacy-ABCs are very complex and hard to understand for non-specialists.

Since then, a lot of progress has been made in addressing the above problems. The goal of this paper is to describe this progress and show that Privacy-ABCs are now attractive to be incorporated in eID solutions. In particular we report on the progress being made by the EU-funded project ABC4Trust in bringing together different Privacy-ABC technologies and abstract away their differences. Then we discuss how one of these technologies (namely U-Prove) can be integrated with the German eID card, given the current infrastructure, in order to show that today's eID systems can enjoy some of the benefits of Privacy-ABCs. Finally, we report on the new results we got from experimenting with U-Prove directly on contactless smart cards, indicating that both issuance and presentation can be brought down to the order of milliseconds, making Privacy-ABCs perfectly practical on smart cards.

The rest of the paper is organized as follows. Section 2 analyses the current solution for authentication through the German eID card and discusses the most important privacy and security threats that relate to this. Section 3 introduces Privacy-ABCs and shows their significant potential in addressing these problems. Section 4 shows a case of how U-Prove can be integrated with the German eID system and finally Sect. 5 takes a step further and discusses the possibility of putting Privacy-ABCs directly on smart cards.

2 Current eID Solutions for User Authentication

The German eID card translates privacy into a set of features. First of all, services must authenticate themselves to citizens. The possibility to choose certain attributes, so that the user can controls the transmission of his/her data is another important feature. Moreover, citizens must consent to every access. On-card verification supports uses such as age verification, while releasing minimum information. Finally, restricted identification creates service-specific pseudonyms that are unlinkable across services [4].

However, the authentication scheme based on the German eID card still raises security and privacy concerns. In Sect. 2.2 we elaborate on them, but before that we need to understand the entities that are involved in the authentication protocol, as well as the steps of the protocol. We do so in the following subsection.

2.1 The eID Function

The German Federal Office for Information Security's technical guideline TR-03127 [5] specifies the eID card system's architecture. Figure 1 shows an overview of this architecture for online authentication. Three main components participate in the protocol, namely the *user*, the *service provider* and the *eID server*.

The user wants to use an online service through the use of his browser. For that, he must provide part of his personal data to the service provider in order to authenticate. For that purpose, he uses his personal eID and the accompanying software on his personal computer. The service provider offers online services that can be used only by authenticated users. For authenticating the user, the service provider uses the services of a trusted eID server, through which it can query the data in the eID card of the user. The eID server operates as an Identity Service Provider and answers requests for the personal data of users by service providers. It might be operated by the service provider itself or by a third party as an external service. In the latter case, the eID server offers its services to the service providers who want to support the eID functionality within their Web applications. In this case, the eID server reads the data on the eID that are required by the service provider. Furthermore, it stores and manages the authorization certificates and revocation lists.

Figure 1 shows the involved entities, as well as the phases of the authentication process that are executed, when a citizen wants to use an online service and authenticates to the service provider through her eID. As the figure shows, the authentication process takes place according to the following steps:

1. The citizen wants to authenticate with the use of her eID card to the service provider. The service provider forwards the authentication request to the associated eID server. Corresponding to that, the user is presented with the list of functions and data that the service wants to read.
2. A secure channel between the eID server and the eID client is established by use of cryptographic protocols (PACE, terminal and chip authentication).
3. The eID client displays the requested data to the user.

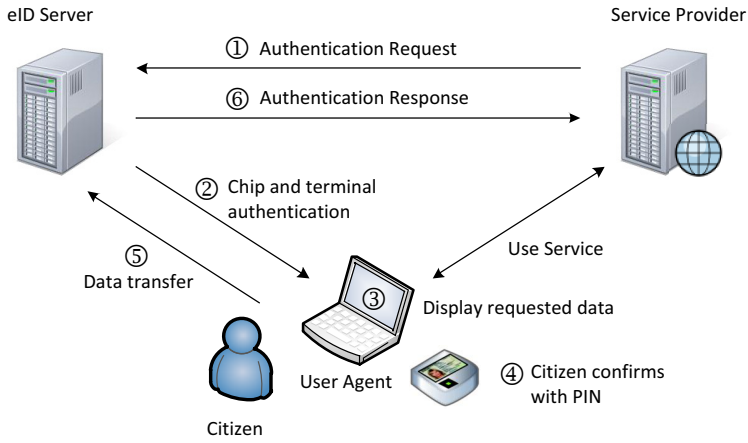


Fig. 1. The steps of online authentication to a service using the German eID card

4. After reviewing the service information and restricting which data the service provider is allowed to get, the user enters her eID PIN to express consent.
5. The data is transmitted to the eID server. The eID server reads the subset of the eID data according to the corresponding authorization, on behalf of the service provider.
6. The eID server forwards the data back to the service provider as part of the authentication response. Corresponding to that, the service provider verifies the results and decides whether the authentication is successful.

From the process described above, we should note that it adheres to two important privacy features, namely notice and selective disclosure. Indeed, the user is duly informed of the scope of the transaction, i.e., of which identity information is transferred to the application owner and for which purposes the data will be processed. The user is also given the possibility to decide on which identity attributes to disclose and to what extent. These features are in accordance to the privacy requirements of electronic ID cards, as defined by ENISA in 2009 [6].

2.2 Security and Privacy Problems

When compared to the privacy features offered by other European eID card specification [3], the German new eID card is one of the most privacy-friendly solutions. However, it follows the passive authentication protocol with bearer tokens that we described in the previous section. Bearer tokens (security tokens) containing user's claims are delivered by the eID server to the service provider without user intervention. This model is subject to several threats [7]. Here we will focus on the most important ones, relevant to security, privacy and availability.

eID Server Knows All User Transactions. Even though the eID server does not necessarily need to know where the user is authenticating and which service she is requesting, this knowledge is passed by design to the eID server in the current eID solution. More specifically, the eID server is involved each time a user authenticates to a service provider using her eID, and is able to keep track of the user actions. This enables the eID server to trace and link all communications and transactions of each user.

This pattern is followed in most federated identity management systems today, and it can be also observed in STORK's architecture for eID interoperability between different European countries [8]. In the physical world we might have to show a government issued picture ID on different occasions. The issuer of those picture IDs is not aware that we show those at this specific location. In the digital world however, the default case is that the issuer knows when you present your ID.

eID Server Knows All Customers of the Service Provider. Reversing the above threat, the involvement of the eID server in every user authentication constitutes a privacy threat for the service providers as well, since the eID server learns all the customers trying to access a service. Especially if the eID server is operated by a private company, it might be a competitive threat, if it can learn all the customers of another company (i.e. the service provider).

User Impersonation. Since the user does not perform an active role in the information exchange between the eID server and the service provider, there is a high security risk of user impersonation by insider attackers at the eID server or outsider intruders when they would gain access to the eID server's resources.

An eID server under control of an attacker (insider or outsider) has the ability to impersonate every user at applications using eIDs for authentication. For example, insiders can copy or alter user's credentials and as such steal the identity of a user. In general, in a federation scenario, the insiders or outsiders who learn a user's credentials can impersonate the user and get access to the assets at different applications involved in the federation.

Availability. The eID server becomes a business critical component as it is needed for every transaction the user does with the applications. Denial of Service attacks towards the eID server will impact all applications using the service. Attacking this component may have a huge economic impact because the attack spreads over different services.

All of the above problems become critical when there are currently only a couple of eID servers operating, despite the view of the German government that this service will be offered by multiple servers. Meanwhile, the requirement that the eID providers are not able to track the behaviour of eID holders is becoming more prominent. In the evaluation assessment of the recent proposal of a Regulation "on electronic identification and trusted services for electronic transactions in the internal market" [9] it is stated that a solution to this tracking problem should be aligned with the current ongoing revision of the Data

Protection Directive and include specifically privacy-by-design rules. In the next section we discuss specifically how the above threats can be addressed with the privacy-by-design model.

3 Privacy-ABCs to the Rescue

To alleviate the above threats and offer more flexibility, governments can turn to a claims-based architecture [10]. The claim-based architecture is a design pattern used by system architects to implement claims-based identity. The main purpose of claims-based identity is to externalize authentication. The service provider's interest is not to authenticate the user, but rather receive verified claims about the user, based on which access to the service is decided. That is, the service provider publishes a policy on accessing a specific resource and expects to receive claims and identity tokens from trusted sources that satisfy this policy. The trusted sources that issue such security tokens are the *identity service providers* (IdSP), sometimes also called identity providers for simplicity. In the particular eID system that we are studying in this paper, the eID server has the role of the IdSP.

The claimed-based architecture allows separation between service providers and identity service providers, so that there is no direct exchange of information between them. Instead, the user lies in the middle, having control of the exchange of his identity information. Then, on one side identity providers authenticate the user and issue security tokens, and on the other side service providers consume tokens. Because a service provider relies on the IdSP to provide authentic information about the user, it is called the *relying party* (RP).

An example of claim-based architecture is the Identity Metasystem [11]. Claim-based architectures can use privacy-respecting credential systems (Privacy-ABCs) to provide untraceability and minimal disclosure. Examples of such credential systems are Idemix [12] and U-Prove [13]. Over the last few years, Idemix and U-Prove have been developed to offer an extended set of features, even though these features are named differently and they are realized based on different cryptographic mechanisms. Recently, the European research project ABC4Trust [14], was initiated with the goal to alleviate these differences and unify the abstract concepts and features of such mechanisms. In particular, it brings them under the common name *privacy-preserving attribute based credentials*, or *Privacy-ABCs* [15]. So, Privacy-ABCs are privacy respecting credentials that are defined over these concepts and features and are independent from the specific cryptographic realization beneath. Overall, Privacy-ABCs offer the following advantages [15]:

- Privacy-ABCs are by default untraceable. Even when they are obtained on-demand, IdSPs are not able to track and trace at which sites the user is presenting the information.
- Privacy-ABCs can be obtained in advance and stored by the user while still being able to disclose the minimal amount of information needed for a

particular transaction. So, the real-time burden of the IdSP is diminished, improving scalability.

- To prevent identity theft and "credential pooling", i.e., multiple users sharing their credentials, credentials can be bound to a *secret key*, i.e. a cryptographically strong random value that is assumed to be known only to a particular user. A presentation token derived from such a key-bound credential always contains an implicit proof of knowledge of the underlying secret key, so that the verifier can be sure that the rightful owner of the credential was involved in the creation of the presentation token. As an extra protection layer, the credentials can also be bound to a trusted physical device, such as a smart card (i.e. the eID card itself), by keeping the secret key in a protected area of the device. That is, the key cannot be extracted from the device and so it is not possible to make a presentation proof without the device.
- Instead of complete anonymity, if desired, users can generate an unlimited number of pseudonyms or a batch of Privacy-ABCs and use them at the same or different relying parties. Presentations of pseudonyms or different Privacy-ABCs are cryptographically unlinkable, meaning that given two different presentations of the credentials, one cannot tell whether they were generated by the same user. In cases where it is undesirable that users are able to generate multiple identities on the same site, the relying party can impose a *scope-exclusive pseudonym*, meaning that for a scope string (e.g. a URL) the user can only register a single pseudonym. This feature is useful in applications where the user should not be able to create multiple identities based on a single credential, like for example in online petitions.

So, privacy-ABCs have significant potential to enhance existing eID card privacy functions. Their integration is perfectly realizable today, and does not necessarily require modifications at the current infrastructure of the eID server and the eID cards. This is demonstrated in the next section, where we take as a paradigm one of the Privacy-ABC technologies, namely U-Prove, and show how it can be integrated in the German eID system.

4 Integrating Privacy-ABCs to Existing eID Systems

U-Prove has been integrated with the German eID system and it has been demonstrated in a typical e-Participation scenario [16]. In particular, it was demonstrated in a local referendum application, where the citizens had to prove their eligibility to participate, by proving properties of their identities, while at the same time their anonymity is preserved. In this section, we generalize the discussion and show the entities and the protocol involved. Even though in our discussion below we still use U-Prove, the same would apply for other Privacy-ABC systems as well (e.g. Idemix). We only use U-Prove here as an instance, since it was used in the initial implementation [16].

Compared with the standard German eID system we discussed in Sect. 2, the entities remain the same only that a new entity has been introduced and in

particular the U-Prove issuer. The U-Prove issuer has two responsibilities: first to validate the claims issued by the eID server and second to issue a U-Prove token that contains these claims. By deploying the U-Prove issuer, applications can leverage U-Prove Tokens providing unlinkability and anonymity to the users.

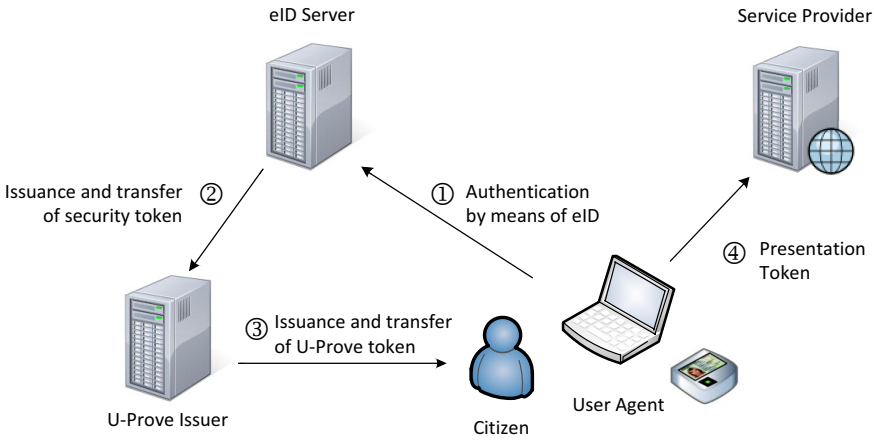


Fig. 2. U-Prove Integration with the German eID system

As Fig. 2 shows, the authentication process takes place according to the following steps:

1. The user wants to authenticate by means of eID card to the application. For that, the user is redirected to the eID server of her choice and she is prompted to present the eID card and PIN-code. The eID server then validates the identity of the user.
2. The eID server produces a security token containing the claims on user attributes that are requested by the application. This token arrives at the U-Prove issuer.
3. At the U-Prove issuer, the current security token is exchanged for a U-Prove token, after the U-Prove issuer has validated the received token.
4. The user presents the U-Prove token to the application through a U-Prove Presentation Proof.

The above protocol achieves a challenging combination of high assurance on the identity of individuals through their eID and full anonymity while using a service. User anonymity is made possible since the presentation token cannot be tied back to the true identity of the user. The true identity of the user was established during the authentication at the eID server, but at Step 3, the eID token was exchanged for a U-Prove token. U-Prove ensures the unlinkability between the issuance of the U-Prove token and its usage through the presentation

proof (Step 4). Even if the U-Prove issuer and the service provider collude, there is no way to link the two together. Actually, this offers the advantage to assign an extra role to the U-Prove issuer, if desirable: that of the validator of the U-Prove tokens at the RP. That would alleviate this extra burden from the RP without having to introduce an extra entity in the architecture and without having to make any compromise in terms of security and privacy.

The example above illustrates how Privacy-ABCs and eID systems can be combined. The idea of having high assurance on the identity by means of a smart card and being anonymous in the actual transaction on the RP sounds esoteric, but it can be easily accomplished by means of Privacy-ABCs. This combination is possible by leveraging e.g. a service in the cloud. This service will learn the user's attributes coming from the eID server but will not learn where the user is using them. The U-Prove issuer will learn as much information as the eID server and so both services are equal from a privacy threat modeling perspective and should be protected in a similar way.

However a few issues of trust management need to be addressed, if a new type of entity is introduced into the system. The U-Prove issuer learns about the attributes that are requested and so (over time) can build a profile of the user and the attributes that the user needs for her service at which time, e.g. (being adult, asking for the respective credential Friday night, having been checked for AIDS recently with no AIDS having been detected, being eligible for medical consultation via a special type of assurance).

One could say, that the cleanest solution would be to simply regulate the U-Prove issuer to not store any attributes after credential issuance and to audit him for this. However experience with ID issuers raises doubt, that a "no-records-taken"-policy would be accepted by all stakeholders, who want to look after the U-Prove issuer and who may require some record-keeping, though the authors do not really see a hard reason for record keeping. Quick re-issuance of credentials after a user has lost them does not seem to be so important, that it would justify record-keeping with that many privacy implications.

Actually, one measure to mitigate the privacy risk at the U-Prove issuer lies in the heart of Privacy-ABCs philosophy. The eID server should always issue tokens containing claims for *all* attributes in the eID cards, and let the user decide which of these attributes to reveal to the RP, during the presentation proof. In this way, the eID server cannot draw any conclusion on the type of the application the tokens are being used for. At the same time, it is important that several U-Prove issuers are available; this allows users to spread the knowledge of certain types of attributes over their selection of providers.

Besides privacy reasons, the organizational setting of the U-Prove issuer needs to assure that no monopoly situation can arrive, as a monopoly would also be risky from an availability and cost perspective. While monopolies can achieve scale effects in network-based industries and therefore can have cost advantages, the cost risk from the point of user is the lock-in situation, that comes with the monopoly and that allows the monopoly provider (in this case the issuer) to dictate prices. Actually a U-Prove issuer could and should fall under the rulings

of the recently proposed regulation on electronic identification and trusted services for electronic transactions in the internal market [17]. Article 11 (4) therein explicitly mentions pseudonyms (i.e. a special type of attributes), which could be issued by eID servers (and also U-Prove issuers). The requirements set out in the proposed regulation can be expected to establish appropriate trust into the token issuers.

The most important aspect of the U-Prove issuer is that its cloud instances do not learn the relationship between the user and RP. It is actually this relationship that affects the privacy of the user because it makes profiling of the user possible. It is in the interest of the eID server to apply Privacy-ABCs, in order not to be seen as a "monitoring beacon". Privacy-ABCs will also protect the privacy of the RPs because now there is no third party (eID server) which learns all their customers. Especially when the eID server is run by a private company, learning all the RP's customers is seen as a competitive threat for the RP.

Germany has gone a long way in adding privacy to the eID card, much further as any other system. While this is certainly to the right direction, most eID systems being deployed in Europe would benefit even more from Privacy-ABCs. For that, it is crucial that Privacy-ABCs become part of the eID card itself. So the delivery of the claims to the RP is done under control of the user and the eID card. This removes immediately many of the threats discussed in Sect. 2.2 and increases the privacy of the user and RP. Can we put Privacy-ABCs on eID cards? Could they run efficiently on the smart cards? This will be discussed in the next section.

5 Privacy-ABCs on Smart Cards

There have been several approaches to implement Privacy-ABCs on smart cards. Bichsel [18] and Balasch [19] focus on providing the arithmetic functionality required, i.e. fast modular arithmetic. Balasch implemented the arithmetic using AVR microcontrollers, whereas Bichsel used the JCOP platfom. Later, Bichsel et al. presented the first practical implementation of a Camenish-Lysyanskaya-based Direct Anonymous Attestation scheme on a Java Card 2.2.1 [20] with a performance close to 7.5 seconds. Tews and Jacobs [21] considered U-Prove and succeeded in performing a presentation proof in about 5 seconds for 2 attributes and 8s for 4 attributes. Batina et al. [22] suggest to use self-blindable certificates and put forward an implementation that requires about 1.5s to perform presentation for 1 attribute. In 2011, Mostowski and Vullers implement U-Prove on a MULTOS card and reach about 0.5s (resp. 0.8s) for 2 (resp. 5) attributes. Up to our knowledge, no implementation of a Privacy-ABC system is available on a contactless smart card at this time.

We have chosen to focus on the full-fledge version of U-Prove as opposed to the device-binding version [23], thus showing the applicability and user-friendliness of a complete Privacy-ABC system running on an eID card. The chosen smart card platform is a 32-bit chip made available by Invia [24]. The component features a Sparc v8 Leon II core and embeds a lightweight public-key coprocessor

called MEXPA running at 33MHz. We assumed an ISO/IEC 14443 contactless interface running at a pessimistic baudrate of 106 Kbits per second.

U-Prove describes an issuance phase and a presentation phase. The two protocols may employ either a group of integers modulo a prime number or an elliptic curve defined over a field of large prime characteristic. We have considered the case of an implementation based on elliptic curves for increased flexibility at the algorithmic level, although some of our optimizations are readily applicable to groups of integers.

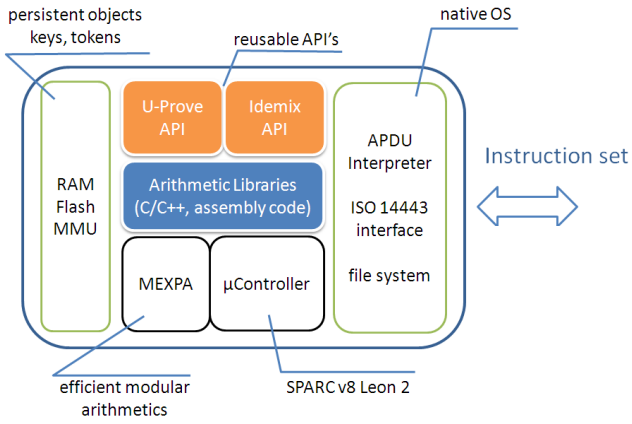


Fig. 3. Schematics of a contactless smart card integrating U-Prove and eID application

At a high-level view, we have undertaken and combined the following approaches:

1. reformulation of U-Prove's protocol flow to identify the critical operations performed by the smart card and minimize on-board computations;
2. maximal delegation of unsensitive computations to the Issuer and Verifier within the limits of neutrality towards cryptographic security;
3. off-line/on-line optimization using a maximal number of precomputed values stored in non-volatile memory as coupons. Coupons are internal variables that can be generated by the smart card beforehand, thus reducing the total latency when transactions (either issuance or presentation) take place;
4. boosting point operations using the best suited coordinate system (Jacobian, affine or mixed Jacobian/affine) for each operation performed on the curve;
5. optimizing scalar multiplications by aggregating multiple products: computing a double multiplication $([k_1]B, [k_2]B)$ with the same basepoint B allows to share intermediate variables even when B changes from one execution to the next;
6. finely compare implementation strategies and trade-offs given the performance of low-level hardware operations to find optimal settings.

U-Prove Issuance. Further optimizations are made possible by the use of NIST curves as recommended in the specifications of U-Prove [23]. Taking the NIST curve P-256, we get an estimated cost of about 4.1 million clock cycles for the most critical part of the issuance phase, neglecting modular additions and subtractions. The memory size required remains moderate, namely of the order of 1KB of RAM. Under a clock frequency of 33MHz, 4.1 million cycles correspond to 124ms. We add a 30% overhead to take into account modular additions, pointer management and minor CPU-operated instructions. A pessimistic additional 20ms is added to reflect one-time minor operations such as hash computations.

The total bitsize of transmissions in the issuance protocol amounts to $10 \log_2 p$ where p is the field characteristic which, assuming a baudrate of 106 Kbps (slowest configuration), leads to an estimated 78ms. Putting it altogether, and neglecting the execution time of off-board operations, we end up with an expected running time of about 259ms for the complete issuance protocol.

Presentation Proof. For the presentation phase, we take the typical case where the user generates the presentation proof, in which some attributes are disclosed and some are not. Similar to the issuance phase, there are many possible algorithmic options for this phase as well and we may rely again on precomputations (coupons) and various algorithmic optimizations of aggregated scalar multiplications.

Overall, we find that the cost of the presentation proof amounts to $38.42 \times (n - |D|)$ milliseconds on the target chip, where $n - |D|$ is the total number of undisclosed attributes. We upper bound the extra time needed by the remaining computations by about 45 to 50ms. This gives a typical presentation phase of 434ms for 10 unrevealed attributes, thus providing evidence that both the issuance and the presentation phase of U-Prove can be efficiently implemented on a state-of-the-art contactless card.

Areas for Further Optimization. Operations on the elliptic curve could be made faster by using efficiently computable endomorphisms over the group of points as with the GLV method [25]. For instance, curves over a field extension allow to use the Frobenius map to speed-up scalar multiplication. Also, curves with coefficients $a = 0$ or $b = 0$ that have endomorphisms that one can evaluate using roots of unity are quite appealing (other examples with specific values for a and b are known). Also, field arithmetics can be boosted using extended fields. Taking a curve over $\mathbb{F}_{p^2} = \mathbb{F}_p/(p^2 + 1)$, a field multiplication which usually requires two operands of n bits boils down to 3 multiplications with half-size operands and one can replace a modular reduction from $2n$ bits to n bits with two reductions from n bits to $n/2$ bits.

On a general-purpose 32-bit CPU, taking a field extension with a pseudo-Mersenne characteristic and a sparse irreducible polynomial would probably be the best possible choice as one can rely on both fast arithmetics and efficient endomorphisms. Also, selecting an Edwards curve would slightly improve speed. As investigated in [25], the best timings on a 64-bit Intel processor when no crypto-coprocessor is available are realized with Edwards curves over $\mathbb{F}_{p^2}$ using endomorphisms as per the GLV and GLS techniques.

6 Conclusions

A potential future deployment of Privacy-ABCs in eID schemas would allow going beyond the existing privacy-preserving capabilities of the German model. In this paper we have showed the benefits of such an integration in terms of preserving the privacy of the user. Overall, based on several properties, Privacy-ABCs bear a high potential to challenge what must have been considered as necessary processing of personal data in the past. If deployed broadly, Privacy-ABCs would enable the revision of the understanding of necessary processing and require reassessment of existing systems. Integrating them into the upcoming European framework on electronic identification and trust services for electronic transactions in the internal market seems possible, though further details need to be analysed. For example, one could extend further the discussions in this paper on establishing the appropriate trust on the token issuers, as well as continue the analysis of further optimizations of the performance on smart cards, as discussed in the last section.

Acknowledgements. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement no. 257782 for the project Attribute-based Credentials for Trust (ABC4Trust).

References

1. Ahlswede, S., Gaab, J.: eIDs in Europe, Deutsche Bank Research. Tech. Rep. (September 2010)
2. A Strategy for ICT R&D and Innovation in Europe: Raising the Game. Commission Communication, COM 116 (2009)
3. Naumann, I., Hogben, G.: Privacy Features of European eID Card Specifications, ENISA, Position Paper (January 2009)
4. Poller, A., Waldmann, U., Vowe, S., Turpe, S.: Electronic identity cards for user authentication – promise and practice. IEEE Security & Privacy 10, 46–54 (2012)
5. Architecture electronic Identity Card and electronic Resident Permit, German Federal Office for Information Security. Technical Report TR-03127, Version 1.13 (2011)
6. Naumann, I.: Privacy and Security Risks when Authenticating on the Internet with European eID Cards, ENISA, Risk Assessment Report (November 2009)
7. Bjones, R.: Architecture serving complex Identity Infrastructures, Trust in Digital Life. Tech. Rep. (November 2011)
8. Krontiris, I., Leitold, H., Posch, R., Rannenberg, K.: eID Interoperability. In: Fumy, W., Paeschke, M. (eds.) Handbook of eID Security. Publicis Publishing (2011)
9. Impact Assessment accompanying the proposal for a regulation of the European Parliament and of the council on electronic identification and trust services for electronic transactions in the internal market. In: European Commission, SWD, 136 (2012)

10. Cameron, K., Posch, R., Rannenberg, K.: Proposal for a common identity framework: A User-Centric Identity Metasystem. In: Rannenberg, K., Royer, D., Deuker, A. (eds.) The Future of Identity in the Information Society – Opportunities and Challenges. Springer (2009)
11. Cameron, K., Jones, M.B.: Design Rationale behind the Identity Metasystem Architecture. Microsoft. Tech. Rep. (February 2006)
12. Camenisch, J., Van Herreweghen, E.: Design and implementation of the idemix anonymous credential system. In: Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002), pp. 21–30 (2002)
13. Brands, S.: Rethinking Public Key Infrastructures and Digital Certificates; Building in Privacy. MIT Press (2000)
14. ABC4Trust: Attribute-Based Credentials for Trust, `https://abc4trust.eu`
15. D2.1 Architecture for Attribute-based Credential Technologies - Version 1, ABC4Trust, Deliverable D2.1 (2011)
16. Bjones, R.: eParticipation Scenario Reference Guide. Microsoft. Tech. Rep. (October 2010)
17. Proposal for a regulation of the European Parliament and of the council on electronic identification and trust services for electronic transactions in the internal market. In: European Commission, COM, 238/2 (2012)
18. Bichsel, P.: Theft and Misuse Protection for Anonymous Credentials, ETH Zürich, Switzerland, Master's thesis (2007)
19. Balasch, J.: Smart card implementation of anonymous credentials, K. U. Leuven, Belgium, Master's thesis (2008)
20. Bichsel, P., Camenisch, J., Groß, T., Shoup, V.: Anonymous credentials on a standard java card. In: Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS 2009), pp. 600–610 (2009)
21. Tews, H., Jacobs, B.: Performance issues of Selective Disclosure and Blinded Issuing Protocols on Java Card. In: Markowitch, O., Bilas, A., Hoepman, J.-H., Mitchell, C.J., Quisquater, J.-J. (eds.) WISTP 2009. LNCS, vol. 5746, pp. 95–111. Springer, Heidelberg (2009)
22. Batina, L., Hoepman, J.-H., Jacobs, B., Mostowski, W., Vullers, P.: Developing efficient blinded attribute certificates on smart cards via pairings. In: Gollmann, D., Lanet, J.-L., Iguchi-Cartigny, J. (eds.) CARDIS 2010. LNCS, vol. 6035, pp. 209–222. Springer, Heidelberg (2010)
23. Microsoft, U-Prove Cryptographic Specification V1.1 (February 2011)
24. Invia, Modular Exponentiation IP, `http://www.invia.fr/Modular-Exponentiation-21.html`
25. Longa, P., Gebotys, C.: Efficient techniques for high-speed elliptic curve cryptography. In: Mangard, S., Standaert, F.-X. (eds.) CHES 2010. LNCS, vol. 6225, pp. 80–94. Springer, Heidelberg (2010)

Federated Identity as Capabilities

Harry Halpin[1] and Blaine Cook[2]

[1] W3C/MIT, 32 Vassar Street Room 32-G515 Cambridge, MA 02139, United States
hhalpin@w3.org
http://www.ibiblio.org/hhalpin
[2] 30 Ballater Road, London SW2 5QR, United Kingdom
romeda@gmail.com
http://blog.romeda.org/

Abstract. The problem of federated identity, the ability to sign-in across multiple services, has not been solved in a privacy-respecting or secure manner. We briefly analyze the design of OpenID Connect, as implemented by Google and Microsoft, and BrowserID as implemented by Mozilla Personae. Then we consider a capabilities-based approach to federated identity that posits identity to be a set of capabilities that a user can prove to a service that they possess, such as possession of the capability to check a particular email address. Then we show how we can extend existing federated identity approaches can be re-designed using capabilities verified by the use of key material.

Keywords: identity, capabilities, federation, BrowserID, OpenID.

1 Introduction

As the Web leads to a proliferation of services that each require some form of authentication, the problem of sign-in has returned with a vengeance: Current approaches that federate a single identity across multiple services ("federated identity") solve this problem at the cost of sacrificing user privacy. Due to an inability to solve sign-in and data-sharing in a user-centric manner that is usable by developers, currently the most popular identity system is the non-federated Facebook Connect [13]. Also, most federated identity solutions are not standardized via a standards body like the IETF or W3C, so there is not a single alternative federated model capable of being easily deployed across browsers. As we can not determine the security or privacy characteristics of proprietary solutions or iterate through the large list of federated identity systems produced in both the academic literature and in enterprise computing, we will confine our privacy analysis to two open-source efforts that have attracted attention and user-bases: OpenID Connect [2], as currently implemented by Google and Microsoft, and BrowserID [1], also called Mozilla Personae. After discussing the design of both of these federated identity solutions, we outline a "capabilities-based" approach that considers identity simply to be a set of capabilities that a user can prove to a service that they possess without revealing their identity. Then we show how

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 125–139, 2014.

we can extend existing federated identity approaches such as OpenID Connect [9] and BrowserID [1] with capabilities based on using key material, and outline the future work necessary to integrate such work with anonymous credentials and zero-knowledge proof systems.

1.1 Terminology and Assumptions

In this work, we assume there is a user that is sending some kind of information to a *relying party*, a services that wish to access verified identity claims. The source of the identity claims is called an *identity provider*, a service that stores and can possibly verify identity claims on behalf of a user. The common example would be having a user send their username and password combination to Facebook via Facebook Connect, the identity provider, to sign-on to a third party service such as music-sharing service Spotify, the relying party [13]. Spotify also may require some information from Facebook, such as the full name of the users and the songs they like in their Facebook profile, in order to customize their service. This information required by the relying party from the identity provider are considered *identity claims*.

We assume that users don't care about sign-in, they just want to access a service as quickly as possible. It is well-known from anecdotes that forcing users to register a new account and sign-in leads to almost half of users being so discouraged that they do not even bother to register a new account. Developers shouldn't care either: For most developers, all they want is the ability to authenticate a user in order to provide them with some sort of access to a system, possibly with a persistent state tracked by their system. Unless their particular service happens to be in the business of buying and selling user's personal data - as many "social" services are - the developer simply wants to authenticate and store just enough state about the user in order to have the service fulfill its functionality. This state is usually thought of as a set of verified identity claims, such as that a user's first name is "Bob" and that has the email address *bob@example.org*. These claims are verified in a number of different ways, such as having the user entered them into a form when registering a new account (as is the most common practice today) ranging or having the claimed stored locally encrypted in a device and transferring them with some time-limited scope that includes a digital signature. In general, *identity* can be broadly construed as a set of claims about a user or a particular persona of the user, where a *persona* is a set of claims that exposes only a (possibly "fictional") facet of the user.

We also assume that users would like to minimize the amount of disclosure of personal information to be verified to the relying party. For example, minimizing the collection of personal data is sometimes required by regulatory and government authorities, as exemplified by NIST's "Guide to Protecting the Confidentiality of Personally Identifiable Information" [12]. However, there are also everyday examples that do not involve regulation. For example, a user may want to sign-into a site such as Youtube that features videos that have content that require the user to identify themselves as over a certain minimum age, such as 18. However, we assume the user would prefer not disclose their entire name and

other personal information in this case to the relying party. Not only will we assume that the relying party is potentially untrusted, but we would also like to assume that the trust in the identity provider should be minimized. So, we would like to minimize the amount of information the identity provider observes about the actions of the user at various relying parties, which can be considered the linkability of the transactions with relying parties by the identity provider. We will also note that this can be generalized to any outside passive adversary that is observing the identity provider.

1.2 Related Literature

Indeed, these problems are not new, but have already to a large extent theoretically been solved by the advent of anonymous credentials [4] and accompanying work related to zero-knowledge proofs [8], work that often goes by the name 'attribute-based credentials.' The European Commission has funded a number of large-scale projects in this area that have produced open-source implementations of anonymous credentials with zero-knowledge proof systems, such as *PrimeLife*[1] and *ABC4Trust*[2] as well as commercial systems such as Microsoft's UProve.[3] However, the concepts behind cryptographically anonymous credentials and zero-knowledge proofs so far have not yet been introduced into common federated identity schemes, much less incorporated into the code of popular browsers. Another reason is likely that the conceptual complexity of anonymous credentials and zero-knowledge proofs seems not to match the desire by companies to define the flow of information of personal data. Lastly, the ability to create anonymous credentials with efficient zero-knowledge protocol signature schemes has still been an active area of research, although efficiency gains are very promising [3].

The most likely reason for lack of take-up of these technologies is that these products is not compatible with existing browsers. Currently attribute-based credentials require browser plug-ins, and as such plug-ins are being phased out by browser vendors and are incompatible with the increased speed of browser releases, we are unlikely to see any attribute-based credential scheme work. If it was required by law, it would likely cause more harm than good by tying users to older browsers, as shown in the case of the Korean e-banking certificate requirements that trap users with Internet Explorer and an ActiveX plug-in with security risks. Requiring browser plug-ins has been such a catastrophe in Korea that it has even impacted elections[4] Attempting to avoid any new code being added to browsers has been the strategy of increasingly high-profile federated identity schemes such as OpenID Connect[2] and BrowserID[1], but these schemes have not considered any compatibility with attribute-based credentials or having higher-security and privacy requirements.

[1] `http://primelife.ercim.eu/`
[2] `https://abc4trust.eu/`
[3] `https://research.microsoft.com/en-us/projects/u-prove/`
[4] `http://blogs.wsj.com/korearealtime/2012/11/13/ahn-pledges-to-end-outdated-encryption-standard/`

So in this paper, we focus on existing deployed Web-based identity systems and their common (if simple) attributes such as e-mail verification, although we present e-mail verification and identity claims within a new framework of generalized capabilities verified by digital signatures. Second, our proposed simple extensions of existing federated identity scheme should be capable of using *any* digital signature given that these cryptographic operations are now natively being coded into browsers. We'll investigate also including both conventional digital signatures from identity providers like Facebook and Google as well as the possibility of using the kinds of signature schemes needed by anonymous credentials [5].

1.3 Capabilities

There has long been a debate between access-control lists and capabilities as a way of providing authorization restrictions [6]. As put by Laurie, "a capability can be thought of as a 'handle' representing some operation on some object. Possession of the capability is all that is required to perform that operation" [11]. While it may seem counter-intuitive to remove the explicit identity of the user from an identity-based system, the framework of capabilities is extremely useful as it can also provide a much more fine-grained level of control than per-user access control. Furthermore, this lack of an explicit user identifier is nonetheless possibly a requirement for data minimization principles and anonymous credentials. In the next section, we show that current federated identity-based systems can be considered as proving the possession of a capability by a user (rather than a program), with the possession of an email address and a password to access being the 'handle' that defines the operation of a user authenticating their identity.

2 Identity as Capability to Check Email

2.1 Problems and Example

Currently, despite years of effort around solutions like OpenID [9], the primary form of identity on the Web is still a HTML form for filling out claims, which in return delivers a cookie that tracks a user's session state. Between sessions, users have to enter a password, which is an easily phishable symmetric shared secret. There are two attack vectors, the first from outside attackers and the second from a compromised or malicious identity providers. First, Without proper security considerations (such as enforcing TLS) the cookie that proclaims the session state is easily swiped by a man-in-the-middle attack by an outside attacker. On the other hand, the identity provider itself gets to observe all transactions with any relying party, and may even commit transactions with user notification.

Not only is the security poor, but so is the user-experience: For users, sign-in is not only just a means to an end, it's often a confusing annoyance. For developers, the situation is unsatisfactory: Instead of having to re-implement sign-in code for their services, most developers would prefer to use a commonly available library.

2.2 Using Email Addresses as Identifiers

In practice identity for Web-based services can be reduced to a number of rather simple primitives. Let us walk through a simple example. Assume a user wanting to access a service on a web-site. The user who is visiting a site knows who they are. The service provider does not, yet. The first element of a sign-in system allows the user to tell the service provider who they are, i.e. to allow the user to identify themselves. This is thought of both in federated systems and non-federated systems as requiring the user to present an identifier to the service provider. What the identifier hopes to identify is a single individual human (or perhaps also a personae thereof).

However, natural language names will not work as these can be ambiguous. Assigning numbers (as done with credit cards, telephones, and the like) is a possibility, but are difficult to remember and require a centralized database lest they also suffer from ambiguity. Currently, email addresses of the form *name@domain* seem to be the best distributed system for identity, as they are generally associated with individual humans and tend to be globally ambiguous. Multiple, distributed identities can be created easily by creating a new email address for each identity. Email addresses tend to be rather natural identifiers for individual humans, as they seem to roughly approximate other naming schemes that attach some local context in order to disambiguate a name. For example, "Jesus of Nazareth" is approximately equal to *jesus@nazareth.israel*, or to use a more modern example, "Blaine Cook, 30 Ballater Road, London" can be thought of as *blaine.cook@30.ballater-road.london*. In the global system of the Internet, the domain name is the necessary context needed to disambiguate a string identifier that names an individual person. Although the domain name system is as centralized under as phone numbers or credit card numbers, it offers a number of additional capabilities and a well-understood extension mechanism, i.e. the purchase of a new domain name from a domain registrar and the setting up of an SMTP server.

There are numerous other options for what identifier can be used beyond email addresses. Often a simple string for a user name is all that is required, assuming the user has previously registered with the service. Yet users often forget the particular string they may use as their identifier for a service (or be forced to used, due to some other user registering the same string before them). More importantly, usernames are usually "disguised" email addresses, as they are directly tied to email addresses in the registration process.

Certain systems such as earlier versions of OpenID [9] and WebID [14] have claimed that a URI by itself (such as *http://www.example.org/bob*) is a good identifier, however, in practice these URIs are in general tied to e-mail addresses as well (the owner of the domain of the URI as determined by the email of whoever registered the domain name or controls on behalf of the user, as easily determined by protocols such as WHOIS). Also, as users do not remember URIs and do not imagine that they themselves are URIs, federated identity systems such as OpenID 2.0 based on URIs confused users (and thus, were not used) despite considerable deployment by companies such as Google and Yahoo! [9].

In comparison to using self-signed certificates like WebID [14], email addresses tend to be better than cryptographic credentials stored in the browser as the browser is not what is being identified in a sign-on process: The human is the thing that is being identified rather than the browser, and the human may use multiple browsers across different devices. If a system is based entirely on using cryptographic credentials in the browser but does not include a prompt for an email address, then the user is unable to sign in if they lose their device. This is one of the major problems with losing private keys in general, unless keys are synchronized amongst multiple devices, but in this case the user is at the behest of whoever runs the key hosting service that manages the keys of a user in absence of their client device. This is not to say the approaches are completely incompatible: When a client device goes away (seized by police, lost in a river, or borrowed from a friend) and with it the user's cryptographic tokens, a service can still use e-mail addresses to identify a user and authenticate claims, and even revoke and re-provision key material for client devices.

2.3 Checking E-mail as a Capability

The last step is to authenticate the user, in which the service needs to prove that the user is who they say they are. This is done by proving that the user controls the e-mail address and can respond to an email, usually a temporary URI embedded in an email with some secret code. This also addresses concerns that certain email addresses can be checked by more than one person, such as a couple (such as "Sheri" and "Stan" in *sherinstan@gmail.com*) or that e-mail addresses can be recycled by different users over time (as done by Yahoo!), but these are cases can be dealt with by determining if someone can check the email. Thus, the vast majority of the time, all sign-in systems do from a the perspective of security is validate that a user controls an email address. Traditional username-password combinations are an often unnecessary convenience as if a site has a link to recover a forgotten password via checking for an email from the service, then the password saved on the site for a user is irrelevant from a security perspective, since anyone with access to their email account can reset the password. Whenever a site uses a *username-password*, this almost always becomes a *email-password-reminder* token. The majority of existing identity systems in the wild are at their foundation simply systems that prove the user has the capability to check email.

3 Privacy Issues in Federated Identity Systems

Indeed, existing federated identity systems simply hope to re-use the same identity, an *email-password-reminder* token, across different sites. There are two different systems currently under development that have attracted attention: OpenID Connect [2] and BrowserID [1]. These two systems are not designed (or at least thought to be) capability-based systems, and do not vary in terms of

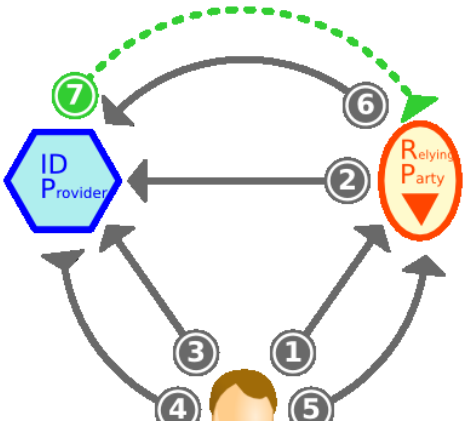


Fig. 1. Flow of OpenID Connect (OAuth 2.0)

authentication, with BrowserID directly using email addresses and OpenID Connect (in its usual deployment) simply redirecting the user to an identity provider for signing-in with a *username-password* token that can be reduced to an email address. Instead, OpenID Connect and BrowserID vary primarily in how they conceive of the flow of information from the identity provider to relying parties and these differences in information flow have important ramifications for privacy. Note that federated identity by nature imposes a security risk due to single point of losing control over identity claims at the identity provider, but this risk already present in *username-password* re-use, and putting the user "in-the-loop" via the information flow of identity claims in BrowserID mitigates this risk.

3.1 OpenID Connect

OpenID Connect is a popular federated identity system meant to provide much of the same functionality as Facebook Connect, and is deployed by large identity providers (email providers) such as Google and Microsoft [2]. OpenID Connect builds upon the well deployed base of OAuth 2.0 standard for server-side claim exchange [10], but optimizes certain elements of OAuth for server-side exchange of identity claims and requires no changes to current browsers. OpenID Connect uses OAuth 2.0 for the authorization flow while adding a small number of non-opaque identifiers in the response between an identity provider and relying party as well as adding more detailed hooks for using cryptographic signing. Once the user authenticates to an identity provider (usually via a HTTP redirection and a username-password) and so provides the relying party an access token, the identity claims are passed directly from the identity provider to the relying party, and the user is out of the loop. The flow of OpenID Connect is illustrated in Diagram 1 (the transfer of identity claims is given in *green* in this and subsequent diagrams) and outlined below:

3.2 Flow

1. A user visits a relying party that needs identity claims.
2. The relying party makes a request for identity claims to the identity provider.
3. The user is redirected to the identity provider from the relying party.
4. The user authenticates to the identity provider (typically using a username-password combination) and is granted a bearer token.
5. User is redirected back to relying party and grants authorization token to relying party.
6. The relying party sends the authorization token to the identity provider and receives an access token (a bearer token with a scope and limited lifespan).
7. While the access token is valid, the identity provider sends identity claims to the relying party.

OpenID Connect gives the identity provider ability to observe all requests for identity claims by relying parties, which is the primary flaw from a privacy standpoint as identities cannot be delinked from the identity provider. As the traffic of identity claims flows directly between identity provider and relying party, the interaction between the user and a relying party can be logged by the identity provider, as well as traced by third-parties via traffic analysis between the relying party and identity provider. Although this particular flow has the advantage of possibly authorizing requests for personal data when the user is not online and thus unable to directly intervene at the time of the request, which we call the *offline-server flow*, as the user can be off-line at the time of the interaction. Although this offline server flow is a distinct advantage for some use-cases (such as when the user authorizes the requests ahead of time or on a regularly occurring basis), it is also a danger, as the identity provider may exchange user data with relying parties without the consent of the user, leading to the possibility of identity interactions being unknown to the user. As regarding anonymity, although the architecture of OpenID Connect does not require that identifiers be persistent when sent to the relying parties (and thus allows anonymity to relying parties), the authentication mechanism to the identity provider does not authenticate particular capabilities but instead identifies the entire user or personae on a coarse-grained manner to the identity provider, and so the identity provider is aware of all relying party requests even if the user is anonymous to the relying party. Thus, OpenID Connect can be thought of as absolutely trusting the identity provider, which may be a reasonable assumption in some circumstances but this seems to be a poor choice for use-cases that require a higher degree of privacy.

3.3 BrowserID

BrowserID, also called Mozilla Persona, is backed by Mozilla as its primary identity system [1]. In particular, BrowserID breaks the direct connection between the identity provider and the relying party given by OpenID Connect [2]. Instead, the identity provider can instead mediate the transfer of identity

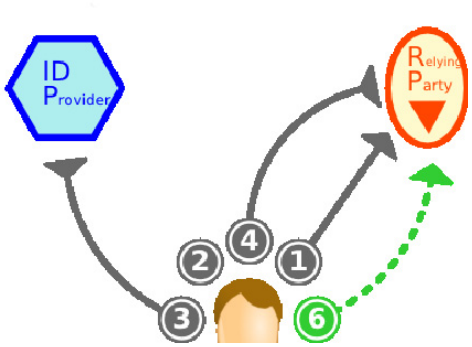


Fig. 2. Flow of BrowserID

claims via the browser, which provides better unlinkability compared to OpenID Connect. Separately, BrowserID allow users to authenticate their identity via a "verified" email as their primary authentication scheme. However, beyond traditional email, the authentication of both the user and the identity provider (which in the case of BrowserID is always the email provider) is done via the use of key material. The authorization flow of BrowserID is outlined below and illustrated in Diagram 2.

3.4 Flow

1. The user attempts to identify themselves by giving an e-mail address, and wants to bind that address to a particular set of key material (for which the user then provide a proof of possession) by having the identity provider attest to that binding.
2. The browser checks to see if a private key is present in the browser associated with that email address.
3. If no key exists for the email address locally in browser, the browser generates key material and registers the public key with the identity provider.
4. The browser sends a signed authentication credential to the relying party.
5. The relying party verifies the authentication credentials with their locally stored database of identity provider public keys and authenticates the user if verification succeeds.
6. The user sends signed identity claims to the relying party from browser (for example, possibly using HTTP POST or PUT).

From a privacy standpoint, the flow of BrowserID is superior as it avoids the observability of relying party transactions to identity providers, although it is a requirement that the user be "in the loop" via the browser, so the flow only works for use-cases where the user is actually online. Thus, we call this

particular authorization flow the *online-browser flow*. Although the relying party knows the identity provider's key and will have to check at least once with the identity provider to determine the public key of user's email address, it does not have to check more than once per e-mail (given some reasonable caching time limits). Currently, there is only one identity provider (*browserid.org*), but regardless any identity provider would have the ability to transfer of identity claims to the browser and then from the browser to the relying party. So for observability and linkability, both the identity provider-browser and browser-relying party connections would have to be observed by a third-party in order to make the entire transaction observable, and thus the transaction is not easily logged by the identity provider (although it could be by the browser). Lastly, BrowserID is still ultimately using a coarse-grained version of identity rather than capabilities, for using the email address as identifier enables an observer the ability to possibly link all observed transactions across different relying parties if they have access to these parties. Email addresses can be multiple (as would be required for personae), pseudonymous, or throw-away, but most email addresses have value for users insofar as they are also long-term valuable identifiers. So while BrowserID has superior privacy compared to OpenID Connect as regards observability of transactions and assumes the identity provider is not to be fully trusted, BrowserID assumes the browser (and all plug-ins) are trusted and reveals the email address to the relying party for every identity claim-based transaction. While the capability to check an email address is the mainstay of authentication today, this may not be enough: the reduction of the security in federated identity to checking email can itself be problematic, for example with many email clients not producing an error message when STARTLS fails to upgrade to TLS.

4 A Privacy-Preserving Capabilities-Based Approach

Earlier, we demonstrated that identity on the Web can be considered as a capability to check email and showed how BrowserID's information flow breaks the observability of the identity provider to relying party requests. In this section, we combine these two insights by outlining how an identity provider could allow the verification of any signed claim to have a capability. The crucial component of BrowserID is not the use of an email address per se, but that it implements a flow for transmitting personal data about a user as verified by signature. However, BrowserID only exposes the "user can check this email" capability, but it hints at a much bigger opportunity: the ability for an identity flow to handle general-purpose capabilities. In essence, the "signature" from the identity provider is the proof of the possession of the capability.

The key is that the identity provider can verify other capabilities about the user by virtue of their access to the personal data of the user. For example, a user may have the capability to use a smart-card to verify their identity or complete two-factor authentication via a mobile device. In fact, one can also generalize user-centric capabilities to a more standard intuitive understanding of capabilities, with the identity provider having information about what

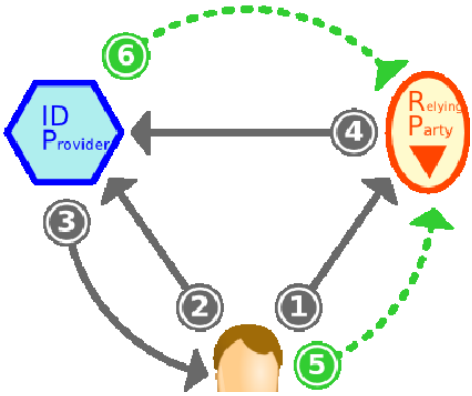


Fig. 3. Flow of Capabilities-Based System

programs ('apps') a user possesses and can run. Thus, by virtue of handing a signed token over to the relying party, the identity provider is granting the relying party that particular capability, in a similar manner as is done with bearer tokens in OAuth [10]. Capabilities provides important privacy advantages as a capability-based approach does not have to disclose an email address (and so reduce the security of the system that of SMTP), as all that is revealed is the presence or absence of a capability. The use of digital signatures in the verification step is crucial, as otherwise the client could simply state any capability was possessed. In this particular variation on federated identity, there is still an assumption of trust between the relying party and identity provider, as essentially the identity provider is responsible for the veracity of the claim and any capabilities associated thereof.[5]

In addition to BrowserID's "online-browser" flow, we can combine the idea of signed tokens as capabilities to deal with the "offline-server" flow of OpenID Connect by simply having OAuth 2.0 deliver a token that contain signed assertions of capabilities instead of the usual OAuth bearer token [10]. Thus, the capability approach allows for the user not to be online, although the "online-browser" approach would be superior for privacy-respecting use-cases. The flow is illustrated in Diagram 3:

[5] The creation of such an API is also straightforward; in Javascript code comparable to existing BrowserID implementations, the generic framework would then be thought of as *navigator.capabilities.get(domain,capability_type[required_capability_value])*. So the generic capability to check email could then be applied using *navigator.capabilities.get('gmail.com', 'email-ownership', 'romeda@gmail.com')*. Asserting a particular user is over the legal drinking age by citing the appropriate domain name but without necessarily giving away the precise age or the identity (e-mail) of the user is done by asserting the value: *navigator.capability.get('gov.uk', 'above-legal-drinking-age', 'yes')*.

1. A user attempts to authenticate to a relying party, and the relying party requests a capability.
2. The browser checks to see if a capability, i.e. a signed token, exists in the browser.
3. If no capability is present, the browser requests a capability from the identity provider, to be sent using either with the "offline-server" flow or using the "online-browser" flow depending on whether or not the user is online and their level of privacy requirements.
4. With conventional signatures, the relying party can check to the signature with the public key of the identity provider, which they can request if the key of the identity provider is not locally cached.
5. For the "offline-server" flow, the relying party requests the capability from the identity provider. The identity provider sends capability to the relying party, which then the relying provider can verify using the public key material from the identity provider it has already received in the previous step.
6. For the "online-browser" flow, the browser sends the capability to the relying party directly, which again the relying provider can verify using the public key material from the identity provider.

There are also a number of variations on the flow. For example, we may want the user to grant the identity provider permission themselves by giving them a signed token for a capability (signed by the client device itself using a local private key) to act on their behalf before the identity provider sends another signed token (signed by the identity provider) to a relying party. In this case, the role of an identity provider is shared between a browser and a service provider for the user, which can then contact other services on behalf of the user if authorized by the correct capability by a user via their client device. We can imagine this being extremely useful in some socially-oriented use-cases, as we illustrate in the following example that demonstrates the "offline-server" flow with a step of user-verification that ensures the the ability to grant that capability to the relying party is explicitly given by the user:

1. User A to User A's service provider: "I want to follow User B."
2. User A's identity provider to User A's browser: "Give me a signed token that says I have the capability to act on User A's behalf."
3. User A's browser to User A: "Is your service provider allowed to act on your behalf?"
4. User A to User A's browser: "Yes."
5. User A's browser to User A's service provider: "Here's the signed token that says you're allowed to act on User A's behalf."
6. User A's service provider to User B's service provider: "Here's a signed token that says I'm allowed to act on User A's behalf, and they'd like to follow User B. Here's a URI you can contact me at to let me know if User B consents or declines consent and publishes content."
7. User B's service provider to User B: "Can User A follow you?"
8. User B to User B's service provider: "Yes"

9. User B's service provider to User A's service provider: "User B says you can follow her, and here is some recent content published by User B."
10. User A's service provider to User A's: "You're now following User B, and here's some recent content published by User B."

5 Conclusions

Rather than invest in an entire new identity system, we have argued that simple tweaks to existing identity systems may provide many advantages in terms of security and privacy by considering digitally signed tokens as capabilities. In fact, even the bearer tokens used by implementations of OpenID Connect and OAuth implementations that avoid digital signatures can be considered as capabilities, as the token giving the capability would just be the symmetric shared secret given by the "string" of the bearer token. However, digital signatures present a much more secure manner to use key material to guarantee the possession of the capability.

From an implementation stand-point, the work is already done and so such as system is merely a manner of tweaking current deployments: One can use Javascript Web Tokens[6] in either BrowserID or OpenID Connect flows [2], and soon digital signature operations will be performed by the Web Cryptography API native in the browser.[7] However, one problem is then tying the capabilities to the token in a uniform manner. Given the immense number of possible capabilities, a central registry is probably not feasible. Commonly used services could just publish their capabilities in their documentation, and using a protocol such as WebFinger[8] would then allow relying parties to inquire about the capabilities in an automatic and decentralized manner. Capabilities actually simplifies both OpenID Connect and BrowserID, as it gets rid of making the relying party having to do requests for the precise personal data specified by OpenID Connect and does not require using BrowserID's verified email in use-cases where it is unnecessary.

Although the proposed system presents a simple methodology to create a kind of anonymous credentials for users, the proposed system does this by rooting the capabilities in the key material of the identity provider or the key material of the user on their client device. However, the blind signature schemes and interactive proof systems needed by anonymous credentials and zero-knowledge proof systems can also at some point be added to the capabilities-based federated identity system. The next step would be using blind signatures such as those suggested by the PseudoID system, with a blinding service being part of the information flow and committed either by the identity provider as a service or a third-party service [7]. This would eventually allow, if the correct cryptographic operations were revealed by the browser, a zero-knowledge proof system that selectively disclosed only some part of their capability without allowing identification.

[6] `https://tools.ietf.org/html/draft-ietf-oauth-json-web-token-06`
[7] `http://www.w3.org/TR/WebCryptoAPI/`
[8] `https://code.google.com/p/webfinger/`

The system sketched here is only the beginning as regards a holistic approach for preserving privacy. For example, in order to avoid IP-address based tracking, one could use either the "browser-online" and "server-offline" flows over a proxy system such as Tor.[9] One could forgo the concept of persistent sessions at all by ensuring that no logs of transactions are kept and carefully decoupling the client and server. Future research needs to determine not only what is possible, but how these anonymizing options can be realistically and easily implemented on top of existing systems such as OAuth [10]. Lastly, we have barely scratched the surface of the vast literature on capabilities, and the various research results on capabilities should be investigated to determine if they can be applied to federated identity-style systems.

What we have presented here is not a new system for federated identity, but simply viewing existing systems such as BrowserID and OpenID Connect through the lens of capabilities. We claim that capabilities actually simplifies the apparatus needed by developers and maximizes privacy for end-users. Also, unlike many other proposed privacy-enhancing federated identity schemes, capabilities are easy to implement on top of existing systems as, in their most basic form, capabilities can be considered to be digital signatures provided by an identity provider, although the higher level of privacy required by blind signatures and zero-knowledge proofs will require more complex signature schemes. Nonetheless, federated identity and the minimization of user's data disclosure are fully within the simple and powerful framework of capabilities, a framework that can hopefully be at some point generalized as an identity system and permissions component for the entire Web.

Acknowledgments. W3C/MIT would like to thank the Northrop Grumman Cybersecurity Research Consortium for funding this work.

References

1. Adida, B.: BrowserID (Mozilla Personae) (2012), `https://browserid.org/`
2. Bradley, J., Recordon, D., Jones, M., Sakimura, N.: OpenID Connect (2012), `http://openid.net/connect/`
3. Camenisch, J., Lysyanskaya, A.: Signature Schemes and Anonymous Credentials from Bilinear Maps. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 56–72. Springer, Heidelberg (2004)
4. Camenisch, J., Lysyanskaya, A.: An Efficient System for Non-transferable Anonymous Credentials with Optional Anonymity Revocation. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 93–118. Springer, Heidelberg (2001)
5. Camenisch, J., Van Herreweghen, E.: Design and implementation of the Idemix anonymous credential system. In: Atluri, V. (ed.) Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002), pp. 21–30. ACM, New York (2002)

[9] `https://www.torproject.org/`

6. Chander, A., Mitchell, J., Dean, D.: A State-Transition Model of Trust Management and Access Control. In: Proceedings of the 14th IEEE Workshop on Computer Security Foundations (CSFW 2001). IEEE Computer Society, Washington, DC (2001)
7. Dey, A., Weis, S.: PseudoID: Enhancing Privacy for Federated Login. In: Hot Topics in Privacy Enhancing Technologies, pp. 95–107 (2010), http://research.google.com/pubs/pub36553.html
8. Fiege, U., Fiat, A., Shamir, A.: Zero knowledge proofs of identity. In: Aho, A.V. (ed.) Proceedings of the ACM Symposium on Theory of Computing (STOC 1987), pp. 210–217. ACM, New York (1987)
9. Hardt, D.: OpenID Authentication 2.0 (2007), https://openid.net/specs/openid-authentication-2_0.html
10. Hardt, D.: OAuth 2.0 Authorization Protocol. IETF RFC (2012), https://tools.ietf.org/html/draft-ietf-oauth-v2-31
11. Laurie, B.: Access Control (2008), http://www.links.org/files/capabilities.pdf
12. McCallister, E., Grance, T., Scarfone, K.: Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). P 800-122. Technical Report. NIST, Gaithersburg, MD, United States (2010)
13. Shepard, L.: Facebook Connect (2012), http://www.facebook.com/help/?page=229348490415842
14. Story, H.: WebID (2012), http://webid.info

Privacy Preserving Course Evaluations in Greek Higher Education Institutes: An e-Participation Case Study with the Empowerment of Attribute Based Credentials

Vasiliki Liagkou[1], George Metakides[1], Apostolis Pyrgelis[1], Christoforos Raptopoulos[1], Paul Spirakis[1,2], and Yannis C. Stamatiou[1,2,3]

[1] Computer Technology Institute & Press – "Diophantus", N. Kazantzaki Str., 26504 Patras, Greece
{liagkou,spirakis}@cti.gr, {pyrgelis,raptopox,stamatiu}@ceid.upatras.gr, george@metakides.net
[2] Computer Engineering and Informatics Department, University of Patras, 26504, Rio, Patras, Greece
[3] Business Administration Department, University of Patras, 26504, Rio, Patras, Greece

Abstract. Course evaluations enable educational institutions to adjust their teaching methodologies and curricula in order to suit, best, their students' needs. Such evaluations in Greece are being conducted for many years in higher education institutes based on traditional print questionnaires handed-out to students at the lecture room. Compared to traditional paper-based questionnaires, the introduction of electronic student evaluation procedures has a number of advantages that merit consideration. a) it allows the students to evaluate courses from their home at their ease and beyond privacy breaches (e.g. by avoiding his/her fellow student looking at his/her answers), b) results are automatically archived in electronic format allowing fast further processing for the extraction of useful information, and c) it offers the possibility of using strong cryptographic tools to ensure student anonymity and data confidentiality. In this report we describe a pilot system that is being developed by the Computer Technology Institute & Press - "Diophantus" (CTI) within the context of the project ABC4Trust. The project's main goal is the development of a reference implementation of a privacy preserving eIdentity management framework based on the cryptographic primitives called Attribute Based Credentials. The pilot system will offer to a group of selected students the possibility of evaluating courses they have taken from their homes through the Internet and provide their feedback proving their eligibility to participate in the evaluation while, at the same time, preserving their anonymity. In this paper we describe the architecture and main scenarios of the pilot. CTI's long term vision is to use the pilot as a small scale proof of concept of privacy enhancement technologies in the eParticipation domain in order to introduce, in the future, of these technologies to the educational communities of all levels in Greece. These technologies will be the vehicle for supporting privacy preserving eParticipation in

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 140–156, 2014.

discussion groups whereby participants will provide their opinion anonymously but after proving that they are eligible to participate in group discussions.

1 Introduction

Over the last years the widespread use of the Internet and the new technologies by increasing volumes of population has made possible the creation of public consultation fora and opinion gathering platforms towards the realization of the concepts of eParticipation and eGovernance as integral parts of modern democracies. We have entered in a phase where old government-centric practices are strongly challenged and citizens demand direct involvement in social collective and political issues. One of the main obstacles for the wide adoption of eParticipation tools, such as polling and opinion gathering, is the reluctance of citizens to participate. This reluctance can be, partially, attributed to the, relatively, low penetration of technology among citizens (especially those of higher ages). However, the main reason behind this reluctance is the lack of trust towards ICT, which stems from the fear of citizens that systems implementing eParticipation services and procedures may violate their privacy. Our point of view, which we will discuss in the proposed chapter, is that trust in ICT-based Government should be founded on the emergence of the belief, on the citizens' side, that the systems implementing eParticipatiin respect their privacy. The departure point of our approach is that the emergence of such a belief can be considerably facilitated by designing and building systems in a way that demonstrates the respect in privacy using tools and representations that can be understood and checked by the specialist and, to a certain degree, by the layman alike. These tools and representations should provide sufficient evidence that the target system indeed handles privacy concerns and requirements to a degree that can, at least, eliminate the reluctance towards eParticipation. The ABC4Trust project, to the consortium of which CTI belongs, aims to deepen the understanding of a new privacy preserving, eIdentity management technology, based on Attribute Based Credentials. This technology enables the user to uncover only the elements of his/her eIdentity which are required in order to prove his/her eligibility in using a service. In addition, the project will organize and run the first ever pilots of ABC deployment in two real application environments, collecting useful feedback from the users that will participate in the pilots. One pilot will be performed in Sweden and will develop a privacy respecting public consultation and discussion environment for secondary school pupils while the other pilot will involve, in a privacy preserving manner, a number of University students to perform electronic evaluation of a course they have attended at the University. ABC4Trust will accumulate invaluable experience with ABC applications in two specific environments. Having these two specific pilots will give the opportunity to test credentials use and performance with two user groups of differing skills and needs.

In this paper we will focus on the Greek pilot, which concerns remote course evaluation of university courses. The pilot scenario and set-up, which will be

explained in detail in the rest of the paper, are as follows: The students will be issued credentials that certify a number of facts about them (e.g. year of study, major, number of times students appeared in lectures, etc.), allowing student with proper credentials to anonymously provide feedback on courses and instructors they had during a semester or school year. To be eligible to participate, the students' credentials should prove some facts about them, i.e. whether they have taken the course, the year of their first registration to the university department, and their the exact course attendance information (i.e. exact number of times the students actually appeared in lectures). All this should be done without revealing any elements that uniquely identify the student. This privacy requirement can be satisfied using attribute based credentials over the reference implementation of the project where for each student a set of credentials will be defined in the context of the project that allow proving their eligibility for participating in a specific evaluation (e.g. proof that they are indeed students of the department offering the course, proof that they are registered to the course under evaluation, proof that they have attended sufficient number of classes, indication of grade level (e.g. pass/fail, without indication of exact grade) etc.). The student credentials are stored in their smart cards and are verified (without ever the credentials leaving the card) by the relying party for compliance with the relying party's access policy.

There will be two pilot rounds over the 2012-2013 academic year. The first round will be run within the fall 2012 semester and the other round will be run within the spring 2013 semester. This will assure that the second trial will take into advantage the experience from the first as well as a new version of the reference implementation with corrections proposed during the first trial. CTI's plans is to use this pilot as means of introducing ABCs first to the educational community of Greece as a tool for opinion gathering of certified group members and then to the Greek government as an eParticipation tool that respects users' privacy. Our belief is that innovations, such as ABCs, that can introduce breakthroughs in privacy respecting eIdentity management are more likely to be adopted and trusted by people if they are introduced gradually, in a step-wise manner, from specialized technical groups of people to more general non-technical groups and not in an "all-inclusive manner", which may be viewed by many as an abrupt, intrusive effort to interfere with people's everyday lives and ways of conduct.

2 The Core ABC Ideas and the ABC4Trust Project

2.1 Privacy and Privacy ABCs

Commonly used user authentication methods (e.g. PKI based) that are employed today for controlling access to Internet services most often fall short, with regard to respecting users' *privacy*. In general this situation arises in services in which only a *subset* of a user's full identity profile is necessary to allow access to a service. Such services range from accessing online libraries, where there is no need to give full identity profile to access books but only a proof that you are

subscribed to the library, to online borrowing of movies, where you may have to prove that you are of appropriate age (e.g. older than 18) in order to watch particular films. In such types of applications there is, clearly, a need for a *partial*, and not complete, revelation of the user's identity.

Privacy Attribute-Based Credentials or Privacy-ABCs, for short, is a technology that enables privacy preserving, partial authentication of users. Privacy-ABCs are issued just like normal electronic credentials (e.g. PKI based) using a secret signature key owned by the credential issuer. However, and this is a key feature of this technology, the user is in position to transform the credentials into a new form, called *presentation token*, that reveals only the information about him which is really necessary in order to access a service. This new token can be verified with the issuer's public key.

The main ABC entities are four: the *Issuer*, the *User*, the *Verifier*, and the *Revocation Authority*. In general, the Issuer issues credentials containing certified user attributes, thereby attesting the validity of the attributes. The Verifier, or *relying party*, on the other hand, offers a service with access limited only to those users for which it can verify the possession of certain attributes (or credentials). The Revocation Authority is responsible for revoking issued credentials, i.e. disabling the possibility of creating presentation tokens out of them.

2.2 The ABC4Trust Project

Some proposals of how to realize ABC systems in the literature can be found in [3–5]. Notable in this respect has been the appearance of two technologies, *IBM's Identity Mixer* and *Microsoft's U-Prove*, as well as some preliminary work done in past EU projects. More precisely, the EU-funded projects PRIME and PrimeLife have actually shown that the state-of-the art research prototypes of ABC systems can indeed confront the privacy challenges in today's Internet applications.

However, even though PRIME and PrimeLife showed that ABC technologies can provide, in principle, privacy respecting user authentication, the emphasis of understanding ABC technologies was rather on their theoretical analysis. Moreover, no agreed upon set of functionalities, information formats, and protocols has appeared in a prototype or set of libraries form that can boost the applicability and wide acceptance of ABC technologies further. Accordingly, a gap existed between the theoretical ABC proposals and the real user authentication applications. This gap was, further, magnified due to the lack of standardization in the ABC domain. As a result, the *European Network and Information Security Agency* (ENISA) observed in [2] that although ABC technologies have been available for a long time, no steps have been taken towards their adoption in mainstream user authentication and eIdentity card applications (however, countries such as Austria and Germany have taken some important steps towards this direction).

The ABC4Trust project (see [1]) aims at eliminating the gap between theory and practice in ABC technologies in order to pave the way towards their deployment in applications requiring partial user authentication. In particular,

the project's two main goals are (i) to propose an architectural framework for Privacy-ABC technologies that allows their co-existence and interchangeability (e.g. IBM's Idemix and Microsoft's U-Prove) and (ii) to provide a reference implementation of those ABC components that developers can deploy in order to build privacy enhanced technologies in their own applications. A key element of the project in demonstrating the practical use of ABC technologies is the implementation of two ABC based pilot applications: one to be run in a school in Sweden and one to be run at a University in Greece. This paper focuses on the second pilot.

3 Towards Electronic University Course Evaluations

Course evaluations have become, today, a standard practice in most universities around the world. However, these evaluations are most frequently conducted by traditional means: students who happen to be in the lecture room on the day the evaluation is scheduled are handed paper-based questionnaires which they have to complete *anonymously* while the instructor waits outside. The anonymity and absence of the instructor requirements protect student's privacy. In cases where the evaluations are conducted electronically, the employed infrastructure does not preserve student's privacy since it requires them to prove that they are eligible to participate by asking them to authenticate themselves. The authentication reveals, at least to the authentication server, their full identity which opens the door to linking their identity to the questionnaire form they fill in on-line. These considerations may ever deter students from using electronic course evaluation systems resulting in of paper-based course evaluation means with all their disadvantages (e.g. cost, difficulty in processing and preserving evaluation results, small student participation etc.).

Consequently, the goal of the pilot is twofold: (i) to prove the applicability of ABC technology in a real application environment and (ii) provide the first implementation of an electronic course evaluation platform for universities (and educational institutions in genera) that ensures the participants' privacy while, at the same time, guarantees that only eligible students participate by requiring them to reveal only information that proves this eligibility and nothing else.

To achieve these goals, the pilot system will issue to the students Privacy ABCs attesting that they are students of the University and have registered to the course. These credentials will be stored on smart cards that will be distributed to participating students. Moreover, in order to achieve as much as possible accuracy in the gathered opinions of the students, only students who have attended the course sufficiently many times, beyond a preset threshold (can be set to 0 if all students should participate). Thus, in order to receive *attendance credentials* (one per each student appearance in the lecture room) the students will wave their smart cards close to NFC (Near Field Communication) device (essentially, a contactless smart card reader) installed on a computer located in the lecture room. Thus, the number of class appearances, or class attendance

units, of each students is equal to the number of class attendance credentials stored in the card.

At the end of the semester, the students will be able to use the credentials stored on their cards in order to authenticate towards the Course Evaluation System and prove their eligibility to participate in the evaluation of the course. They only prove to the system that (i) they are students of the University, (ii) they are registered to the course, and (iii) they have sufficiently many attendance credentials. Nothing else is revealed about the students at any stage of the whole process.

The students are also expected to give feedback with respect to the usability and effectiveness of the ABC technology. The goal is to gather information that will be useful for ABC technology developers and will result in suitable adjustments of the technology to meet their remarks and expectations. This opinion gathering has not been done before for ABC technologies which were, instead, evaluated on a theoretical basis by their developers themselves or by their peers and not by actual users.

Beyond the pilot, our vision is that, as a result of the ABC technologies and the success of the pilot, educational institutions in general will be able to run their own trusted online course evaluation platform. Moreover, the institutions will be able to organize fully anonymous polling procedures targeted to specific *user groups*. This will be accomplished by verifying the eligibility of the users (using ABC technologies) that are allowed to participate and disallowing users that should not interfere with the polling from participation to avoid (perhaps even malicious) "contamination" of the polling results.

4 eVoting vs. Privacy ABCs

Electronic Voting, or eVoting, techniques have given the opportunity to employ computer systems in order to perform electronically national elections as well as less critical types of opinion gathering from public consultation and discussion fora, with low cost, potential for large participation as well as convenient and fast processing of the election results. Of course, participants' *anonymity* is the common denominator of all these processes as one of the major anchor points for protecting the individual's privacy. In this respect one can argue, rather superficially, that both eVoting techniques and Privacy ABCs cover the needs of all such electronic opinion gathering processes. There is subtle difference between them, however, that goes unnoticed at first sight and may reveal the fact the eVoting *end* Privacy ABC techniques are rather complementary to each other and can potentially be used both in opinion gathering applications. The eVoting techniques focus on allowing a voter to *submit*, securely, a voter's opinion rather than *authenticating* her for eligibility to vote. In this respect eVoting targets the *confidentiality* of the vote rather than authenticating the voters. Voter authentication is accomplished by usual PKI based techniques (using suitably drafted election catalogues) that lead to uncovering fully the voter's identity to the system. To say the least, the system *knows* that a particular voter has

cast her vote, a fact that the user might not want to reveal. Privacy ABCs, on the other hand, do not target confidentiality of information but user privacy preserving user authentication. In their context, a user can authenticate herself to an eVoting system without revealing her identity at all, and then proceed to cast electronically her vote, in a confidential manner (i.e. vote is encrypted) using eVoting techniques. Therefore, eVoting and Privacy ABCs target different security requirements and can act complementary in order to support *privacy* and *confidentiality* preserving electronic opinion gathering processes.

Moreover, and with an eye towards enhancing the course evaluation pilot later (beyond the scope of the project), there are situations where knowledge of certain characteristics (or attributes) of individuals' profiles may enhance the conclusions drawn from analyzing only the responses of the individuals to an opinion gathering process. For instance, the ministry of education may initiate a discussion as to whether general entrance examinations at the Universities should be abolished and all students enter at the University schools of their choice, depending on their grades only. Then an discussion result indicating that 90% of the participants support the abolition of general examinations may provide some clue as to what the feelings of society are towards the examinations but a closer examination, according to individuals' profiles, may indicate that only 10% of individuals who are university professors support the abolition and, thus, governmental authorities should be careful in implementing such a radical change without further discussion and elaborations. Moreover, there are also situations where knowledge of characteristics of individuals' profiles may be mandatory. For instance, the ministry of education may want to start a discussion about whether University infrastructures in a country are sufficient for a normal operation of Universities. In order to take as substantiated opinions as possible, the ministry decides to open the discussion only to 3rd, or more, year students and professors, who have had sufficiently many years of university life in order to be in position to judge more accurately the University infrastructures.

Overall, our view is that Privacy ABCs with their selective identity disclosure properties offer new opportunities for conducting more accurate privacy preserving opinion gathering processes using the confidentiality properties (e.g. encrypted vote or opinion) of eVoting techniques.

5 Pilot Operational Environment and Requirements

In this section we will describe the environment in which the electronic course evaluation system operates and the main requirements that must be satisfied in order to ensure user privacy and personal information protection.

5.1 The Operational Environment

The course evaluation pilot system will be used by students of the Computer Engineering and Informatics Department of the University of Patras in Greece. The department is located close to CTI's premises, where the pilot system will

be installed, operated, and monitored. Figure 1 shows the pilot's system and network infrastructure. Network security relies, partly, on a pair of firewalls which are connected to a high availability configuration (active-standby, without NAT, with automatic fail-over capability between them). The firewalls appear in between the border router and the internal network, inspecting incoming and outgoing traffic and ensuring protection against malicious attacks. For instance, these firewalls can block suspicious source IP addresses in the case of detected DoS attacks as well as traffic directed towards internal servers. However, they alone cannot block packets with malicious content (e.g. viruses) which are taken care of by other components of the security subsystem.

In addition, a *DMZ* subnet exists in CTI's network infrastructure. A DMZ (which stands for "Demilitarized Zone") is a physical or logical subnetwork that encompasses and publicizes an organization's computing services to an external, untrusted network, most commonly the Internet. The DMZ offers an additional security layer to an organization's local area network and services since an attacker from the outside can only access the DMZ and not parts of the internal infrastructure of the organization. The DMZ contains all the servers, such as web and Virtual Private Network (VPN) servers, that offer public services and and do not have (for security reasons) any connection to CTI's internal network. The VPN servers will allow secure remote administration of the Course Evaluation and the University Registration systems. These systems are also in the DMZ and have their own publicly available services. All http/https requests (which obey access control lists and rules) to these servers pass through the DMZ.

The Course Evaluation system supports several different user groups and roles with corresponding remote access rights: students, professors, and possibly members of the *HQAA* (Hellenic Quality Assurance Agency for Higher Education - the organization responsible for ensuring quality in higher educational institutes). For instance, the administrators can access the DMZ via https/http/ssh/ldap connections from CTI's internal network or, remotely, through the Internet (via VPN connections). CTI's domain controller server authenticates the administrators and then control passes to an AAA (Authentication, Authorization, and Accounting) server that will finally give access to the internal network and the two pilot systems.

Finally, at the perimeter of CTI's network infrastructure a border router exists which is placed between the firewalls and the external network and performs some basic checks on incoming and outgoing network activity, such as *ingress* and *egress* filtering that may be helpful in blocking some Internet-based worms from reaching the firewall. In computer security terminology, *ingress filtering* refers to techniques which are employed to verify that incoming traffic actually comes the originators that the traffic packets claim to be from. Complementarily, *egress filtering* refers to techniques of monitoring and, possibly, restricting the type of outgoing traffic from one network to another. Most commonly, this outgoing traffic may contain information from private LANs which may be maliciously directed to the outside network (e.g. the Internet) and should be intercepted and, perhaps, blocked. This border router also implements some generic access

list based control in order to increase the level of security and handle some types of attacks like DoS (Denial of Service) or DDoS (Distributed Dos). Additionally, access control lists are maintained by the two internal routers. Through these lists one can specify which processes can access which system objects, as well as what operations are allowed on the objects themselves.

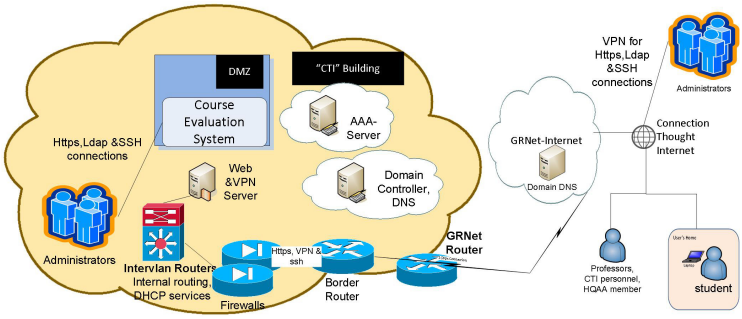


Fig. 1. Course Evaluation System Network Infrastructure

5.2 Pilot Requirements

As electronic course evaluations should be privacy processes, having a focus only on the traditional security requirements, i.e. confidentiality, integrity, and availability ([6]), does not suffice, as it was argued earlier. In the pilot's context, we further need to take care of three additional privacy oriented requirements which are unlinkability, intervenability, and transparency (for further details see [7] and [9]):

– *Unlinkability*: Unlinkability requires that all data processing is done in such a way that the user's actions are unlinkable to each other. Unlinkability is a key element for imposing user data minimization (see [8]) because it encompasses all kinds of separating identity elements from users (i.e. full identity), e.g., by means of anonymization, pseudonymisation, data erasure or simply by not keeping identity data at all. In addition, unlinkability aims at separating different subsets of a user's identity, if they can serve different purposes, thereby supporting the principle of *purpose binding*. Unlinkability, in this wide definition, encompasses the criteria from the Privacy Class in the Common Criteria, i.e. anonymity, pseudonymity, unlinkability (in a stricter definition), and even unobservability in the sense that any observation by another party cannot attest to the action or non-action of a particular user. The main objective of the unlinkability requirement is to minimize risks due to the misuse of user's identity elements and to prohibit or restrict user profiling efforts (see [9]).

However, in some cases there is a requirement for some form of "consumption control" (e.g. one-time coupons, online voting), where users should remain anonymous to the Verifier, but should not be able to use a service more than once by creating multiple, *unlinkable*, identities. This is true in the course evaluation scenario of the pilot where students should not be allowed to participate more than once in the course evaluation. For such scenarios, Privacy-ABCs offer the concept of *scope-exclusive* pseudonyms which are still re-usable but they are unique.

Scope-exclusive pseudonyms are cryptographic pseudonyms derived from a user secret that underlies an issued credential and a scope string (e.g. the URL of a web service). Such pseudonyms are cryptographically guaranteed to be unique per scope string and per user secret. When a Verifier requests from the User to present a scope-exclusive pseudonym for a specific scope, he can be sure that only a single pseudonym can be created per user.

This Privacy-ABC feature is employed in the Patras scenario. When a User (student) interacts with the Verifier (Course Evaluation System), he receives a presentation policy that requests from him to present a scope-exclusive pseudonym with the scope string "urn:patras:evaluation" along with the rest of credential attributes or predicates. This way, a student course evaluation is stored in the system's database along with his scope-exclusive pseudonym. If a student desires to re-evaluate a course, he has to present again his scope-exclusive pseudonym and thus his previous evaluation is overwritten in the database.

- *Transparency*: Transparency requires that all parties involved in any privacy critical data processing operation clearly agree upon and understand the legal, technical, and organizational conditions behind this processing. Satisfying this requirement entails the clear and comprehensible statement of involved regulatory measures such as laws, contracts, or privacy policies, as well as the description of employed technologies, of the organizational processes, and the corresponding responsibilities, among other things. The involved parties parties should understand the risks and have sufficient information on potential countermeasures for privacy regulations as well as on their usage and limitations. This information should be given before the data processing takes place (ex-ante transparency) which is, in particular, necessary if data subjects are being asked for consent or if data controllers want to decide on the usage of a specific system. But also after the processing has taken place, transparency is required on what exactly happened to the data so that all involved parties can keep record of the processing that took place (for more details see [9]).
- *Intervenability*: Intervenability requires that all the parties involved in any privacy critical data processing operation, including the individual whose personal data are processed, have the opportunity to intervene, where necessary, and interrupt the operation. The goal of this requirement is to offer user corrective measures and counterbalances towards unwanted data operation processes. Intervenability supports the individual's rights to corrective actions and personal data erasure as well as the right to file a complaint or

to initiate a dispute in order to claim amends when undesirable effects have occurred. For data controllers, intervenability allows them to have efficient means to control their data processors, as well as the employed ICT systems, in order to prevent undesirable effects. This includes, for example, the ability to stop a running process in order to avoid further damage, the right to initiate an investigation, ensuring secure erasure of personal data (including data items stored on backup media), manually overriding automated decisions, or applying "breaking glass" policies (for more details see [9]).

These requirements, together with the security requirements we discussed earlier, form a complete set of six user privacy protection requirements. However, as these individual protection goals act complementary to each other, it is possible that sometimes they may act contradictory too. In such a case, an optimum balance should be sought, depending on the application in hand (see [7, 9]).

6 High Level Description of the Pilot System Architecture

The architecture of the pilot course evaluation system is shown in Figure 2. As it can be seen, the architecture is based on various components that have different functionalities and roles. In what follows, we will describe their properties and interactions within the pilot's context.

– *Patras Portal*: This component is web base information portal. Through this portal, the students will get information about the pilot system and functionality as well as information about its usage. Moreover, this portal also contains the necessary links to the components of the system (e.g. University Registration System, Course Evaluation System) that the students should access.

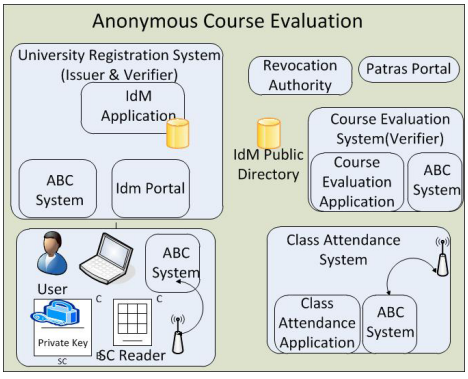


Fig. 2. High Level Architecture

- *University Registration System*: This component issues credentials to the students. It is comprised of the ABC Engine, the IdM (Identity Management) Application, and the IdM portal (the web page that explains to the users how the IdM operates and helps them to create their credentials). The IdM application is a web application whose users are the students and the university registration office employees.
  In particular:
  - A university registration officer is authorized to insert student information in the database of the University Registration System.
  - A university registration officer can revoke a student credential, e.g. when a student graduates from the university or upon student request (e.g. smart card loss).
  - Students are issued credentials that certify that they are, indeed, students of the University of Patras.
  - Students are able to browse their personal data that is stored in the IdM database.
  - Students are able to manage, themselves, a limited subset of their personal information.
  - Students are issued credentials that certify that are registered to the university course that will be evaluated in the end of the semester.

  When a user requests a credential, through the IdM portal, the IdM application invokes the ABC System in order to initiate the issuance protocol.
  Finally, the parameters of the University Registration System (e.g. system parameters, public key information, revocation information) should be stored in a public repository, so that all the other system components can access them. This repository is the IdM Public Directory, as it can be seen in Figure 2.
- *Course Evaluation System:* This component is responsible for the realization of the anonymous course evaluation procedure Its sub-components are an ABC System and a Course Evaluation Application. The ABC System sub-component performs access control to the Course Evaluation Application. This access control is achieved by presenting a policy to the students stating what credentials they must possess in order to proceed. Only users who own the required credentials are given access to the Course Evaluation Application. The Course Evaluation Application is a web application that implements the course evaluation procedure. Potential users of this application are the students, the university professors, and HQAA members.
  More precisely:
  - Course professors can upload questionnaires regarding their courses.
  - Students are able to anonymously evaluate the courses to which they are registered and have attended sufficiently many times.
  - When the evaluation procedure is completed, professors and HQAA members can access the course evaluation results.
- *Class Attendance System:* This component is responsible for issuing an attendance credential to a student's smart card each time the student attends

a lecture. It consists of an ABC System and a Class Attendance software application. The equipment that is required for the Class Attendance System is a laptop with a contactless smart card reader attached to it. The Class Attendance Application runs on the laptop and is responsible for transferring (through the contactless reader) to the students' smart cards the required information during an ABC credential issuance protocol.
- *User's Home Application:* This component needs to be installed on students' personal computers which should, further, be equipped with a contactless or contact smart card reader (if the chosen contactless smart cards support contact operation too). Its main sub-component is an ABC System that enables the user to perform ABC-related operations on their credentials (which are stored in their smart cards) and initiate credential issuance and verification protocols. There is also another software component, called *User Agent*, which is provides the user with an interface that enables her to browse the credentials stored on her smart card, delete credentials and backup credentials on her computer.

With respect to role mapping, the following have been defined:

- *Issuer:* The system component that issues credentials to users. In our pilot this component is the University Registration System. The users of the system, i.e. students of Patras University, interact with this component in order to collect credentials that can be, later, used to prove that they belong to the university and are registered to the course. A second Issuer in our architecture is the Class Attendance System. This component issues an attendance credential to a User, each time she attends a lecture of a course.
- *User:* The entity (human) that collects credentials from an Issuer in order access services offered by a Verifier. The Users in the Patras Pilot are the students that will participate in the trial. In order to interact with Issuer and Verifier components, the students use the User Agent. The User Agent runs locally on their computers and enables them to perform various ABC related operations, e.g. participate in credential issuing and verification protocols as well as use, browse, and delete credentials stored on their smart cards.
- *Verifier (relying party):* The system component offering a service. This component defines restrictions on the credentials that legal users of the service must have and which items from their credentials need to reveal in order to prove their eligibility to use the service. The Verifier accepts credentials from Issuers that she trusts. In our architecture, the component that acts as a Verifier is the Course Evaluation System. This component allows access to a specific course evaluation only to those Users (i.e. students) that satisfy certain properties e.g. students that have booked this course and have attended a minimum number of its lectures. The Issuers that this Verifier trusts are the University Registration System and the Class Attendance System.
- *Revocation Authority:* This component is responsible for revoking issued credentials upon request of the revocation requestor. In our architecture the component that implements the Revocation Authority is the University Registration System. Upon request, a university registration office employee, can

use the University Registration System to revoke the requested credential. Revocation is required in case a student has graduated from the University or when a student loses the smart card containing his credentials.

7 The Realization of the Pilot

We will now take a closer look at the stages involved in the realization of the pilot. We will, first, describe the Setup Phase and how the involved credentials are obtained (i.e. University, Course Registration, and Class Attendance credentials). In addition, we present the basic steps that a student has to follow in order to back-up and restore Class Attendance credentials so as to not miss the opportunity to participate in the course evaluation if she loses her card. Finally, we describe the Course Evaluation process and how student credentials can be revoked.

7.1 System Setup

This section gives a high level description of the generation procedure of pilot and system parameters. The Setup Phase consists of the following steps:

(a) A smart card and a smart card reader is given to each student. The University Registration office provides each student with a smart card in a sealed envelope (which is marked with the corresponding smart card ID and, also, contains the card PIN and a PUK numbers) and a suitably protected slip of paper with a unique student password. In this phase, the smart card does not contain any student information. Furthermore, the slip of paper contains a unique correspondence between the provided smart card ID and the student password. A list of student names and their corresponding identification numbers and passwords is maintained by the University Registration office.

(b) The Course Evaluation System, the University Registration System, and the Class Attendance System are started: The information of students participating in the Pilot (first and last name of the student, University Name, Department Name, and Matriculation Number) is provided to the IdM database by CTI in collaboration with a University Registration office employee. In addition, the administrators of the University Registration System and the Class Attendance System generate issuer parameters and the issuance keys for the involved issuer components. Subsequently, the issuer parameters of the University Registration System are stored in the IdM, so as to he accessible by the ABC System components.

After the above steps have been completed, students install the appropriate software on their computers (e.g. User Agent) and can proceed to the next step in order to obtain their credentials.

7.2 Obtaining University and Course Registration Credential

In order to obtain their University and Course credentials, the students follow instructions provided to them in the Patras portal. These instructions are contained in a User Manual that explains, briefly, the goals and set up of the pilot, the main ABC concepts, as well as the steps that the students need to follow and how they can verify the correctness of their actions. As explained in the manual, each student need to log on to the University Registration System running the IdM. In order to collect a University and a Course credential the student places her smart card into (or near, if the reader is contactless) the reader. After the system has authenticated the student, the credentials issuance protocol is triggered and the generated credentials are transferred in the smart card (see Figure 3).

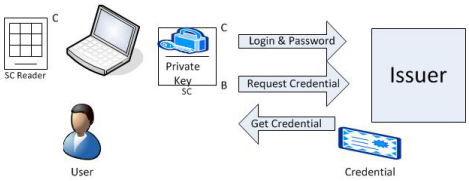


Fig. 3. Obtaining the University or Course Credential

7.3 Obtaining Class Attendance Data

We, will, now describe how a student can collect attendance credentials using the Class Attendance System located in the lecture room. At this point, it is assumed that the System Setup procedures have finished successfully and students have obtained their University and Course credentials.

Fifteen minutes before the lecture begins, an authorized CTI employee will place the Class Attendance System in the lecture room. The Class Attendance System contains the course specific information given by the lecturer (this information includes room number, lecture date, and lecture start/finish times). As soon as the student enters the lecture, she waves her smart card in front of the Class Attendance System in order to collect an attendance credential. This is done by the execution of a credential issuance protocol between the smart card and the Class Attendance System. The generated attendance credential is stored in the card.

Each student can, also, back up attendance credentials and browse the credentials stored on her smart card by running an application locally on her computer (User Agent application). Moreover she will be able to restore the backed up information in case the smart card is lost.

7.4 The Course Evaluation

The pilot involves two rounds, in each of which one course will be evaluated. One round will take place after the end of Fall 2012 semester and the other round will take place after the end of the Spring 2013 semester.

The eligibility criteria, and the corresponding credential verifications, for participating in the evaluations are the following three: (i) the student has a University Credential (i.e. the student belongs to the University) (ii) the student has a Course Credential (i.e. the student has registered the course), and (iii) the student has sufficiently many course attendances (i.e. she has gathered in the smart card sufficiently many attendance credentials). All these criteria are checked for each in a privacy respecting fashion using the ABC technology. In what follows, we describe the course evaluation scenario preparation and execution.

Near the end of each semester, the course evaluation questionnaires are prepared for each course in collaboration with the course lecturer and uploaded in the Course Evaluation application. After the end of the semester, and for a strictly specified evaluation course period, students log on the application and provide, anonymously, their evaluation, after their eligibility is verified according to the three criteria mentioned above. It should be stressed here that each student is allowed to access the Course Evaluation server and provide her evaluation several times. However, only her last evaluation is taken into account due to the use of scope-exclusive pseudonyms (see the discussion on scope-exclusive pseudonyms in Section 5.2).

The Course Evaluation application contains a database for storing eligibility policies and course evaluation data for subsequent analysis. As an important privacy enhancement, the system is configures in such a way that if the employed student eligibility policy for a specific question leads to the emergence of a small and, thus, potentially identifiable subset of students, then the system prevents the student to proceed with this question.

7.5 Student's Privacy-ABCs Revocation

Under certain circumstances, the University registration officials should be able revoke students's credential. This is especially needed in the case a student graduates or finishes a course as well as if her smart card is lost or damaged. In the cases of graduation of course completion, the University Registration System Administrator revokes the credential and eliminates the corresponding information from the university system. In the second case, after the student officially declares smart card loss, the administrator revokes the student University credential and deletes her private information from the ABC system. Subsequently, the student gets a new envelope (containing PIN, PUK) and a smart card form the University Registration office, which she can use to obtain, from the beginning, University and Course credentials in order to restore backed up attendance credentials from her computer (in case she has already performed a back up).

8 Beyond the Pilot

CTI's vision is to be able to extend the ABC4Trust pilot scenario to a full-fledged environment for supporting public consultation and discussion fora targeted at specific educational community groups. The vehicle for this will be the *Greek School Network*, which CTI manages, that connects all Greek schools' local networks together as well as with the Internet. Members of this large community, equipped with ABC based iIdentity cards, will be able to prove their participation eligibility by uncovering the elements of their identities which prove their eligibility in a way that does not uncover their full identity.

The next step is for CTI to promote the use of this consultation environment by the Greek government for enhancing eParticipation in Greece based on gradual introduction of ABC4Trust technology in a small number of Internet based interactions between the citizen and the government and then extend ABC-based privacy preserving services to a wider spectrum of applications useful to the citizens thus contributing to the gradual enhancement of eParticipation and eIdentity management in Greece.

Acknowledgements. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement no. 257782 for the project Attribute-based Credentials for Trust (ABC4Trust).

References

1. Project Description, ABC4Trust-Attribute-based Credentials for Trust, `https://abc4trust.eu/`
2. European Network and Information Security Agency, Privacy Features of European eID Card Specifications. Position Paper (February 2009), `http://www.enisa.europa.eu/act/it/privacy-and-trust/eid/eid-cards-en`
3. Camenisch, J., Lysyanskaya, A.: An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 93–118. Springer, Heidelberg (2001)
4. Camenisch, J., Lysyanskaya, A.: Signature schemes and anonymous credentials from bilinear maps. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 56–72. Springer, Heidelberg (2004)
5. Chaum, D.: Security without identification: Transaction systems to make big brother obsolete. Commun. ACM 28(10), 1030–1044 (1985)
6. Federrath, H., Pfitzmann, A.: Gliederung und Systematisierung von Schutzzielen in IT-Systemen. Datenschutz und Datensicherheit (DuD) 24(12), 704–710 (2000)
7. Hedbom, H., Schallaböck, J., Wenning, R., Hansen, M.: Contributions to standardisation. In: Privacy and Identity Management for Life, pp. 479–492 (2011)
8. Pfitzmann, A., Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management (2010), `http://dud.inf.tu-dresden.de/Anon_Terminology.shtml`
9. Zwingelberg, H., Hansen, M.: Privacy protection goals and their implications for eID system (2012)

Collection and Storage of Personal Data: A Critical View on Current Practices in the Transportation Sector

Eleni Kosta[1], Hans Graux[2], and Jos Dumortier[2]

[1] TILT-Tilburg University, Tilburg, The Netherlands / time.lex law offices, Brussels, Belgium
e.kosta@tilburguniversity.edu, eleni.kosta@timelex.eu
[2] ICRI-KU Leuven, Leuven, Belgium / time.lex law offices, Brussels, Belgium
{hans.graux,jos.dumortier}@law.kuleuven.be/timelex.eu

Abstract. This paper is based on a 2011 ENISA study that aimed at the analysis of two core principles that can be considered as key manifestations of privacy by design: on the one hand the *principle of minimal disclosure* (which is also known as the data minimisation principle), and on the other the *duration of the storage of personal data* (which is also known as conservation principle). It focuses on the data collected for two specific application areas: online ticket booking and purchasing, and the collection and exchange of so-called Passenger Name Record (PNR) data in the European air travel sector and it provides a summary of its findings in relation to the transportation sector across the EU Member States. The analysis shows that it is worrisome to observe that so many systems deployed in real life do not follow a privacy by design approach, and insufficiently consider the data minimisation and data conservation principles. There is a need for these principles to be strengthened in practice, through legislation and governance mechanisms that favour privacy by design, including a clear assessment of privacy impacts and the identification of more privacy conscious implementation alternatives, in order to ensure that the personal data of European citizens is proactively protected, instead of having to modify operational systems only after privacy problems come to light.

1 Introduction

The European legislative approach to protecting personal data against abuses is based on a number of core principles. Most of these primarily target human behaviour, by specifying what persons can and cannot do with personal data. However, because of the strong role that modern technologies play in enabling the processing of personal data – collecting, analysing and disseminating it – the realisation has grown that the design of information processing systems themselves should be impacted by data protection concerns as well. Technology should become a tool that prevents data protection abuses, instead of enabling them. The clearest manifestation of this shift in focus is the so-called "privacy by design" principle.

The privacy by design principle is understood as meaning that "privacy and data protection are embedded throughout the entire life cycle of technologies, from the

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 157–176, 2014.

early design stage to their deployment, use and ultimate disposal".[1] This principle has been promoted as a fundamental tool for ensuring trust and security in European public policy, including notably through the recent Digital Agenda for Europe: "The right to privacy and to the protection of personal data are fundamental rights in the EU which must be – also online - effectively enforced using the widest range of means: from the wide application of the principle of "Privacy by Design" in the relevant ICT technologies, to dissuasive sanctions wherever necessary."[2]

Recently, the European Commission discussed the "privacy by design" principle in the frame of the current review of the European Data Protection Directive, with a view of explicitly codifying the principle into European data protection rules, along with the issues that need to be examined in order to develop a "comprehensive and coherent approach guaranteeing that the fundamental right to data protection for individuals is fully respected within the EU and beyond"[3]. The European Commission admitted that "the 'Privacy by Design' principle could play an important role in [ensuring compliance with data protection rules], including in ensuring data security"[4], and announced its intention to examine possibilities for the concrete legislative implementation of the principle.

While the importance of the *privacy by design* principle as a way of protecting personal data is becoming clearer every day, it is much less clear to what extent the principle is observed in practice. This is especially true in the world of online service providers, where unbridled and excessive data processing is easy, cheap, and relatively risk free. To examine this tension, ENISA recently commissioned an analysis of two core principles that can be considered as key manifestations of privacy by design: on the one hand the *principle of minimal disclosure* (which is also known as the data minimisation principle), and on the other the *duration of the storage of personal data* (which is also known as conservation principle).

The study, entitled the "Study on data collection and storage in the EU"[5] was not intended as a theoretical legal study that identified and analysed national legislation, but instead focused on a limited number of relevant use cases, attempting to discover if and how the aforementioned principles were expressed in concrete legal or regulatory provisions applicable to these cases, and how they were observed in practice. To achieve this objective, the study collected detailed legal information from expert correspondents in all 27 Member States, who were also asked to identify and analyse

[1] European Commission, Communication from the Commission to the European Parliament, the Council the European Economic and Social Committee and the Committee of the Regions "A Digital Agenda for Europe" COM (2010) 245, 19 May 2010, p. 17 (fn. 21).

[2] Idem, p. 17.

[3] Idem, p. 4.

[4] European Commission, Communication from the Commission to the European Parliament, the Council the European Economic and Social Committee and the Committee of the Regions "A comprehensive approach on personal data protection in the European Union" COM(2010) 609 final, 04 November 2010, p. 12.

[5] See: `http://www.enisa.europa.eu/act/it/library/deliverables/data-collection/at_download/fullReport`

three real life examples of use cases in their own country, covering three different sectors: social networking, transportation, and electronic communication.

This paper is based on the ENISA study, and provides a summary of its findings in relation to the transportation sector. In the sections below, we will first provide an overview of the European regulatory backdrop of the data minimisation and conservation principles, and then assess how these principles are being observed in the transportation sector across the EU. This will be done based on the data collected through the aforementioned study for two specific application areas: online ticket booking and purchasing, and the collection and exchange of so-called Passenger Name Record (PNR) data in the European air travel sector. Finally, we will present our conclusions on the current implementation of these principles in the transportation sector.

2 Data Collection and Storage of Personal Data in the European Union

The Data Protection Directive refers to basic principles for the processing of personal data, commonly known as *data protection principles*. These principles are implemented through obligations that data controllers should comply with in order to protect the data they hold, reflecting both their interests and those of the data subjects.[6] The collection and processing of personal data has to be carried out in compliance to the data protection principles, as they are specified in Article 6 of the Data Protection Directive[7]. In relation to the principles of *minimal disclosure* and the *duration of the minimum storage of personal data*, the Data Protection Directive stipulates that personal data must be "adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed"[8] and they must be "kept in a form that permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed".[9] In practice these rules implement the concept of the aforementioned *principle of minimal disclosure* in a binding legal text, and they will be referred to interchangeably throughout this paper.

The data controller generally decides both on the types and amount of data that should be collected, processed and possibly further processed, as well as on the minimum period during which the data can be stored. These decisions will (or should) be based on the *proportionality principle* and after carrying out a 'balance test' between the various interests at stake, for instance the protection of the individual and the commercial profit of the service provider. At least in theory, the data controller does

---

[6] Walden Ian., "Data Protection", in Reed Chris, Angel John, *Computer Law*, 5th edition, Oxford University Press, 2003, p. 432.

[7] European Parliament & the Council of the European Union, Directive 1995/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31.

[8] Article 6(1)(c) Data Protection Directive.

[9] Article 6(1)(e) Data Protection Directive.

not have full autonomy in making this decision: the data controller will need to be able to justify why certain data was collected and/or retained for processing, when requested by the relevant Data Protection Authority or by the data subject himself when exercising his rights. If the data controller cannot provide an adequate justification, then the processing of personal data will be in violation of applicable data protection rules, and might therefore result in the liability of the data controller. Thus, the Data Protection Directive provides a theoretical incentive to data controllers to conduct this assessment responsibly.

The importance of the *principles* of data minimisation and of conservation, which are in practice specific aspects of the proportionality principle, has been demonstrated in a recent Eurobarometer survey on the attitudes on data protection and electronic identity in the European Union.[10] According to the survey, 43% of Internet users say they have been asked for more personal information than necessary when they proposed to obtain access to or use an online service and 70% of Europeans are concerned that their personal data held by companies may be used for a purpose other than that for which it was collected. Moreover, 75% of Europeans want to delete personal information on a website whenever they decide to do so.[11] However, the 2010 Annual Report published by the Irish Data Protection Commissioner presents a different picture, by examining the actual complaints registered with the Commissioner (rather than measuring consumer opinion, as the Eurobarometer does). When looking at these complaints, only 0.64% of the total complaints received by the Commissioner refer to the requesting of excessive data, while a greater concern is expressed in relation to the disclosure of personal data, as this represented the third highest category of complaint – making up 10.47% of total complaints.[12] Thus, the stated consumer concern does not appear to be reflected in consumer protest. The same observation was made in an ENISA study on the economics of privacy[13].

There may be a need in particular cases to specify the principles of data minimisation and of conservation, either in a legal provision, or via an opinion of the Data Protection Authority or in another way, such as via the request for specific authorisation by a competent entity, for instance in order to acquire the authorisation for secondary processing of personal data. In Sweden, for example, the Swedish Data Inspection Board has issued several decisions where companies were ordered to delete or anonymise personal data before the time when they generally used to delete or anonymise them. The Swedish Data Inspection Board published for example specific

[10] Eurobarometer, Attitudes on Data Protection and Electronic Identity in the European Union, Special Eurobarometer 359, available online at http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf (last accessed on 16.12.2011)

[11] Idem, p. 6.

[12] http://www.dataprotection.ie/documents/annualreports/2010AR.pdf (last accessed on 18.12.2011)

[13] ENISA, Study on Monetizing Privacy. An Economic Model for Pricing Personal Information To be publish at the beginning on 2012 on ENISA web page: http://www.enisa.europa.eu/act/it/library

decisions on the deletion of data by the Postal Office[14], by travel agents[15], in the context of video surveillance in grocery stores[16]. The Swedish Data Inspection Board has also published a decision on the storage of customer data by the Swedish train company SJ, which is analysed in detail below in section 3.1.

Indications on acceptable storage periods are sometimes also provided through indirectly related legislation. According to the Dutch Act on Personal Data Protection[17], any automated processing of personal data has to be notified to the Dutch Data Protection Authority. As notifying every automated processing of personal data would be excessive at times, the Dutch legislator provided for various exemptions from the notification obligation. To this end, the so-called Exemption Decree[18] lays down certain categories of data processing which are unlikely to infringe the fundamental rights and freedoms of the data subject and which are therefore exempted from the notification requirement referred to in the Data Protection Act. This Exemption Decree provides an indication of a reasonable storage period for certain personal data. For instance data of customers and suppliers and entities that have a similar role, such as retailers and their standard clients, libraries and readers etc must be deleted two years after the carrying out of the relevant transaction.[19]

3 The Collection and Storage of Personal Data in the Transportation Sector: From Principle to Practice

3.1 Online Booking and Purchasing of Tickets

The booking and purchasing of tickets for public or private transportation is an everyday activity that can be carried out by natural persons either online or offline. The procedures established in various Member States for the booking of the ticket, as well as for its actual payment, differ significantly depending on the type of the means of transportation. The ENISA study examined the purchasing of a ticket online from a private or public transportation company in each of the twenty seven European Member States. Seventeen railway company cases were identified, along with three bus company cases, three airline companies, two online travel agencies, a ferries company

[14] http://www.datainspektionen.se/press/nyhetsarkiv/2008/posten-lagrar-personuppgifter-onodigt-lange/ (last accessed on 16.12.2011)

[15] http://www.datainspektionen.se/press/nyhetsarkiv/2009/charterbolagen-lagrar-kunduppgifter-och-resehistorik-for-lange (last accessed on 16.12.2011)

[16] http://www.datainspektionen.se/Documents/beslut/2011-06-20-lidl.pdf (last accessed on 16.12.2011)

[17] Wet bescherming persoonsgegevens (WBP), 06.07.2000 (O.J. 302/2000); see http://wetten.overheid.nl/BWBR0011468 (last accessed on 20.12.2011)

[18] Vrijstellingsbesluit WBP, http://www.cbpweb.nl/hvb_website_1.0/vwc11.htm (last accessed on 16.12.2011)

[19] Idem.

and finally a ski ticket purchasing process. All but one surveyed transportation companies offer also alternative ways of purchasing tickets, i.e. by telephone or in person at the offices of the company.

The booking of a ticket online from a transportation company gave the opportunity to examine the obligatory types of personal data of the customers that were collected for the completion of the booking in relation to the principle of data minimisation and to examine whether transportation companies carry out excessive collection of personal data during the booking process (Figure 1). All transportation companies required the first and last name of the passenger and all but one required a valid e-mail address. The e-mail address did not need to belong to the passenger, but could also e.g. belong to the person that realised the purchasing. Sixteen of the surveyed companies required a fixed or mobile phone number, while ten of them asked for a postal address.

Interestingly, six of the surveyed companies required an identity card or passport number. These six companies did not belong to the same category of transportation companies, but offered various types of tickets online, i.e. railway tickets, bus tickets and ferries tickets. Depending on the surveyed transportation company several other types of data were required for the booking of a ticket online, such as the gender or the title of the passenger, date of birth, nationality etc. Additional information on the passenger was sometimes required in order to justify discounts (for instance age of the passenger for youth or senior ticket). The fragmentation on the types of data that were required by various transportation companies for the booking of a ticket online revealed a challenge for the principle of minimal disclosure. Although the transportation companies may wish to collect as much personal data about their customers as possible (e.g. to conduct market research into key consumer profiles), this cannot be justified under the principle of data minimisation which stipulates that only the necessary information should be collected and stored.

The study also examined the options that transportation companies (either private or public ones) offered to their customers with regard to the processing of customer data for the sending of information and for marketing purposes (Figure 2). The majority of transportation companies processed as a default the personal data of their customers for the sending of information about their products and services as well as for marketing purposes. In several websites there was a tick-box already pre-checked, which the users would have to uncheck if they did not wish to receive such information.

In some other cases, information about the processing of the personal data of the customers was contained in the privacy statement or the Terms and Conditions of the website. The users were given the opportunity to refuse the processing of their information for such purposes via sending an e-mail to a dedicated e-mail address or via configuring the relevant option in their account on the website. In almost one third of the surveyed companies the users could consent to the processing of their data in order to receive promotions and news of the company or for marketing purposes by ticking a checkbox. In two of the surveyed companies the fact that data can be used for marketing purposed only after the explicit consent of the user, is mentioned in the privacy policy. In these cases, the users have to explicitly give such permissions via their account.

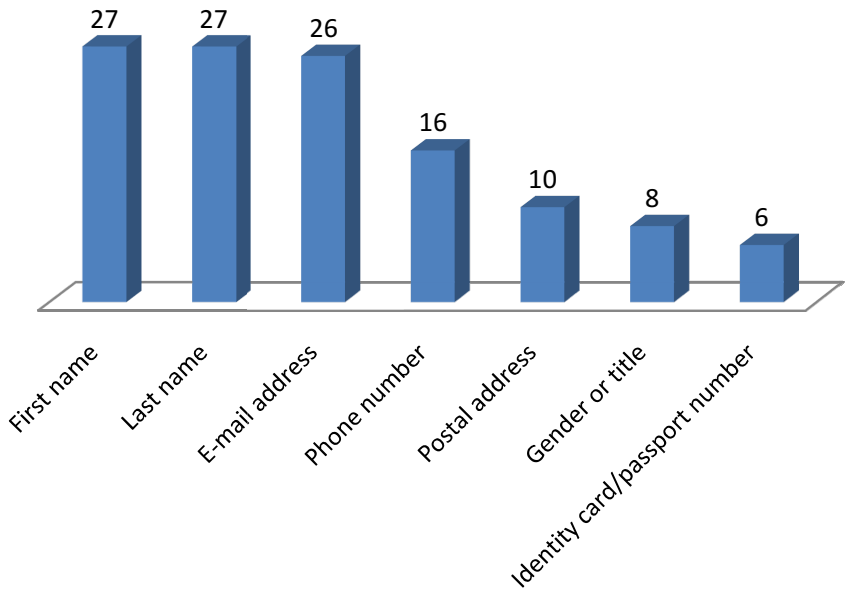


Fig. 1. Types of personal data collected when booking the ticket online in 27 MS

Only one surveyed company collected personal data from its customers only in order to process purchasing requests without collecting data for any other marketing purposes. To the contrary, another surveyed company (Malta), which by default implemented a right to sell or otherwise communicate the contact details of an individual to third parties for marketing purposes, did not even allow the users to unsubscribe or refuse the processing of their data for specific purposes. Specifically, the privacy notice of the transportation company mentioned that "We also reserve the right to send all customers of our service email communications from time to time regarding updates and changes to our goods and services, new links to our website and any technical, administrative or legal notices important to our website, our products and services that we consider essential. **Customers are not able to unsubscribe from these notices**." (emphasis added). Finally, one of the surveyed companies did not offer any kind of option and did not inform its users with regard to the processing of their data for marketing purposes and for the sending of promotions and news of the company.

The ENISA study showed that the lack of specific legislation or policy documents on the collection and storage of personal data in the transportation sector has led to a lack of harmonisation in relation to the storage period of the personal data of the users and the customers. At least four of the surveyed companies (in Greece, Hungary, Romania and Slovakia) did not even have a privacy policy that would inform the users of the types of data that are collected and their storage period, while in the majority of the cases where a privacy policy did exist, the users were informed about the

use of cookies on the website, but not about the storage period of their data. In one of the surveyed companies offering online purchasing of bus tickets, the personal data of the passenger, more specifically the first and last name of the passenger, their phone number and birth date, were stored for a maximum period of ten years. It was surprising to note that several of the online transportation companies surveyed did not contain a privacy policy or any kind of document that would inform their users about the processing of their personal data.

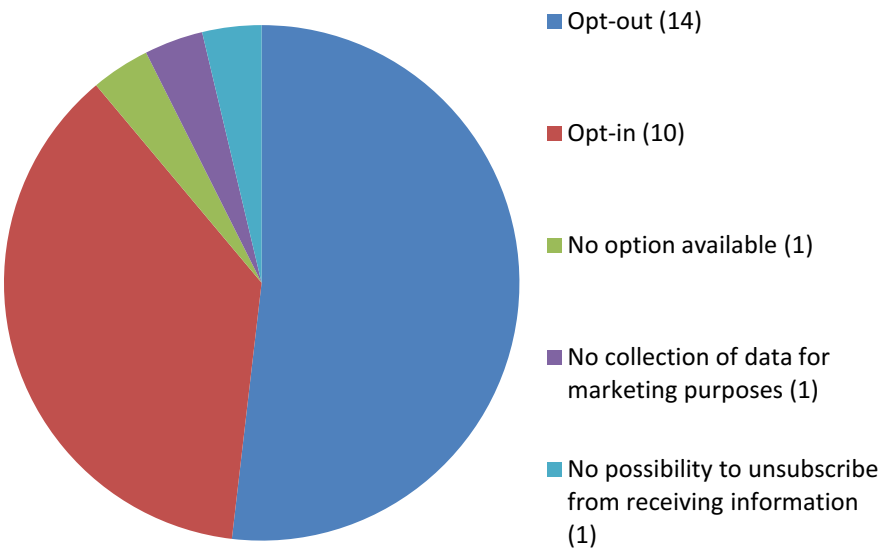


Fig. 2. Options offered by transportation companies to customers regarding the sending of information by the company in the future based on the data they collect on them

The Swedish train company SJ was investigated in 2008 by the Swedish Data Inspection Board as it stored customer data on certain travel cards. SJ was storing personal data on the travel history of the passengers for statistical purposes and for customer complaints. The Swedish Data Inspection Board adopted on 22 December 2008 a decision ordering SJ to anonymise the data relating to travel history 90 days after the departure date at the latest.[20] As highlighted by the Swedish report, in earlier decisions, the DIB had ordered maximum retention periods of 60 days, but in SJ's case the period for customers to reclaim a journey is 3 months, so 90 days were deemed adequate.

[20] Decision no 711-2008, available in Swedish at `http://www.datainspektionen.se/Documents/beslut/2008-12-23-sj.pdf` (last accessed 17.12.2011)

In 2009, the Belgian railway company[21] (Belgium's national railway company) introduced a "ticketless" way of travelling on their railway system, by enabling their customers to link their citizen's National Registry number with the ticket number via their eID card[22]. When travelling, the user will have to show his eID card to the train attendant in order to verify the purchasing of the train ticket. The transfer of personal data in this case is inherently excessive, as the Belgian eID belongs traditionally to the first generation of eIDs and has implemented an "all or nothing" model.[23] This means that the citizen, when he wishes to use his elD, has to disclose all the personal data that are stored in his card and does not have the opportunity to choose which types of personal data he would like to disclose. In this way the citizen reveals an abundance of personal information for the purchasing of the train ticket, which is undoubtedly not necessary for the purpose of purchasing a train ticket and the verification that it has been paid. Such an application puts the respect to the principle of minimal disclosure into question. The Belgian DPA issued a recommendation on transport e-ticketing in 2010, stating that e-ticketing should never allow transportation companies to trace the travel route of individual travellers.[24]

3.2 Payment for Purchasing on Online Tickets

The data that are collected either by the transportation company or by an intermediary company that carries out the payment of the ticket (following the booking/purchasing process as described in section 3.1 directly above) are to a large extent common throughout the European Union. For instance for the payment by Visa (or MasterCard or Maestro) the following personal data are required: the card holder's name, the card number, the expiration date of the card and the secure code (CVV2 or CVC2). In Hungary, the name of the bank issuing the card is also required.

Electronic Ticket Cards. The online purchasing of tickets in the transportation sector poses challenges in the way how (and whether) the principle of minimal disclosure is respected in this field. Similar concerns have been raised for the use of electronic travel cards, which require the user to reveal a number of personal information when purchasing the card online. Users tend to reveal a large amount of personal information and leave traces of their location at various time points for the sake of "convenience". The traditional paper ticket used for public means of transportation is gradually being replaced by electronic cards, such as the Oyster card in London or the MoBIB card in Brussels, which allow the user to use the public transportation system in an easy and uninterrupted way. However, the unique number that is stored on the card allows for

[21] http://www.b-rail.be (NMBS/SNCB) (last accessed on 25.01.2012)

[22] http://mobile.b-rail.be/en/Novelties/Use-your-Belgian-e-ID-as-ticket (last accessed on 25.01.2012)

[23] Van Alsenoy Brendan & De Cock Danny, Due processing of personal data in eGovernment? A case study of the Belgian electronic identity card, Datenschutz und Datensicherheit 3/2008, p. 178.

[24] http://www.privacycommission.be/nl/docs/Commission/2010/aanbeveling_01_2010.pdf (last accessed on 25.01.2012)

the tracking of the location of the user and, when combined with the identification data of the user that may be revealed when the electronic ticket card has been purchased via a credit or debit card, it offers a rich amount of personal information that can be used for user tracking and user profiling. In Denmark, a new national electronic travel card is planned to be launched in 2012. According to information in the press, travellers will have to provide their name, address and e-mail address, but also bank account information and their personal identification number. The card scheme foresees the possibility for travellers to get an anonymous travel card, but at a higher cost. This approach, in which privacy protection is essentially treated as a common barter, has created a heated debate in Denmark. In this section some further prominent examples will be presented in greater detail, along with the challenges they pose to the principle of minimal disclosure.

The London Oyster Card. The London 'Oyster' card was implemented in 2003 and has been severely criticised over the collection of excessive data of the users, as well as for enabling their tracking and tracing. Transport for London (TfL) collects the following information about the users of the Oyster card: title (Mr/Mrs/Ms/Miss etc), first name, middle initial and surname, address and a password. When a user applies for a card online, their telephone number and email address also have to be supplied. When a user is purchasing the Oyster card using a debit or credit card, the encrypted bank details are stored. When the user is making use of the service for an automatic top-up, then TfL also stores the history of the transactions, including location, date and time. Finally the Oyster ticketing system records the location, date and time an Oyster card was used to validate a journey on TfL's network or on National Rail services where Oyster is accepted.[25]

The amount of personal data collected by Transport for London through the Oyster card service has been criticised, especially in relation to children that wish to travel at a discounted rate. They must apply for a photocard ID and provide their name, date of birth, address, school name and telephone number, data that have been deemed as excessive in relation to the purpose of issuing a transportation card.

The data are stored for a period of **eight weeks**, a time period that was agreed in consultation with the Information Commissioner's Office (the UK data protection authority), when the card was first implemented in 2003.[26] The data are then anonymised and used for research purposes. According to the website of TfL, the Oyster ticketing system is being changed so that it will retain customers' names and contact details for **two years** after the customer last used their card or bought an Oyster product.[27]

The details of debit or credit cards that are used to buy Oyster products are retained for a maximum period of 18 months.[28] When a user is issued a penalty fare notice or

[25] http://www.tfl.gov.uk/termsandconditions/12321.aspx#page-link-what-personal-details-are-held-about-oyster-customers- (last accessed 05.11.2011)

[26] Idem.

[27] Idem.

[28] Idem.

prosecuted for fare evasion, their personal details and relevant journey and transaction history will be retained for a longer period, which is not specified.[29]

There is an ongoing debate in the United Kingdom about how long TfL should hold the data and to what extent is it acting proportionately when it decides to either comply with data requests from the police or withhold information in order to protect peoples' privacy. This debate has been stimulated by the increasing number of requests for data on Oyster card passenger movements from the Metropolitan Police in connection with criminal investigations.

The Paris Navigo Pass. The adoption of the 'navigo pass' for the Paris region, which is similar to the Oyster card, has been in the centre of similar debates in France. Due to the fact that the user could be banned from the use of the 'navigo pass' in cases of delayed payments, the processing of the personal data of the user in relation to the 'navigo pass' had to be authorised by the French data protection authority, the CNIL. The CNIL issued in 2008 a single authorisation[30] for ticketing systems, which was updated in 2010,[31] covering any kind of data processing in the context of ticketing systems that should comply with a series of guarantees defined by the CNIL. The single license for ticketing systems is directed to those systems that imply the processing of personal data for the following purposes: management, delivery and use of transportation tickets, fraud management, statistical analysis of the use of the network, quality assessment of the functioning of the system. The CNIL specified the types of personal data that should be processed, depending of the type of ticketing, enforcing in this way the principle of data minimisation in the transportation sector.

---

[29] http://www.tfl.gov.uk/termsandconditions/12321.aspx#page-link-how-long-does-tfl-keep-oyster-information-- (last accessed on 05.11.2011)

[30] Single authorisation AU-015 - Decision No. 2011-107 of 28 April 2011 authorizing single implementation of automated processing of personal data relating to the management of ticketing applications by operators and public transport authorities (Autorisation unique n° AU-015 - Délibération n° 2011-107 du 28 avril 2011 portant autorisation unique de mise en œuvre de traitements automatisés de données à caractère personnel relatifs à la gestion des applications billettiques par les exploitants et les autorités organisatrices de transport public), available online at http://www.cnil.fr/en-savoir-plus/deliberations/deliberation/delib/136/ (last accessed on 17.12.2011) Single authorisations may be issued by the CNIL in accordance with Article 25II of the French Data Protection Act when the processing of personal data meets a single purpose, relating to categories of the same data and have the same recipients or categories of recipients.

[31] In 2010, a working group of the CNIL in collaboration with GART (Grouping of transport authorities) was formed to identify new practices in public transportation (http://www.cnil.fr/dossiers/deplacements-transports/actualites/article/les-systemes-billettiques-evoluent-lautorisation-unique-n15-aussi/, last accessed on 17.12.2011) New practices such as post-payment or access to multiple services with the same media called for additional guidelines, while led the CNIL to amend on 28 April 2011 its single license on three topics: anonymity, media tickets and post-pay.

According to the CNIL authorisation, all customer data are kept for the full duration of the contractual relationship and upon the end of it for two years for commercial and statistical purposes. The validation data that reveal information about the movements of the users, should be anonymised 'shortly'. The anonymisation can take place either by completely removing the card number or the joint date, time and place of the journey, or by applying a cryptographic algorithm (a public 'hash') that is deemed safe to the card number. However, the validation data containing information about the movement of people associated with the card number or subscriber and referring indirectly to the identity of a user, may be retained for forty-eight hours and solely for the fight against technological fraud.

During the 2010 amendments, the CNIL distinguished three types of tickets, depending on the anonymity achieved for the user:

- the nominative ticket, such as the 'navigo pass' in the Paris region,
- the declarative ticket which allows anonymity and cannot be replaced if lost or stolen, and
- the anonymous ticket, which in practice only allows the loading of single tickets.

The study showed that some authorities, for financial and practical reasons, do not offer special rates (reduced rates or free) on declarative tickets. The CNIL however considers that software vendors are now developing and maps declarative tickets that would support such solution. The name, first name and photograph of the holder of the pass can be scanned on the support (without being integrated into the customer file) and a receipt is issued at the time the ticket is loaded (linking the identity of the holder and number to the declarative password). Such a solution reduces the risk of fraud and helps preserve the anonymity of travel for recipients of social tariffs and the resurfacing of the past in case of loss or theft. The CNIL recommends that special rates are also made available on declarative support.

With regard to the principles of data minimisation and data conservation, the amendment to the CNIL authorisation on 'post-payment' is of particular interest. Transportation authorities in France are developing public transportation services where the billing is based on the actual journeys conducted and it takes place after the service has been offered. As certain information on the journeys made will be needed for the billing of routes and for the resolving of customer complaints, the CNIL specified that only data that are strictly necessary to calculate the price should be collected. Therefore, the information revealing the place where the ticket has been purchased (the station of validation) is not justified to be processed as it is not necessary for the calculation of the price and it would not be in line with the right of the citizens to come and go anonymously. With regard to the storage period of the processed personal data, the CNIL specified that they may be retained for a period of four months from the date of the events –and not from the moment when the billing takes place. Finally, information on the management of overdue payment should be immediately removed from the black list from the moment the amounts due are paid and by default, within maximum two years from registration.

The Brussels MoBIB Card. In 2008, the Brussels public transportation company[32], launched the 'MoBIB' card.[33] The MoBIB card is equipped with a Radio Frequency Identification (RFID) chip, on which the name, last name, date of birth and postal code of the user are stored. The information relating to the programme that the user has chosen (10-journeys ticket, 1 day ticket etc) is also stored on the card, along with the information on the last three uses of the card. A photo of the user is printed on the card.[34]

The Brussels public transportation company claims that the location information of a user is never processed, while such processing only takes place based on encoded or anonymous information. However, the implementation of the MoBIB has been criticised as violating the Belgian legislation on the protection of personal data.[35]

The Belgian Privacy Commission adopted a recommendation in March 2010 in which it pointed out that the direct or indirect processing of personal data of the users in order to trace the route they are following via their electronic ticket is not allowed.[36]

The Brussels public transportation company mentions in the terms of use of the MoBIB card that the data will be stored for limited periods of time as necessary for the specific foreseen processing. No exact storage period is however specified.[37] The Belgian Privacy Commission has advised in its recommendation 01/2010 that the data

[32] http://www.mivb.be (STIB/MIVB) (last accessed on 25.01.2012)

[33] http://www.stib.be/mobib.html?l=en (last accessed on 25.01.2012)

[34] http://www.mivb.be/pointdevue_Standpunt.html?l=nl&news_rid=/STIB-MIVB/INTERNET/ACTUS/2010-05/WEB_Article_ 1274963883674.xml (last accessed on 05.11.2011)

[35] http://www.brusselnieuws.be/artikel/garandeert-nieuwe-mobib-chipkaart-anonimiteit-van-reiziger
http://www.brusselnieuws.be/artikel/liga-mensenrechten-mobib-schendt-het-priv%C3%A9leven
(last accessed on 05.11.2011)

[36] Commissie voor de bescherming van de persoonlijke levenssfeer (Belgian Privacy Commission), Aanbeveling nr 01/2010 van 17 maart 2010, Aanbevelling over de na te leven basisbeginselen bij het gebruik van e-ticketing door de openbare vervoersmaatschappijen (Recommendation 01/2010 on the fundamental principles that have to be respected during the use of e-ticketing by the public transportation companies) (A-2010-003), 17 March 2010, available online at http://www.privacycommission.be/nl/docs/Commission/2010/aanbeveling_01_2010.pdf (last accessed on 05.11.2011). The Belgian Privacy Commission adopted also in 2009 an Opinion on the application of the Belgian Data Protection Act to the processing of personal data in RFID systems: Commissie voor de bescherming van de persoonlijke levenssfeer (Belgian Privacy Commission), Advies nr 27/2009 van 14 oktober 2009 uit eigen beweging inzake RFID (Opinion 27/2009 relating to RFID) (A/2009/003), 14 October 2009, available online at http://www.privacycommission.be/nl/docs/Commission/2009/advies_27_2009.pdf (last accessed on 05.10.2011)

[37] http://www.stib.be/utilisation_gebruik.html?l=nl (last accessed on 05.11.2011)

that are collected for travel ticket administration should be deleted at the latest after six months.[38] The Belgian Privacy Commission also recommended that the client data of the users should be deleted within 12 months after the last use of the card, or after the time when the customer has returned the card.[39]

The Prague 'Opencard'. In 2008 the Prague City Hall launched an electronic card called 'Opencard', which can be used for public transportation in Prague, can function as a library card for the municipal Library or as the means for discount programmes, and also includes an application for payment of parking fees.[40] The card can be issued with a monthly, quarterly or annual validity.

For the issuing of an Opencard, a number of personal data of the traveller are processed and stored. The first name, the last name and a photograph of the card holder are printed on the card. According to the Opencard website, these data serve for the verification of the card holder's identity during some operations such as public transport inspections.[41] In addition, the date of birth of the traveller is stored in an encrypted way in the contactless chip of the Opencard. The justification for the processing of this information is that the date of birth is needed when applying for age-related discounts.[42]

Following the introduction and widespread deployment of the Opencard, the Czech Office for Personal Data Protection issued a statement urging the Prague City Hall to offer, besides the traditional Opencard, an anonymous alternative for which no personal data of the traveller need to be processed. The Prague City Hall complied with this request and launched in December 2011 an anonymous Opencard that does not contain any personal data and is transferable. The anonymous travel cards in Prague were introduced in full respect of the data minimisation principle, allowing citizens to exercise their right to come and go anonymously.

---

[38] Commissie voor de bescherming van de persoonlijke levenssfeer (Belgian Privacy Commission), Aanbeveling nr 01/2010 van 17 maart 2010, Aanbevelling over de na te leven basisbeginselen bij het gebruik van e-ticketing door de openbare vervoersmaatschappijen (Recommendation 01/2010 on the fundamental principles that have to be respected during the use of e-ticketing by the public transportation companies) (A-2010-003), 17 March 2010, p. 5, available online at `http://www.privacycommission.be/nl/docs/Commission/2010/aanbeveling_01_2010.pdf` (last accessed on 05.11.2011)

[39] Commissie voor de bescherming van de persoonlijke levenssfeer (Belgian Privacy Commission), Aanbeveling nr 01/2010 van 17 maart 2010, Aanbevelling over de na te leven basisbeginselen bij het gebruik van e-ticketing door de openbare vervoersmaatschappijen (Recommendation 01/2010 on the fundamental principles that have to be respected during the use of e-ticketing by the public transportation companies) (A-2010-003), 17 March 2010, p. 6, available online at `http://www.privacycommission.be/nl/docs/Commission/2010/aanbeveling_01_2010.pdf` (last accessed on 05.11.2011)

[40] `http://opencard.praha.eu/jnp/en/home/index.html` (last accessed on 17.12.2011)

[41] `http://opencard.praha.eu/jnp/en/about/security/index.html` (last accessed on 17.12.2011)

[42] Idem.

Dutch OV-Chipcard. The OV-chipcard has recently been introduced in the Netherlands, as a smart card with a built-in chip for public transportation. There are currently three types of OV-chipcards: a personalised one, which mainly aims at season ticket holders; a disposable card which can be used for a certain period of time; and an anonymous one. The Dutch Data Protection Commission carried out an investigation with regard to the processing of personal data relating to the use of student OV-chipcards. The Commission found that four companies[43] were storing personal data for a longer period than was necessary. The transportation companies modified the **storage period** of the personal data[44] they were collecting in relation with the student OV-chipcards in order to be in line with the conservation principle and adopted storage periods mainly varying between 18 and 24 months depending on the purposes.[45] The Commission imposed an order for incremental penalty payments if the companies do not comply with the order.

3.3 Airline Companies and PNR Data

Concept and Legal Background. The purchasing of airplane tickets, either online or offline, especially for flights into the U.S. (or even Canada or Australia) requires the revealing of a large number of personal information of the user. The most well known example of such a data transfer mechanism is the so-called Passenger Name Record (PNR) data. PNR data[46] is information that is provided by passengers and is collected by carriers for enabling reservations and carrying out the check-in process.[47]

---

[43] The Amsterdam-based transportation company GVB, the Rotterdam-based transportation company RET, the transportation company NS and the cards issuer TLS.

[44] http://www.cbpweb.nl/Pages/pb_20110726_OV-chip_LOD.aspx (last accessed on 17.12.2011)

[45] http://www.cbpweb.nl/downloads_pb/pb_20110726_OV-chip_LOD_TLS.pdf, http://www.cbpweb.nl/downloads_pb/pb_20110726_OV-chip_LOD_NS.pdf, http://www.cbpweb.nl/downloads_pb/pb_20110726_OV-chip_LOD_RET.pdf, http://www.cbpweb.nl/downloads_pb/pb_20110726_OV-chip_LOD_GVB.pdf (last accessed on 25.01.2012)

[46] It should be noted that PNR data are different from Advance Passenger Information (API), which has to be communicated by air carriers at the request of the authorities responsible for carrying out checks on persons at external borders (Council Directive 2004/82/EC of 29 April 2004 on the obligation of carriers to communicate passenger data, OJ L261/24, 06.08.2004). API data are the biographical information taken from the machine-readable part of a passport and contain the name, place of residence, place of birth and nationality of a person.

[47] European Commission, Communication from the Commission on the global approach to transfers of Passenger Name Record (PNR) data to third countries, COM(2010) 492, Brussels, 21.09.2010., p. 3, available online at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0492:FIN:EN:PDF (last accessed on 07.10.2011)

The record that is created on each of the passengers contains data, such as the dates of travel and the travel itinerary, ticket information, contact details, address and phone numbers, the travel agent that was involved in the booking of the ticket, payment information, seat number and baggage information.[48]

The European Union has signed agreements for the transfer of PNR data with the U.S., Canada and Australia. In 2004, the Council of the European Union adopted a Decision concerning the conclusion of an agreement between the European Community and the United States of America on the processing and transfer of passenger name record (PNR)[49] data by air carriers to the United States Bureau of Customs and Border Protection (CBP) and a Decision was also adopted by the European Commission on the adequate protection of those data[50]. The 2004 PNR agreement of the transfer of personal data of passengers between the European Union and the United States Government foresaw that 34 data elements has to be provided to the US Customs Bureau for each passenger. The European Court of Justice in a judgement adopted in 2006[51] annulled the aforementioned decisions.

The Court ruled that the "transfer of PNR data to CBP constitutes processing operations concerning public security and the activities of the State in areas of criminal law"[52]. Although the data have been initially collected for commercial purposes, the Court found that the actual purpose of their transfer falls within a framework established by the public authorities that relates to public security and thus the processing falls outside the scope of protection of the data protection directive. The Court followed the argumentation of the General Advocate and distinguished between the activities of collection of data and the purpose of the (further) processing based on public safety needs, in order to exclude the latter from the scope of application of the data protection directive. The Court judgement can be briefly described as admitting that the data collected for commercial purposes fall within the protective ambit of the Data Protection Directive but when the same data are further transferred for public security reasons, they no longer enjoy the same protection. The Judgment of the European Court of Justice created a substantial *lacuna legis* in the protection of PNR data, raising the general problem of protection of personal data that are not covered by

---

[48] Idem. See below Section 3.3.2 for the detailed list of PNR data in the context of the EU-US PNR draft agreement.

[49] Council of the European Union, Council Decision of 17 May 2004 on the conclusion of an Agreement between the European Community and the United States of America on the processing and transfer of PNR data by Air Carriers to the United States Department of Homeland Security, Bureau of Customs and Border Protection (2004/496/EC), [2004] OJ L183/83.

[50] Commission of the European Communities, Commission Decision of 14 May on the adequate protection of personal data contained in the Passenger Name Record of air passengers transferred to the United States' Bureau of Customs and Border Protection [2004] OJ 235/ 11.

[51] Judgment of the Court of Justice in Joined Cases C-317/04 and C-318/04 (30 May 2006), ECR 2006, p. I-4721.

[52] Paragraph 56 of the Judgment of the Court of Justice in Joined Cases C-317/04 and C-318/04 (30 May 2006).

the Data Protection Directive[53]. The European Parliament had raised issues relating to the respect to the proportionality principle, although the Court did not consider this issue.

The European Commission recently proposed a Directive of on the use of Passenger Name Record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime (PNR Directive)[54], as well as a Proposal for a Council decision on the conclusion of the Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Records to the United States Department of Homeland Security.[55] The Article 29 Data Protection Working Party, as well as the European Data Protection Supervisor, have criticised the European Commission initiatives on PNR data with regard to the list of data that have to be transferred, as well as on the storage period of the PNR data.[56]

---

[53] See also the analysis made by Hielke Hijmans, in HIJMANS Hielke 'De derde pijler in de praktijk: leven met gebreken Over de uitwisseling van informatie tussen lidstaten'. SEW 2006.91, under chapter 4.1.

[54] European Commission, Proposal for a Directive of the European Parliament and of the Council on the use of Passenger Name Record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, COM(2011) 32 final, Brussels, 02.02.2011.

[55] European Commission, Proposal for a Council decision on the conclusion of the Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Records to the United States Department of Homeland Security, COM(2011) 807 final, Brussels, 23.11.2011.

[56] European Data Protection Supervisor, Opinion of 09.12.2011 on the Proposal for a Council Decision on the conclusion of the Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Records to the United States Department of Homeland Security; European Data Protection Supervisor, Opinion of 15.07.2011 on the Proposal for a Council Decision on the conclusion of an Agreement between the EU and Australia on the processing and transfer of PNR data by air carriers to the Australian Customs and Border Protection Service; Article 29 Data Protection Working Party, Opinion 10/2011 on the proposal for a Directive of the European Parliament and of the Council on the use of passenger name record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, WP181 (05.04.2011); European Data Protection Supervisor, Opinion of 25.03.2011 on the Proposal for a Directive of the European Parliament and of the Council on the use of PNR data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime; Article 29 Data Protection Working Party, Opinion 7/2010 on European Commission's Communication on the global approach to transfers of Passenger Name Record (PNR) data to third countries, WP 178 (12.11.2010); European Data Protection Supervisor, Opinion of 19.10.2010 on the global approach to transfers of PNR data to third countries; European Data Protection Supervisor, Opinion of 20.12.2007 on the Proposal for a Council Framework Decision on the use of PNR data for law enforcement purposes; Opinion 5/2007 on the follow-up agreement between the European Union and the United States of America on the processing and transfer of passenger name record (PNR) data by air carriers to the United States Department of Homeland Security concluded in July 2007, WP138 (17.08.2007); Article 29 Data Protection Working Party, Opinion 4/2003 on the Level of Protection ensured in the US for the Transfer of Passengers' Data, WP78 (13.06.2003).

The Principle of Data Minimisation and PNR Data. According to the recent proposal for a Council decision on the transfer of PNR data from the European Union to the United States Department of Homeland Security (DHS), an abundance of personal data of all passengers that are flying to and from the European Union have to be collected irrespective of the fact whether they are suspected of any wrongdoings. According to the Annex to the agreement, the following nineteen types of data would have to be collected by the airlines companies and be transferred to the DHS: (1) PNR record locator code, (2) date of reservation/issue of ticket, (3) date(s) of intended travel, (4) name(s), (5) available frequent flier and benefit information (i.e., free tickets, upgrades, etc.), (6) other names on PNR, including number of travellers on PNR, (7) all available contact information (including originator information), (8) all available payment/billing information (not including other transaction details linked to a credit card or account and not connected to the travel transaction), (9) travel itinerary for specific PNR, (10) travel agency/travel agent, (11) code share information, (12) split/divided information, (13) travel status of passenger (including confirmations and check-in status), (14) ticketing information, including ticket number, one way tickets and Automated Ticket Fare Quote, (15) all baggage information, (16) seat information, including seat number, (17) general remarks including OSI, SSI and SSR information, (18) any collected Advance Passenger Information System (APIS) information, (19) all historical changes to the PNR listed in numbers 1 to 18.

The European Data Protection Supervisor (EDPS) noted that the aforementioned types of data would be collected and stored not only for passengers, but also for prospective passengers who may cancel their trip. The list of data was considered as excessive and disproportionate compared to the purposes pursued via the proposed Council decision. The EDPS proposed limiting the data to the following information: "PNR record locator code, date of reservation, date(s) of intended travel, passenger name, other names on PNR, all travel itinerary, identifiers for free tickets, one-way tickets, ticketing field information, ATFQ (Automatic Ticket Fare Quote) data, ticket number, date of ticket issuance, no show history, number of bags, bag tag numbers, go show information, number of bags on each segment, voluntary/involuntary upgrades, historical changes to PNR data with regard to the aforementioned items".[57] As for the processing of sensitive data, the EDPS recommended that airline carriers should not transfer any sensitive data to the DHS.[58]

The Maximum Period of Storage and PNR Data. According to the proposal for the PNR Directive of 02.02.2011, the PNR data would have to be retained for a period of 30 days in a database at the Passenger Information Unit[59] for a period of 30 days after

[57] European Data Protection Supervisor, Opinion of 09.12.2011 on the Proposal for a Council Decision on the conclusion of the Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Records to the United States Department of Homeland Security, p. 5.

[58] Idem.

[59] A Passenger Information Unit is a single designated unit that should be created in each Member State and will be responsible for handling and protecting the data (if the PNR Directive is adopted).

their transfer to the Passenger Information Unit of the first Member State on whose territory the international flight is landing or departing. Upon expiry of the period of 30 days after the transfer of the PNR data to the aforementioned Passenger Information Unit the data shall be retained, masked out, at the Passenger Information Unit for a further period of five years.[60] The Article 29 Data Protection Working Party considers the retention period of five years as disproportionate.[61]

The European Commission proposal for a Council decision of 23.11.2011 on the transfer of PRN data from the EU to the US DHS foresees even longer storage period for the PNR data. In accordance with Article 8 of the proposal, DHS retains PNR data in an active database for up to five years. The data will be depersonalised and masked after the initial six months of this period, but the passenger will still be able to be identified. After this five-year period, the PRN data will be transferred to a dormant database for a period of up to ten years. According to the European Data Protection Supervisor, and similar to the position taken by the Article 29 Data Protection Working Party, the maximum retention period of fifteen years that is foreseen in the Proposal is disproportionate and excessive. Rather a retention period of six months is recommended.[62] The position of the EDPS requiring for a retention period of six months instead of the period of fifteen years that is currently proposed illustrates a significant challenge on defining what the appropriate storage and retention period would be for specific types of data. The general data protection principle on the conservation of data stipulating that personal data must be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed"[63] allows room for broad interpretation.

4 Conclusions

As shown by the examples above, new technologies can greatly improve the efficiency, user friendliness and security of transportation systems. However, from a data protection perspective it is worrisome to observe that so many systems deployed in real life do not follow a privacy by design approach, and insufficiently consider the data minimisation and data conservation principles. This can be seen in online ticket purchasing systems, where data collection practices vary quite widely between the

[60] Article 9 of the proposal for a PNR Directive.

[61] Article 29 Data Protection Working Party, Opinion 10/2011 on the proposal for a Directive of the European Parliament and of the Council on the use of passenger name record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, WP 181 (05.04.2011), p. 6.

[62] "The data should therefore be anonymised (irreversibly) or deleted immediately after analysis or after a maximum of 6 months": European Data Protection Supervisor, Opinion of 09.12.2011 on the Proposal for a Council Decision on the conclusion of the Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Records to the United States Department of Homeland Security, p. 5.

[63] Article 6(e) Data Protection Directive.

countries, despite the fact that data needs behind these services are relatively homogeneous. This would seem to indicate that these systems are designed without systematically considering how data collection can be minimized. This can also be seen in the mechanisms for obtaining the consent of data subjects, particularly for direct marketing purposes, where communication towards data subjects is often ambiguous and clearly slanted towards facilitating data collection and data processing, rather than towards protecting the privacy of the users of such services.

Similar observations can be made with respect to travel cards and PNR data. A commonly recurring trend appears to be that such technologies are developed and deployed with optimal usability and usefulness in mind, but without duly considering data protection implications. Only after these concerns are brought to light – either by data protection authorities, court cases or consumer complains – are the systems reviewed and updated to improve privacy friendliness. In some cases – PNR data being a prime example, as are several travel card deployments – no conclusive answer to privacy questions has been found yet, and existing practices still fail to appropriately observe the data minimisation and conservation principles.

Globally, the examples illustrate that there is a need for these principles to be strengthened in practice, through legislation and governance mechanisms that favour privacy by design, including a clear assessment of privacy impacts and the identification of more privacy conscious implementation alternatives, in order to ensure that the personal data of European citizens is proactively protected, instead of having to modify operational systems only after privacy problems come to light. Hopefully, the ongoing revision of the Data Protection Directive and its future successor will take some steps in that direction.

Acknowledgments. The authors would like to acknowledge the kind contributions and keen insights offered by ENISA in the course of conducting the Study on data collection and storage in the EU, including notably Dr. Rodica Tirtea, Dr. Demosthenes Ikonomou and their team, as well as the national correspondents who provided extensive information on the legal systems and administrative practices in their countries for the study.

ICT and Privacy: Barriers

Antonio Kung

Trialog, Paris, France
antonio.kung@trialog.com

Abstract. This paper identifies barriers for the handing of privacy issues caused by *Information and Communication Technologies* (ICT). It reports first on experience gained in addressing privacy issues in *Intelligent Transport systems* (ITS). It discusses two applications, eCall and Pay-Per-Use. It identifies barriers for privacy and suggests recommendations. These barriers are at the application level (conflict of interest, lack of consensus on protection policies), at the design level (agreement on the meaning of Privacy-by-Design, neglect of architecture impact, lack of practice) and at the implementation level (leaks created by ICT infrastructures, lack of flexibility).

Keywords: Privacy-by-Design, Minimisation, Enforcement, Transparency, PET, Architecture, PEAR.

1 Introduction

The European Commission adopted in 2010 the directive 2010/40/EU [1] in order to address the compatibility, interoperability and continuity of Intelligent Transport Systems (ITS) solutions across the EU, for areas such as traffic and travel information, eCall emergency systems, and intelligent truck parking. The directive was preceded by the adoption of an Action Plan [2] in 2008. This action plan included four application areas, (1) optimal use of road, traffic and travel data, (2) continuity of traffic and freight management, (4) road safety and security, and (4) integration of vehicle and transport infrastructure. The eSafety initiative [3] provides more information on the many projects that were undertaken. The action plan also included a specific transversal area: data protection and liability, for which a series of research projects were started: SeVeCom [4], PRECIOSA [5], EVITA [6], OVERSEE [7] and PRESERVE [8]. They addressed secure communication, privacy, protection against vehicle intrusion, secure platforms, and validation through field operational tests respectively. The eSecurity Working Group [8] involving data protection and ITS stakeholders was also created. Work is continuing as the European Commission is currently carrying out a study to assess data protection in ITS [10] while keeping the new privacy regulation [11] in perspective.

This paper reports on the insight gained from these undertakings. It will first report on experience gained in the study of two applications, eCall and Pay-Per-Use insurance, and two R&D projects, SeVeCom and PRECIOSA. It will then describe barriers to ICT which are not necessarily specific to ITS and provide recommendations for

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 177–186, 2014.

ICT in general. These barriers are at the application level (conflict of interest, lack of consensus on protection policies), at the design level (agreement on the meaning of Privacy-by-Design, neglect of architecture impact, lack of practice) and at the implementation level (leaks created by ICT infrastructures, lack of flexibility).

2 Experience Gained from ITS

2.1 Applications

eCall is a European initiative intended to bring rapid assistance to motorists involved in a collision anywhere in the European Union. The development of solutions for eCall rapidly led to concern about the location tracking of vehicles. The Article 29 Working Group Party published a working document on eCall in 2006 [12] which recommended the possibility of switching off the eCall capability. This recommendation raised further issues that were discussed in a meeting organised by the technology subgroup of the Article 29 WG Party in 2009. During this meeting a second generation eCall product, developed in liaison with the French data protection authority (CNIL) was presented by PSA. This solution included privacy preservation features, e.g. blurring vehicle location data to avoid calculation of speed, removing physically collected data on a daily basis. The solution also coped with liability issues created by the proposed approach of providing a switch-off capability: what if a vehicle eCall capability is switched off by one person and then the vehicle is used by another person which is not aware that it is switched off. It was suggested that a systematic check be included along with a request that the driver maintain the eCall capability off. The eCall case was a wake-up call within the ITS community on the need to address location issues.

Pay-As-You-Drive Insurance is a type of automobile insurance whereby the costs of insurance are dependent upon time, distance, behaviour and location. Further to the Article 29 WG party document on eCall, the European Commission organised a workshop in privacy in ITS in 2007. During this workshop a person from the French protection agency (CNIL) presented the case of MAAF, a French insurance company which requested to deploy a Pay-As-You-Drive solution but was denied authorisation. The same year, a study was started at research level that led to the specification of a privacy friendly solution called PriPayd [13] based on an approach whereby location data was kept in the vehicle. Instead of having internet based systems collecting the data and calculating invoices, only minimum billing data are provided by the vehicle. The striking characteristics of the solution was that by focusing on the physical minimisation of data (i.e. data is not collected on the internet, but is kept in the vehicle), the architecture of an application was profoundly changed.

2.2 ITS Technology

Rather than focusing on the impact of privacy on ITS applications, SeVeCom and PRECIOSA focused on how Privacy-by-Design applied to ITS applications could impact on the underlying ICT technology.

SeVeCom was an FP6 project that ran from 2006-2009 [4, 14, 15]. It focused on security for communication systems involving cooperating vehicles (i.e. car-to-car

and car-to-infrastructure communication). It therefore focused on privacy leaks that can occur in this kind of communication.

SeVeCom made the following analysis: communication includes application data and protocol data. Application data need to be transmitted with different levels of security for integrity or confidentiality reasons. Protocol data might also need to be transmitted in a secure way since they can lead to privacy infringement. This is the case of the communication of the MAC[1] address in car-to-car communication. The car-to-car MAC address was initially devised as a fixed unique address. Similarly to fixed IP addresses, the MAC address could therefore be considered as personal data, because it could be used to track a vehicle.

SeVeCom contributed the following technology:

- a mechanism in the form of proof of concept for secure communication, with pseudonym change management to address the fixed MAC address problem,
- a contribution to flexibility in the form of an implementation structure to allow for easy integration in existing implementation protocols[2].

Future deployed ITS infrastructure could reuse the SeVeCom implementation to solve the fixed MAC address issue.

PRECIOSA was an FP7 project that ran from 2009-2010 [5]. It addressed the problem of protecting collected data in ITS applications. It therefore focused on measures for privacy leaks that can occur when collecting data, in particular in ICT deployment based on common platforms.

PRECIOSA work was heavily influenced by discussions that took place in the eSecurity Working Group [8] concerning data protection stakeholders as well as privacy enhancing technology stakeholders and the need to adopt a Privacy-by-Design approach[3]. Since the meaning of Privacy-by-Design lacked clarity, the concept was investigated by the project. PRECIOSA concluded that it involves three principles, minimisation, enforcement and transparency [17] which are defined as follows.

Minimisation is related to the collection limitation principle for privacy of the OECD guidelines [18]. Applied to Privacy-by-Design, it means that the collection of personal information should be kept to a strict minimum in the design of an application. Applied to current technology trends, it means that the design process should start with the default option that no identifiable data is collected. Moreover, whenever possible personal data should be replaced by equivalent minimised data. For instance, birthdate information can be replaced by a computing proof that a person is over eighteen. Minimisation leads to requirements on what shall not be collected, on where it is collected, and on the use of specific minimisation technology. This approach was applied in Pay-As-You-Drive insurance [13].

Enforcement is related to the security safeguards principle for privacy in the OECD guidelines, which states that personal data should be protected by reasonable security

[1] MAC stands for Medium Access Layer. The MAC address identifies a communication entity in a physical network.

[2] Based on the so-called *hooking architecture* [15].

[3] The term Privacy-by-Design was coined by Ann Cavoukian, the Ontario Data Protection Commissioner [16]. In January 25, 2012, the European Commission proposed a comprehensive reform of the EU's 1995 data protection rules to strengthen online privacy rights and boost Europe's digital economy. The reform integrates the concept of Privacy-by-Design [11].

safeguards against such risks as loss or unauthorised access, destruction, use, modification, or disclosure of data[4]. Applied to Privacy-by-Design, it means that an application should be designed to provide maximum protection of personal data during an operation. Applied to current technology trends, it means that the design process should start with the default option that all collected personal data should be protected by technical means. They should automatically ensure that data are accessed by the authorised parties (e.g. location data is only made available to a location based application) and that such data is automatically removed at the expiration of a retention period (e.g. at the end of the day). This approach is exemplified by Hippocratic databases [19]. PRECIOSA took this further by developing a data-centric approach for protecting personal data in a cooperative ITS environment [20]. Enforcement leads to requirements on what must be protected, on how it is protected (which leads to organization decisions), and the use of technology for protection (in order to prevent from leaks due to manual organization of protection).

Transparency is related to the openness, the individual participation, and the accountability principles of the OECD guidelines. Means should be readily available to establish the existence and nature of personal data aa well as the main purposes for their use. Furthermore, an individual should have the right to get information on data collected about him. Finally, a data controller should be accountable for complying with the measures required for privacy preservation. Applied to Privacy-by-Design, this means that applications should be designed and operated so that maximum transparency can be provided to stakeholders about the way privacy preservation is ensured. In particular, the design process should include specific verification procedures (e.g. open design, auditing). Applied to current technology trends, it means that the design process should start with the default option that mechanisms for verification during operation should be included. For instance mechanisms could be included to provide evidence that some location data have been removed. Transparency leads to requirements on what evidences have to be produced, on how these evidences are provided (which leads to architecture decisions), and on the use of technology for evidence provision.

PRECIOSA contributed the following technology:

- a proof of concept data-centric approach for protecting personal data in a cooperative ITS environment,
- an understanding of the meaning of Privacy-by-Design.

Future deployed ITS infrastructure could reuse the PRECIOSA implementation to ensure the right level of enforcement.

3 Barriers to ICT

While investigating the impact of privacy on ITS, barriers were identified. Many of the barriers are general in nature, and they also apply to ICT.

[4] The rationale for the enforcement principle is to prevent accidental or malicious leaking of personal data. The massive deployment of ITS applications for millions of vehicles implies that a single failure or accident could have a huge liability impact.

3.1 Application Level: Conflict of Interest

When applications values are based on the use and exploitation of user data, conflict of interest will occur. This is what is currently happening in today social networks applications, and in ITS when application stakeholders want to make use of collected location based data in order to provide value-added services. There is a risk that privacy regulation and Privacy-by-Design are considered to be an obstacle for deployment of these location-based services, leading to the weakest interpretation on how to apply Privacy-by-Design.

Solving a conflict of interest in a global manner necessitates consensus. This consensus must ensure that an application can be cost effectively developed and deployed, which is the priority of application stakeholders) while protecting personal data efficiently, which is the priority of privacy defence stakeholders. It is recommended to put in place a consensus process supported by policy makers. This was suggested by EDPS [23] who recommends the development of best available techniques through "comitology", i.e. a consensus process. This approach was applied in the case of pollution prevention techniques, through a process supported by the European Commission called the Sevilla Process [24].

3.2 Application Level: Lack of Consensus on Protection Policies

Protection policies require agreement (e.g. whether a data field should be encrypted for confidentiality?). Without such agreement, different policies can be applied, leading to situations where the level of protection reached is that of stakeholders applying the least protective policy. Consider for instance data retention policy where some stakeholders simply do not remove data. A process for agreement on policies must be available but it is currently ill-supported by current standardization processes, since the time frame for standardization is so long. A more agile and flexible consensus process is needed.

3.3 Design Level: Lack of Agreement on the Meaning of Privacy-by-Design

The term Privacy-by-Design has been used widely by policy makers, including in the new privacy regulation [11]. It is generally associated with Privacy Impact Assessments (PIAs), an instrument that has been the subject of much study [25]. But as highlighted in [21], there is a gap between the understanding of this concept by policy makers and by the ICT engineering community. A common technical understanding of Privacy-by-Design is needed as a result of standardization work. [26, 27] are examples of work contributing to this understanding. The contribution of PRECIOSA explained in section 2.2 is an attempt to shape this understanding [17]. The creation of a multidisciplinary working group working in this understanding would be needed.

3.4 Design Level: Neglect of Architecture Impact

The architectural dimension of Privacy-by-Design is currently not well highlighted. Yet almost all privacy preserving solutions devised today have a profound impact on architecture as shown in the Pay-Per-Use case [13], the road charging case [30] and

the smart meters case [31, 32]. Currently research on privacy puts more value on contributions related to crypto aspects, which overshadows the need to assess architecture impacts aspects. For instance Stanford University has a web page which lists Privacy Enhancing Technologies (PET) [33]. However, no equivalent can be found for architectures. We have also observed that the meaning of PET is often narrowed to security and crypto-based features for minimisation. We believe that a broader meaning should be used, i.e. a PET can be a mechanism for minimisation, for enforcement of policies or for transparency.

It is recommended to take a more global architectural view rather than a mechanism centric viewed as suggested by the term PET. Let us switch to *Privacy Enhancing Architectures* (PEARs)!

3.5 Design Level: Lack of Practice of Privacy-by-Design

There is currently little of practice of Privacy-by-Design. We need to create a wealth of architectures (PEARs) and of measures (PETs): Privacy enhancing technologies are not well spread. Minimisation technology is a recent development. Much research is still in progress and a wealth of new results can be expected in the near future. Furthermore, it is also expected that new threats will be discovered as applications are deployed, which will also lead to new measures. Enforcement and transparency measures are currently mostly managed through manual and organisational activities. Industry expertise is not commonly available. Little research work is available on enforcement for privacy, e.g. [19, 20] for run-time protection perimeters. Nevertheless, these efforts could leverage on well-established work on enforcement of access such as the Bell-La Padula model [28].

Another issue is that Privacy-by-Design has to be properly integrated in the development process of applications. This integration is not easy to specify because of the wide variety of engineering processes in use (e.g. automotive, railways, smart meters, etc..)

Finally, Privacy-by-Design is a topic that is not addressed in the standard education curriculum. When current students are employed in the next few years, they will have little understanding of what a Privacy-by-Design process is.

3.6 Implementation: Leaks Created by ICT Infrastructures

ICT infrastructures include technology components which use and possibly transmit system data. Such data are needed for the operation of the infrastructure. For instance a run-time platform will make use of operating data such as computing resource descriptions, or a communication stack will involve the transmission of protocol data. In current industry practices, such data can be easily monitored for conformance testing or for performance monitoring purposes. However, the monitoring capability itself creates problems. Monitoring the content of memory used by an application is obviously a problem if there is a possibility to derive personal data from it. Transmitting protocol data can also be a problem, for instance a fixed IP address is enough to identify a user. The design of recent car-to-car communication systems initially planned to

use fixed MAC[5] addresses. However, this data would be enough to track a vehicle. Even worse, simply tracking communication activity could be enough in many cases to track user activity. For example, a device that transmits data could be a proof that a user is at home.

A novel approach to the design of ICT infrastructure must be taken so that it provides suitable protection of system data and system activity. It must be protection oriented, i.e. system data must be protected against unauthorized access. Isolation features should be integrated to prevent access to system data or activity by unauthorised stakeholders.

3.7 Implementation: Lack of Flexibility of ICT Infrastructures

The deployment of ICT infrastructures currently involves heavy investments and therefore any need for unanticipated modification is difficult to take into account. In fact, it is in general impossible to modify part of an ICT infrastructure while it is operating, in particular when millions of entities are involved[6].

Evolving requirements for data protection necessitate two levels of flexibility of ICT infrastructures. First of all, *policies for protection* could evolve. For instance some data initially transmitted in the clear are now required to be transmitted confidentially. Or policies for pseudonym renewal need now to be changed in a communication protocol. The challenge is to provide support for defining, creating and changing such policies dynamically. Secondly, *privacy requirements* could change. This may be caused by the discovery of privacy leaks, or by changing societal perception. For instance the physical location of smart grid data initially collected in a remote centre could evolve and be kept at the level of a smart meter. The challenge is to provide support for defining, creating and changing architecture parameters such as the physical distribution of data. This in turn has an impact on the definition of interoperability and related standards. For instance interoperability standards for a smart meter could become obsolete as a result of an architecture change.

Currently ICT infrastructures are designed and deployed according to practices which prevent such levels of flexibility. Policies are often totally hardwired, i.e. they are meant to remain unchanged. But perhaps more worrying, current architecture patterns are defined statically and not meant to change, as this would imply modifying interfaces that are frozen and standardized. Supporting the modification of patterns therefore means modifying development and standardisation practices.

Addressing ICT infrastructure flexibility necessitates a long-term research plan to address the following neglected features: *policy as a service*, i.e. the infrastructure should provide support for the flexible deployment of new policies. This should involve a set of consistent technologies in terms of description (policy language), of generation and of deployment (reconfiguring the infrastructure accordingly);

---

[5] The fixed MAC address issue is currently taken into account in ITS standardization activities. See [22].

[6] This kind of barrier must be well anticipated. This is what happened for instance when France switched overnight from 8 digit to 10 digit phone numbers.

architecture as a service, i.e. the infrastructure should provide support for the flexible deployment of new architecture patterns. This should involve features for describing architecture changes, generating modified interoperability specifications and deploying reconfigured items; *agile interoperability*, i.e. new industry practices must be put in place to make sure that reconfigurations of architecture go in parallel with appropriate modification of interoperability specifications.

3.8 Addressing Barriers

The following table provides examples of measures that can be taken to address these described barriers. A feasibility assessment is also provided.

Type of Barrier	Barriers	Recommendation of measures to policy makers	Feasibility
Application level	Conflict of interest	Creation and support of a consensus process	Domain dependent. Could be a short term goal for some domains
	Lack of consensus on policies	Creation and support of a consensus process	
Design level	Lack of agreement on the meaning of privacy-by-design	Create a multidisciplinary working group to define an agreed engineering process.	Short term goal[7]
	Neglect of architecture impact	Switching focus from PETs to PEARs	Short term goal
	Lack of practice of privacy-by-design	Wealth of architectures (PEARS) and measures (PETS)	Short term goal
		Integration into application design processes	Long term goal
		Support in curriculum	Short term goal
Implementation level	Leaks created by ICT infrastructures	Protection oriented design of infrastructure bricks, based on e.g. isolation features	Long term goal
	Lack of flexibility of ICT infrastructures	Research on flexibility Changing standardization practices for interoperability	Long term goal

Acknowledgements. We acknowledge the support of the European Commission in the following FP6 and FP7 projects: SeVeCom, PRECIOSA, OVERSEE, PRESERVE.

References

1. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:207:FULL:EN:PDF
2. ITS Action Plan and Directive, http://ec.europa.eu/transport/its/road/action_plan/action_plan_en.htm

[7] Many contributions on PETs already include implicit contributions on PEARS.

3. eSafety initiative, http://ec.europa.eu/information_society/activities/esafety/index_en.htm
4. SeVeCom, http://www.sevecom.org/
5. PRECIOSA, http://www.preciosa-project.org/
6. EVITA, http://www.evita-project.org
7. OVERSEE, https://www.oversee-project.com/
8. PRESERVE, http://www.preserve-project.eu/
9. eSecurity Working Group, http://www.esafetysupport.org/en/esafety_activities/esafety_working_groups/esecurity.htm
10. EC Study: Assess the security and personal data protection aspects related to the handling of data in ITS applications and services and propose measures in full compliance with Community legislation, http://ec.europa.eu/transport/its/events/2012_06_12_data_protection_en.htm
11. New privacy regulation in Europe, http://ec.europa.eu/justice/newsroom/data-protection/news/120125_en.htm
12. Article 29 Working Group Party working document on data protection and privacy implications in eCall initiative (September 26, 2006), http://ec.europa.eu/information_society/activities/esafety/doc/esafety_forum/ecall/art29wp_ecall_en.pdf
13. Troncoso, C., Danezis, G., Kosta, E., Balasch, J., Preneel, B.: PriPAYD: Privacy-Friendly Pay-As-You-Drive Insurance. IEEE Transactions on Dependable and Secure Computing (to appear), https://www.cosic.esat.kuleuven.be/publications/article-2013.pdf
14. Papadimitratos, P., Buttyan, L., Holczer, T., Schoch, E., Freudiger, J., Raya, M., Ma, Z., Kargl, F., Kung, A., Hubaux, J.-P.: SeVeCom. Secure Vehicular Communications: Design and Architecture. IEEE Communications Magazine 46(11), 100–109 (2008), http://infoscience.epfl.ch/record/129969/files/sevecom1.pdf
15. Kargl, F., Papadimitratos, P., Buttyan, L., Müter, M., Wiedersheim, B., Schoch, E., Thong, T.-V., Calandriello, G., Held, A., Kung, A., Hubaux, J.-P.: SeVeCom. Secure Secure Vehicular Communications: Implementation, Performance, and Research Challenges. IEEE Communications Magazine 46(11), 110–118 (2008), http://infoscience.epfl.ch/record/129970/files/sevecom2.pdf
16. Privacy-by-Desig, http://www.ipc.on.ca/english/Privacy/Introduction-to-PbD/
17. Kung, A., Freytag, J., Kargl, F.: Privacy-by-design in ITS applications. In: 2nd IEEE International Workshop on Data Security and Privacy in wireless Networks, Lucca, Italy (June 20, 2011)
18. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, http://oecdprivacy.org
19. Agrawal, R., Kiernan, J., Srikant, R., Xu, Y.: Hippocratic Databases. In: 28th International Conference on Very Large Data Bases, Hong Kong (August 2002)
20. Mechanisms for V2X Privacy. Deliverable D10. Preciosa FP7 Project (March 2010), http://www.preciosa-project.org/
21. Kung, A.: From PIAs to Engineering Practices. Computer Privacy and Data Protection 2012 (2012), http://www.cpdpconferences.org/I-Q/Resources/KUNG_120127.pdf

22. ETSI ITS WG5, http://docbox.etsi.org/workshop/2011/201102_ITSWORKSHOP/06_INSIDEARCHITECTURE/TC_ITS_WG5_CADZOW_StandardizationActivities.pdf
23. Opinion of the European Data Protection Supervision on an Action Plan for the Deployment of Intelligent Transport Systems in Europe. Official Journal of the European Union (February 25, 2010), http://www.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2009/09-07-22_Intelligent_Transport_Systems_EN.pdf
24. Schoenberger, H.: European Commission. Integrated Pollution Prevention and Control in Large Industrial Installations on the Basis of Best Available Techniques – The Sevilla Process. Journal of Cleaner Production 17, 1526–1529 (2009)
25. Wright, D., de Hert, P. (eds.): Privacy Impact Assessment. Series: Law, Governance and Technology Series, vol. 6. Springer (2012)
26. Spiekermann, S., Cranor, L.: Privacy Engineering. IEEE Transactions on Software Engineering 35(1), 67–82 (2009)
27. Gürses, S.F., Troncoso, C., Diaz, C.: Engineering Privacy-by-Design. Computers, Privacy & Data Protection (2011)
28. Access control based on Bell-La Padula model, http://en.wikipedia.org/wiki/Bell%E2%80%93LaPadula_model
29. Guidelines for Privacy Aware Cooperative Application. PRECIOSA Project Deliverable 11 (November 2010), http://www.preciosa-project.org/
30. Balasch, J., Rial, A., Troncoso, C., Geuens, C., Preneel, B., Verbauwhede, I.: PrETP: Privacy-Preserving Electronic Toll Pricing (extended version). In: 19th USENIX Security Symposium
31. Kursawe, K., Danezis, G., Kohlweiss, M.: Privacy-Friendly Aggregation for the Smart-Grid. In: Fischer-Hübner, S., Hopper, N. (eds.) PETS 2011. LNCS, vol. 6794, pp. 175–191. Springer, Heidelberg (2011)
32. Acs, G., Castelluccia, C.: I have a DREAM (DiffeRentially privatE smArt Metering). In: The 13th Information Hiding Conference (IH) (2011)
33. Stanford Center for Internet and Society PET Wiki, http://cyberlaw.stanford.edu/wiki/index.php/PET

A Method for Analysing Traceability between Privacy Policies and Privacy Controls of Online Social Networks

Pauline Anthonysamy, Phil Greenwood, and Awais Rashid

School of Computing and Communications, InfoLab 21,
Lancaster University, Lancaster LA1 4WA, UK
{anthonys,greenwop,awais}@comp.lancs.ac.uk

Abstract. Privacy management in online social networks (OSNs) is a major concern. However, the complexity of privacy policies and the plethora of privacy controls make it very difficult to assess whether the controls adequately implement the intended policies. This paper proposes a method to assess the degree of traceability between privacy policies and privacy controls in OSNs. The resulting analysis enables one to pinpoint key privacy management gaps that must be plugged. The method can be utilised by privacy watchdogs, user rights groups as well as OSNs themselves to assess the effectiveness of privacy measures.

Keywords: Social Networks, Traceability, Privacy, Policy, Controls.

1 Introduction

Privacy concerns stemming from the vast amounts of personal data collected by online social networks (OSNs) are soaring to unprecedented levels. A 2012 survey [1] revealed that 66% of online adults use OSNs – with recent statistics showing that Facebook alone now has more than 900 million users [2]. As a result of this massive growth, OSNs are facing increasing scrutiny from privacy advocates and regulators around the world. For instance, the Electronic Privacy Information Center (EPIC) [3] stated that it will likely file a complaint to the US Federal Trade Commission [4] over Google's new 'Social Search' which blends photos, comments and news posted on Google+ into its search results. Similarly, the EU's Data Directive [5] has recently mandated that websites (including OSNs) must pro-actively obtain consent from users before they store any information, i.e. cookies, on a web-connected device such as a computer or a mobile phone. Previous legislation, by comparison, had only mandated that websites contain a link to information on how users can opt-out of such practices.

OSN providers have attempted to address some of these privacy concerns. Google, for example, recently introduced a wide range of changes to its privacy policy [6] and Facebook revamped its privacy controls in 2011 to streamline how members manage their personal information [7]. This included 61 privacy settings on 7 different configuration pages [8]. Despite these efforts, a number of

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 187–202, 2014.

studies indicate that OSN users generally know little about how OSNs function in regards to transfer (or use) of their personal information [9,10]. They also find that privacy controls are inherently difficult to understand and configure [11]. Users' trust is further impeded by the privacy violations that follow these recurring changes introduced by OSNs. For instance, a lawsuit was filed against Facebook in which it was accused of violating its privacy policy by purportedly tracking web usage even after a user had logged-out of the social network [12].

Users' trust can be enhanced if it can be demonstrated that OSNs have taken effective measures to protect their privacy. A critical step towards achieving this is to confirm that the actual functionality provided by the OSN is *consistent* with and *reflective* of its privacy policy. This is important both for privacy watchdogs and user rights groups as well as OSNs themselves for the purposes of compliance checking. This paper aims to support this need by proposing a method to asses the degree of traceability between privacy policies and privacy controls of OSNs. The method is rooted in our earlier extensive empirical study of the relationships between privacy policies and privacy controls in OSNs [13]. The novel contributions of this paper are as follows:

1. A taxonomy to classify privacy policy statements in OSNs.
2. A method that provides a systematic means of studying the traceability[1] between privacy policies and controls in OSNs, hence establishing the *degree of traceability* between the two. We define the **degree of traceability** as the level of certainty that we can have about the existence of an externally observable relationship measured using a qualitative 3-point scale.
3. An analysis, based on the *degree of traceability*, that indicates where corrective measures ought to be targeted to improve privacy management.

2 Contributions beyond State-of-the-Art

Existing research related to this area falls into two main categories: (i) Requirements Engineering, Policy and Traceability; and (ii) Social Networks and Privacy. We briefly summarise key works in each of these categories and the novel contributions of our work.

2.1 Requirements Engineering, Policy and Traceability

There has been a significant amount of research to ensure that software requirements comply with governing legal texts and privacy policies. Initial efforts in this involved extracting requirements from privacy policies. Young et al. [14,15] introduced a systematic method for obtaining requirements from privacy policies of healthcare organisations by extracting commitments, privileges, and rights.

[1] Traceability expresses relationships that exist between requirements and other entities of software. In our work we define traceability as the *externally observable relationships* between privacy policies and the run-time functionality of privacy controls.

This method was aimed at aiding requirements engineers to analyse the natural language text in privacy policies rather than using an intermediate representation i.e. goals [16]. Similarly, Breaux et al. [17] developed a method to extract rights and obligations from legal documents – mainly the HIPAA Privacy Rule. While these works focus on deriving requirements from privacy policies and legal documents (of healthcare domains), the focus of our method is on assessing the traceability from statements in privacy policies to the runtime realisation of privacy controls in OSNs.

Much research has been performed in the area of federated digital identity management which focuses on protecting the identities of individuals at an inter-organisational level. In [18], Scquicciarini et al. presented an approach that enables users to trace their personal information across a federated organisation and verify whether it has been managed according to their privacy preferences. The authors present a policy harmonisation mechanism that determines whether the transfer of personal information from one entity to another violates the privacy policy stated by the originating entity. While the overall motivation of this work is similar to ours, it differs in objective and the execution domain. The objective of our work is to determine whether the runtime functionality provided by privacy controls is a true reflection that can be traced back of the statements in the privacy policies in the highly volatile domain of SNS. Researchers have also explored traceability issues between legal documents and requirements. Cleland-Huang et al. [19] use machine learning approaches to automatically generate traceability links between regulatory codes and product level requirements (evaluated against a subset of the HIPAA regulatory document) while other works examine the generation of traces between documentation and code, e.g., [20].

2.2 Social Networks and Privacy

Bonneau et al. [8] conducted a comprehensive study of privacy in 45 OSNs using criteria such as the diversity of data collected by the sites, the types of privacy controls, promotional methods etc. Their study included a general analysis of privacy policies in which attributes like the accessibility of privacy policies, their length, data claims and so on were examined. This work closely relates to ours in the sense that it was performed through passive data collection with conclusions drawn based on the collected data – but differs in terms of the objective which in our case is to analyse privacy from a software traceability perspective rather than an economic one.

There have also been a wide range of user studies [9,11] which highlight users' disposition towards privacy on social networks. Luders et al. [9] performed a user study on the experience and attitudes of the general public, primarily Norwegian users, with regard to personal and consumer protection in social media (focusing on Facebook). This survey showed that users' knowledge of how social media functions in regards to use, disclosure and transfer of their personal information is largely inadequate. The authors also reported that the surveyed users found the privacy controls provided on social networks to be difficult to configure and comprehend. In contrast, [11] investigated users' sharing intentions and actual

privacy controls in search of infringement. The study found that every one of the participants had at least one sharing violation based on his/her stated sharing intentions. Both these studies substantiate our motivation that despite continuous efforts from OSN providers, users are still dissatisfied and are concerned about the consequences of sharing their personal information.

3 Privacy Taxonomy

At the heart of our traceability analysis method is a taxonomy that provides a common classification framework in order to overcome the ambiguity and granularity mismatches that can be present in OSN's privacy policies and controls. Furthermore, our taxonomy also offers a common framework that provides an abstraction layer over the specific functions offered by different OSNs, hence enabling applicability across different OSNs (even providing a means to compare the effectiveness of privacy measures across OSNs).

The objective of our method is to determine whether or not traceability can be established between OSNs' privacy policies and their privacy controls. Establishing such a mapping is a non-trivial task. Multiple statements/paragraphs from a site's privacy policy could potentially be realised as one or more controls and vice-versa. Furthermore, statements relating to a given aspect of privacy could be scattered in different parts of the document. As described in [13], to overcome the ambiguities and granularity mismatches between privacy policies and privacy controls, we used a combination of thematic and content analysis techniques to analyse the privacy policies of sixteen OSNs and derived a privacy taxonomy from this analysis. Thematic analysis refers to the process of categorising segments of qualitative data into meaningful themes and content analysis is a rigorous form of thematic analysis that provides an effective mechanism for analysing and comparing the segmented texts [21].

To develop the taxonomy, we applied thematic and content analysis to the *actions* performed by an OSN. Actions depict the manipulations that are performed with regards to the data provided by an OSN user. Actions can be expressed by any of the verbs or verb phrases in natural language. Verb phrases refer to a combination of a verb and an article e.g. "is required", "will share". We began our analysis cycle by coding Facebook's privacy policy using the action verbs found in its section and sub-section headings. Consider a sub-section heading from Facebook's privacy policy: "How we use your information". The action verb in this heading is '**we use**' and respectively we created a category label called 'Information Use'. We also drew some inspiration for action categories from Anton et al.'s [22] privacy taxonomy such as 'Information Collection'.

Figure 1 illustrates our complete refined set of action categories according to the common characteristics that emerged when examining the privacy policies of OSNs in our survey reported in [13]. High-order categories give an overview of the actions described in a privacy policy, while lower-order categories allow for finer distinctions to be made within and among specific categories of actions.

Privacy Categories	First Level and Second Level Sub-Categories
Information Collection Statements describing what information are collected from the users of a social network. These statements include those that describe direct collection activities – collecting information by directly requesting it from a user, e.g. mandatory registration information needed for becoming a member of a site and other optional information such as hobbies, etc. – and indirect collection activities - collecting information passively, e.g. IP address, user activities and browsing patterns.	*Direct Collection* Personal / Registration Information Optional Information *Indirect Collection* Information gathered passively
Information Use Statements describing how information that is collected from the users of a social network is utilised by an OSN provider.	Internal Use Communication Advertising Personalisation
Information Sharing Statements describing disclosure of information by the users of a social network and the OSN provider. These statements include those that describe why and to whom users/OSN providers share information and how this information is safeguarded.	*Sharing by a member* With other users With 3rd parties *Sharing by the OSN provider* With 3rd parties With law enforcements
Information Management Statements describing the ability of an OSN user to edit or remove his/her information from the site at any time.	Review/Change Personal Information Information Removal
Information Monitoring Statements describing what an OSN provider or third-party organisations may track when a user interacts with the site. These statements include those that describe the type of monitoring technologies that are employed, its purposes and the activities/data that are tracked.	By OSN provider By 3rd parties
Information Protection Statements describing special mechanisms (if any) that are employed to protect minors.	Protection for Minors Age Restrictions

Fig. 1. Privacy Policy Categories (Taxonomy)

4 Traceability Analysis Method

Traceability analysis is an approach for assessing the degree of traceability between artefacts [19], in our case between the privacy policy statements and privacy controls of OSNs. The degree of traceability between these artefacts reflects to what extent the functionality offered by the privacy controls is consistent with and reflective of the privacy policy. This section introduces our method and its constituent steps (shown in Fig. 2) with a running example from our study. Our method is composed of two activity streams that can be performed in parallel: a policy stream and a control stream.

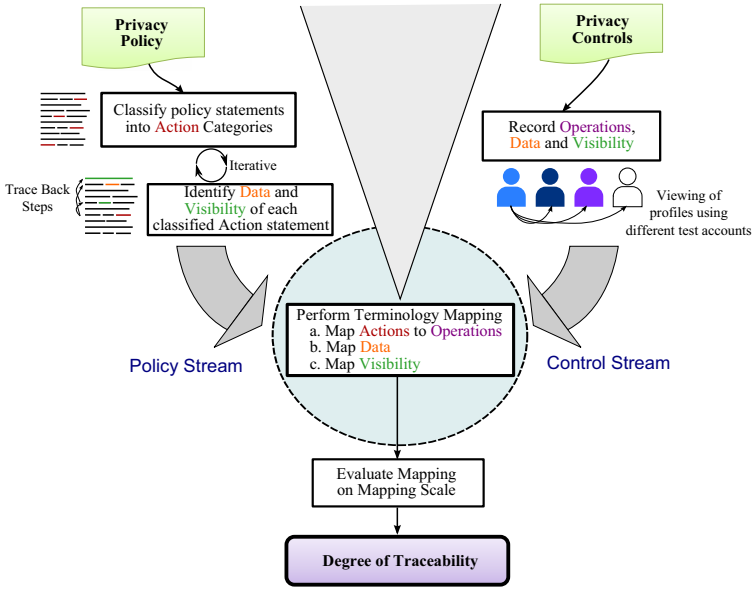


Fig. 2. Traceability Analysis Method

4.1 Policy Stream

4.1.1 Classify Policy Statements into Action Categories

The first step in the policy stream of our method involves using the taxonomy in Section 3 to decompose and classify the statements from an OSN's privacy policy within the action themes of the taxonomy. In this activity the policy statements are decomposed into individual statements and the content of each statement is examined to identify the "actions" (verbs or verb phrases) performed by an OSN. Consider Google's privacy policy statement in Fig. 3. The action verb in this statement is '**use**'. Subsequently, this statement is classified under the 'Information Use' category of our taxonomy. This statement can be further classified within the 'Personalisation' sub-category as it specifies the purpose of the use i.e. verb phrase '**to personalize**'. If a statement does not have an action verb, i.e. statements which merely provide contextual information or examples which elaborate on a preceding statement, it should be classified within the same category as the preceding statement to which the example or contextual information relates.

4.1.2 Identify Data and Visibility of Each Classified Action Statement

Classifying policy statements into homogeneous taxonomy categories not only eases the task of semantically analysing them but also eases another activity in our method: identification of data and visibility. At the core of OSN privacy is the *data* it maintains and the *visibility* of these data. These dimensions are common to both privacy policies and privacy controls and are used as common reference points to derive mappings in our method. To elaborate further, privacy policies specify the data or the types of data being collected by the site, as well as the uses and recipients (visibility) of that data. The fundamental principal behind privacy controls can be described as a mapping between predefined categories of data (e.g. name, contact information, interests/hobbies, etc.) and categories of people (e.g. public, private, friends, etc.) who are allowed to view that data i.e. visibility. For example, a specific piece of data, a user's last name, can be controlled to be visible only to the user's friends.

While classifying the policy statements (Section 4.1.1), each classified statement needs to be examined to identify the specific elements of interest within it – the data and visibility which pertains to an action – to prepare the statement for mapping against the corresponding set of operations found in an OSN (Section 4.2). Schneier's data classifications [23] were used to identify all the data that is associated with both privacy policies and privacy controls.

Consider the same privacy statement from Google+'s privacy policy in Fig. 3 that expresses that the site uses +1's and other profile information (collected from a user) to provide tailored content and ads on third-party websites. The data is '+1's' and 'other profile information' and, the documented visibility is 'non-Google websites' i.e. third parties such as applications accordingly (highlighted).

4.2 Control Stream: Record Operations, Data and Visibility

This activity of our method, which can be performed independently of the previous stream and by different teams or team members, involves recording privacy controls (operations) that are available to a user after signing-up to an OSN. All information (including optional information) that is requested by an OSN during sign-up is also recorded. This enables the data collector to determine the default visibility of optional information if a user chooses to include it.

> *You can choose whether Google may **use** your **+1's** and **other profile information** to **personalize** your content and ads on non-Google websites, including applications or other clients.*

Fig. 3. Snippet from Google+'s Privacy Policy

To determine the complete visibility of data, it is necessary to log-in to each of the OSNs being examined using test accounts (four test accounts are needed:

two for *adults* and two for *minors*[2]) and search for the other accounts using the internal search options (if available) of the OSN. All profile information (of the other accounts) that is publicly visible can then be recorded. All of the privacy controls, including the textual labels (or terminologies) used to describe each privacy control, that are available to a user are observed and recorded along with their *default values*. The other accounts (i.e. an adult and a minor) are then added as friends to capture further differences in terms of information visibility. These steps are repeated by logging into a minor's account. Finally, the data that is visible to non-members and external search engines is then recorded. It should be noted that the default privacy settings which were prescribed by the OSN provider during profile creation must be preserved to ensure consistency of the data collection and recording activity.

4.3 Perform Mapping

In this step of our method, the output of the 'Policy' and 'Control' streams is brought together to perform the mapping. The classified policy statements are mapped to the corresponding privacy controls on the basis of the common reference points to identify whether or not traceability relationships exist between the two. The mapping procedure consists of three key steps:

1. Map Actions:- For each ***action*** statement within a privacy category a corresponding ***operation*** is identified based on the use of matching terminology (terms and words that have the same meaning in describing the actions and operations)[3].
2. Map Data:- Next, verify that the ***data*** item(s) manipulated by the operation match those mentioned in the statement.
3. Map Visibility:- Finally, compare the default ***visibility*** of this data, as indicated by this operation, with that specified in the policy statement.

We illustrate the mapping of Google's policy statement shown in the previous steps of our method. Google's description of their provision of personalisation and ads is shown in Fig. 3. The subset of privacy controls whose operations relate to 'Personalisation' are filtered by matching their terminology against the corresponding text segment. Here, the terms '+1's' and 'personalize' from the privacy policy statement map to the operation described as '+1 personalization', i.e. the two have the same semantic definition.

Next, the data items that pertain to the statements are mapped to the matched operation. This can be a non-trivial process due to the ambiguity in either the text describing the *operations* or the *actions*. In the case of *operations*

[2] There exists a distinct set of regulatory codes that governs the access of minors on OSNs.

[3] In cases where action statements involve negation, such as "we will not collect" or "we will not use", we verified that this negation was reflected by the OSN controls. For example, if a social network states that it will not collect data item X then, we verified that this was complied.

this is caused by instances where the matched operation does not explicitly manipulate the data items specified in the policy statements. In such cases data items based on the illustrations or examples provided by the OSN when describing this operation should be extrapolated if possible. Example in Fig. 4 illustrates that the **name** (data item) of a member is used to display the relevant personalised content. In the case of *actions* this ambiguity is caused by cases where the privacy policy does not explicitly refer to any specific data items, i.e. uses catch-all/generalised terms (for example '**other profile information**'). In these cases, a trace back step is required to see whether these terms were defined anywhere in the privacy policy. Finally, the third sub-step of this activity is the identification and mapping of the visibility. The default visibility value of the matched operation, i.e. non-Google websites (users are opted-in by default for personalisation), as shown in Fig. 4 is mapped to that specified in the policy statement i.e. non-Google websites. We recognize that a small number of statements cannot be verified (mapped) and ultimately rely on users' trust, e.g. "we will not sell your information to 3rd parties". In this work, we choose to focus on the majority of statements that can be subjected to verification.

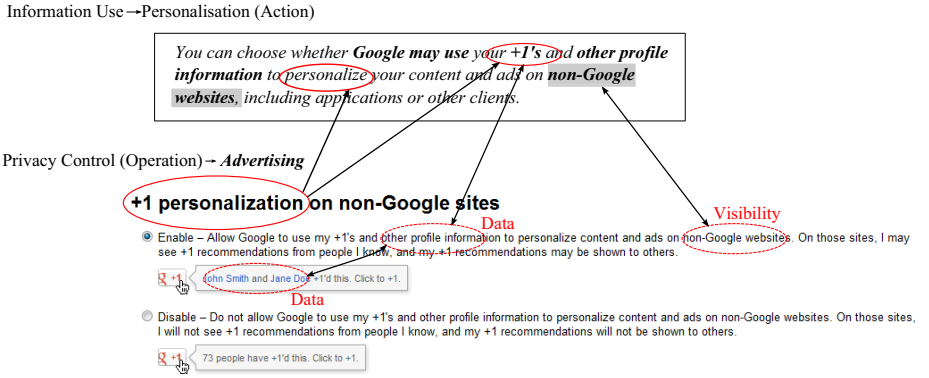


Fig. 4. Terminology matching between action statements and operations

4.4 Evaluate Mapping

Finally, to evaluate the degree of traceability, each mapping outcome (statement ⇔ operation) is assessed on the following linguistic variables[4] [24]:

> ***Complete***: To be considered complete, the mapping between the privacy policies and controls should meet two fundamental properties: (1) Consistent - refers to lack of contradiction between statements in privacy policies and

[4] Variables whose values are not numbers but words or sentences in a natural or artificial language. The motivation for the use of words or sentences rather than numbers is that linguistic characterisations are, in general, less specific than numerical ones [24].

controls. This is evaluated based on consistent use of terminology, where terms and words with the same semantic meaning are used to describe actions and operations. The common reference points are examined to ensure that data items referred to by actions and operations are identical and default visibility matches that of the policy. (2) Unambiguous - the information in a privacy policy can be understood and interpreted in only one way (i.e. explicit specification of actions, data and visibility - no use of generalised terms like "personal information") and the actions easily related to one or more operations of an OSN.

Partial: A mapping is marked as partial if a clear relationship could not be established between the information given in a privacy policy and the privacy control(s) of a site. There are two properties to evaluate this: (1) Imprecise - refers to missing/unspecified and indirect information. Missing/unspecified implies data that is requested by an OSN when interacting with the site but is not included in its privacy policy. Indirect information is information that is specified subtly, i.e. which is there for the purposes of compliance with regulations but potential not intended to be noticed by a user (e.g. opt-in/out links buried within the policy text). (2) Ambiguous - terms or phrases in policy statements that are vague or can be interpreted in more than one manner e.g., use of generalised terms like "personal information" without having an explicit definition for what constitutes such terms. Terms are also evaluated as ambiguous when a trace-back step fails to reveal more specific information.

Broken: A mapping is broken based on one property: Disjoint - which refers to instances where statements (actions) specified in a privacy policy could not be matched to any operations. This implies that a traceability relationship does not exist since the actions are not operationalised into privacy controls.

With this, the degree of traceability of our running example in Fig. 3 and 4 is evaluated as 'Partial'. This is due to imprecise definition of the data items i.e. the use of the catch-all term 'other profile information' without defining what this constitutes. In this paper we intentionally focus on Partial and Broken mappings (in our examples) as these are of most interest to users. The reader is referred to [13] for a more complete analysis including examples of when Complete mappings were observed.

5 Evaluation

5.1 Data Source

The data used in our evaluation is the privacy policies and privacy controls of Google+[5], Meet Me[6] and Zorpia[7]. We focused on these sites as they were not

[5] `https://plus.google.com/`
[6] `www.meetme.com/`
[7] `www.zorpia.com/`

part of the original sixteen OSNs which were analysed in our original study. In addition, these three sites met the criteria defined in our previous survey [13]; the sites have a large number of active users from around the globe [25]; their privacy policies were available online; the sites provided a variety of privacy controls to their users.

Our initial step was to take snapshots of each site's privacy policy (in June 2012). Each snapshot was taken as it would be shown to a non-member and was dated accordingly. Next, we applied our method to each site by: creating four test accounts on each site (two for *adults* and two for *minors*), recording all information requested by the site during sign-up, and maintaining consistency by providing standardised information such as name, birth date, email, etc. across the three sites. Complete profile information can be found at `http://www.comp.lancs.ac.uk/~anthonys/dataset.html`.

5.2 Results

Of the policy documents analysed, Google+ had the most elaborate policy descriptions which were spread across several document types – privacy policy, definition of key terms and advertising principles. Whereas, Meet Me and Zorpia had only one privacy document.

Within the analysed policy documents across the three surveyed sites, a total 150 statements were classified based on our taxonomy. Statements that were not classified included definition of key terms, warning messages and those that referred to documents outside of our analysis scope. Of the 150 statements, 70 statements were then mapped according to our mapping procedure. The fact that more statements were classified than were mapped is primarily because our classification procedure groups successive statements together based on the context surrounding those statements; while our mapping procedure specifically requires that a statement has a clear action or action-verb to trace to a corresponding control. Classifying these extra statements were useful in concretisation of the primary statement which is a candidate for the mapping. The extra statements which were classified but not mapped include statements within paragraphs that provide supporting examples and generic statements such as inviting parents to contact the site about any concerns they may have. The bar chart in Fig. 5 shows the number of statements that were classified and mapped across the three sites.

Figure 5 illustrates the degree of traceability of the mapped statements. 25% of Zorpia's statements were evaluated as having Complete traceability according to our mapping scale. This is followed by 22.8% for Google+ and 14.8% for Meet Me. 42.9% of the mapped statements from Google's privacy policy were evaluated as Partial and 34.3% as Broken based on our traceability analysis method. Although Google had the most elaborate privacy policy, we found its policy statements to be too ambiguous for a user to effectively control how their information is shared or used. In terms of privacy controls, we found that the vast majority of its controls were not very intuitive to a user and the descriptions of how one can use these controls were buried in their Help document which consists of at least 40 different pages.

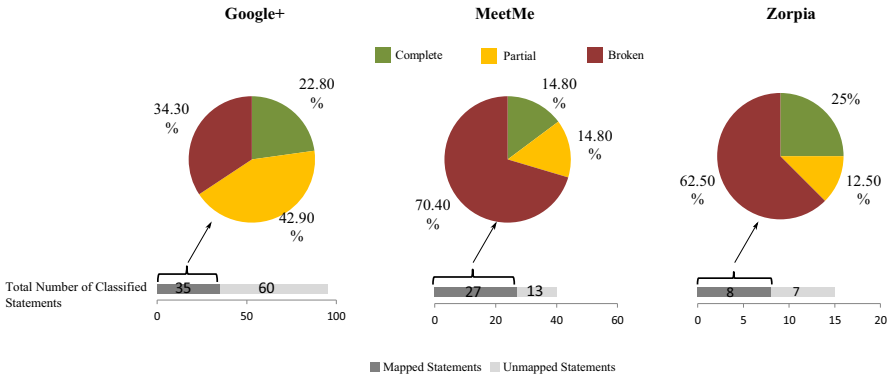


Fig. 5. Degree of Traceability across G+, MeetMe and Zorpia

Meet Me and Zorpia had a higher composition of statements that were Broken (70.4% and 62.5% respectively). Our analysis shows that these two sites nominally comply or perhaps are even in violation of regulatory requirements [5,4]. For instance, Zorpia's privacy policy does not have any statements that describe its data collection practices. As for the controls, the two sites had a very abstract control mechanism. For instance, users' profile visibility was grouped in a coarse-grained manner and there were no opt-in/out options for sharing one's information with third-parties. The statements that were marked Partial composed of 14.8% in Meet Me's policy and 12.5% in Zorpia's.

5.3 Meta-analysis

We now highlight the key insights gained through the application of our traceability analysis method. In particular we discuss the general deficiencies found in establishing a mapping between policy documents and controls based on the degree of traceability. This is illustrated using some of the policy statements as examples in Fig. 6. We also suggest potential remedies and corrective measures that ought to be targeted to improve these deficiencies.

- **Information Asymmetry:** Terminologies used in describing data categories are inconsistent and imprecise. This is exemplified in row two of Fig. 6 with the use of catch-all/generalised terms. This makes it difficult for users to understand the way in which their data is handled. This general class of deficiency can be remedied by using precise and consistent terms across policies and controls, potentially including direct hyper-links from controls to the corresponding statements in policies.

- **Default Opt-Ins:** Although privacy policies claim that they are in favour of privacy, their operationalisation of privacy controls is not privacy-friendly. It is often the case that OSNs do not explicitly require users to opt-in to

	Example of Statements	Degree of Traceability	Deficiencies	Potential Remedies
			Google +	
1	***We may combine personal information from one service with information, including personal information, from other Google services – for example to make it easier to share things with people you know.***	Broken	Lack of precision in describing what constitutes "personal information" and across which services and how.	Explicit specification, for instance, an exhaustive list, of data items that constitute personal information and its usages. Users should be given a choice on whether to opt-in/out of the unification of their data across Google's services.
2	***We may use information that you tell us about yourself to personalize ads on search.***	Partial	Use of catch-all/generalised terminologies in describing the data items that will be used for personalisation purposes.	Clear definitions or familiar metaphors should be used in describing the data items that will used in tailored ads.
			Meet Me	
3	***MeetMe automatically receives and records information on our server logs from your browser, including your IP address, MeetMe cookie information (including without limitation cookies, flash cookies, and beacons), and the pages you request.***	Broken	Passive data collection happens automatically and users do not have a choice regarding which individual data items are collected and processed.	Pro-actively asking users consent on passive data collection. Explicit specification – by using familiar terms - of the types of data that will be collected.
4	***Do not post information in public areas that you want to keep private.***	Partial	Opt-in/out choices are presented at an abstract level, not at the granularity of individual choices made when a user chooses to share his/her information.	Fine-grained controls i.e. per data item control.
			Zorpia	
5	***We may, however, share aggregate demographic statistics about members such as interests or occupations, but advertisers and third parties can never access your individual personal information.***	Broken	Ownership and sharing to third-parties is controlled by OSN providers.	Pro-actively asking users for consent on sharing of their data with third-parties. Provide precise definitions on what is meant by aggregation and what constitutes aggregated information.
6	***Zorpia is not responsible for the privacy practices of other websites/applications that may be linked to Zorpia pages. You are solely responsible for reviewing third party privacy policy terms.***	Partial	The burden is placed on the user for finding/reviewing third-party privacy policies.	The individual 3rd party policies could be reconciled into a consistent representation that illustrates the different contexts (i.e. ads/apps) that draws on the privacy practices set by these 3rd parties.

Fig. 6. Examples of Deficiencies Found and its Potential Remedies

their new features/services, instead using default opt-ins. This is exemplified in rows one, three and five of Fig. 6 with the use of default/forced opt-ins. This kind of deficiency is made worse by the fact that users do not tend to be aware of such practices in their general interaction with the sites. Recent EU legislation [26] has attempted to address this deficiency by requiring explicit consent for the use of cookies, but we argue that this could be expanded much further to cover all kinds of user data.

– **Integration with Third-Parties:** Integration with third-parties represents an important revenue stream to OSNs, but privacy implications of this tend to be poorly defined. This is exemplified in rows two and six of Fig. 6 with poorly defined third-party information sharing. One potential remedy to this deficiency is to reconcile the individual 3rd party policies into a consistent representation that illustrates the different contexts of sharing, e.g. ads or apps, that draws on the privacy practices set by the contributing 3rd parties.

6 Conclusion and Future Work

The method presented in this paper aims to address a key gap - that of evaluating the traceability between privacy policies and privacy controls in OSNs and pinpointing where remedial measures may be targeted to improve privacy management. This is essential not only from a regulatory compliance perspective but also for discharging OSNs' obligations towards their members and addressing users' concerns about privacy of their personal data entrusted to OSNs. As such our method provides a stepping stone towards more systematic and objective means to evaluate the effectiveness of privacy management mechanisms – that is, not just considering privacy policies and privacy controls in isolation from a compliance or functionality perspective respectively but taking a holistic view that studies effectiveness of these measures altogether. Our future work will focus on two key, inter-linked, directions. Firstly, we aim to develop a software (CASE) tool to support privacy auditors to establish the degree of traceability between privacy policies and privacy controls. Secondly, we aim to develop a framework that will (semi-)automate the management and updating of traceability links between policies and controls so that the impact of changes to the one can easily be understood in terms of changes or updates required to the other. Such changes/updates will be automated to as high an extent as possible, hence enabling more effective maintenance of traceability relationships and consequently a high degree of traceability.

Acknowledgements. "Social Media, Social Good: Ultra-Large Scale Public Engagement Systems to Challenge Anti-Social Behaviour" This research is being funded by a Lancaster University 40th Anniversary Research Studentship, EPSRC Grant EP/I016546/1 and EP/I016546/1. We would also like to thank Barry Porter for his valuable feedback on this paper.

References

1. http://www.pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx (last accessed December 4, 2012): Pew internet: Social networking
2. http://www.businessinsider.com/facebook-now-has-900-million-monthly-users-2012-4, Facebook now has 901 million monthly users (2012) (last accessed: December 4, 2012)
3. http://epic.org, Electronic privacy information center (last accessed: December 4, 2012)
4. http://www.ftc.gov/reports/privacy3/, Privacy online: A report to congress (1998) (last accessed: December 12, 20102)
5. http://eur-lex.europa.eu/LexUriServ/LexUriServ Eu data directive 95/46/ec (2011) (last accessed: December 4, 2012):
6. http://www.bbc.co.uk/news/technology-17205754 Google privacy changes 'breach eu law, (last accessed: December 4, 2012)
7. http://blog.facebook.com/ Facebook - new privacy controls (2011) (last accessed: December 12, 2012)
8. Bonneau, J., Preibusch, S.: The privacy jungle: on the market for data protection in social networks. In: Economics of Information Security and Privacy. Springer, US (2010)
9. Brandtzaeg, P.B., Lüders, M.: Privacy 2.0: Personal and consumer protection in new media reality. Tech. Rep. SINTEF A12979 (November 2009)
10. Singh, R., Sumeeth, M., Miller, J.: A user-centric evaluation of the readability of privacy policies in popular web sites. Information Systems Frontiers (2010)
11. Majeski, M., Johnson, M., Bellovin, S.M.: The failure of online social network privacy settings. Technical Report CUCS-010-11 (February 2011)
12. http://www.techspot.com/news/48654-facebook-sued-for-15-billion-over-alleged-privacy//-violations.html
13. Anthonysamy, P., Greenwood, P., Rashid, A.: Can privacy policies be traced to privacy controls on social networking sites?: A qualitative study. IEEE Computer (2012) (accepted and to appear)
14. Young, J.: Commitment analysis to operationalize software requirements from privacy policies. Requirements Engineering (2011)
15. Young, J., Anton, A.: A method for identifying software requirements based on policy commitments. In: 2010 18th IEEE International Requirements Engineering Conference (RE), September 27-October 1 (2010)
16. Antón, A.I., Earp, J.B., Carter, R.A.: Precluding incongruous behavior by aligning software requirements with security and privacy policies. Information & Software Technology (2003)
17. Breaux, T., Antón, A.: Analyzing regulatory rules for privacy and security requirements. IEEE Trans. Softw. Eng. (January 2008)
18. Squicciarini, A.C., Bhargav-Spantzel, A., Czeskis, A., Bertino, E.: Traceable and automatic compliance of privacy policies in federated digital identity management. In: Danezis, G., Golle, P. (eds.) PET 2006. LNCS, vol. 4258, pp. 78–98. Springer, Heidelberg (2006)
19. Cleland-Huang, J., Czauderna, A., Gibiec, M., Emenecker, J.: A machine learning approach for tracing regulatory codes to product specific requirements. In: ICSE (2010)

20. Antoniol, G., Canfora, G., de Lucia, A., Casazza, G.: Information retrieval models for recovering traceability links between code and documentation. In: Proceedings of the International Conference on Software Maintenance (ICSM 2000). IEEE Computer Society, Washington, DC (2000)
21. Marks, D., Yardley, L.: Research Methods for Clinical and Health Psychology, 3rd edn. Sage Publications, Inc. (2004)
22. Antón, A.I., Earp, J.B.: A requirements taxonomy for reducing web site privacy vulnerabilities. Requirements Engineering (2004)
23. Schneier, B.: A taxonomy of social networking data. IEEE Security Privacy (July-August 2010)
24. Moisil, G.: Lectures on the logic of fuzzy reasoning. Scientific Editions, Bucareat (1975)
25. http://en.wikipedia.org/wiki/List_of_social_networking_websites, List of social networks (2011) (last accessed: December 4, 12)
26. http://www.ico.gov.uk/for_organisations/privacy_and_electronic_communications/the_guide/cookies.aspx, New eu cookie law (e-privacy directive) (last accessed: December 4, 2012)

Electronic Footprints in the Sand: Technologies for Assisting Domestic Violence Survivors

Martin Emms, Budi Arief, and Aad van Moorsel

School of Computing Science, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
{martin.emms,budi.arief,aad.vanmoorsel}@newcastle.ac.uk

Abstract. With the rapid growth and spread of Internet-based social support systems, the impact that these systems can make to society – be it good or bad – has become more significant and can make a real difference to people's lives. As such, various aspects of these systems need to be carefully investigated and analysed, including their security/privacy issues. In this paper, we present our work in designing and implementing various technological features that can be used to assist domestic violence survivors in obtaining help without leaving traces which might lead to further violence from their abuser. This case study serves as the core of our paper, in which we outline our approach, various design considerations – including difficulties in keeping browsing history private, our currently implemented solutions (single use URL, targeted history sanitita-tion agent, and secret graphical gateway), as well as novel ideas for future work (including location-based service advertising and deployment in the wild).

Keywords: Privacy; confidentiality; practical security; browsing history; social inclusion; survivors; domestic violence; intimate partner cyber stalking; support groups; system implementation; work in progress.

1 Introduction

As more and more people are embracing online social networking applications, various concerns have been raised with regard to the security and privacy issues associated with such applications. There have been documented cases and reports regarding violations of user privacy by some of the big companies providing these services (e.g. [18] and [19]), although in most cases, their users seem to be rather oblivious to the threats of supplying their details online without much consideration. It is often stated that human players are usually the weakest link when it comes to computer security [4][9][12], so it is very important to provide a system that requires minimum effort from its users.

When it comes to the consequences of privacy violation, a very poignant example can be drawn from our experience in designing and implementing a system to assist survivors of domestic violence – the term "survivors" is used rather than "victims", as it more accurately describes the individuals who have lived with domestic abuse [17].

There are a number of published works which highlight the issues of domestic abuse and that there is a clear link with *intimate partner cyber stalking*

B. Preneel and D. Ikonomou (Eds.): APF 2012, LNCS 8319, pp. 203–214, 2014.

[6][10][13][14]. There are two aspects to the issue: first, an intimate partner (ex or current) has a greater level of access to and knowledge about the habits of the survivor; second, the cyber stalking is a new and powerful weapon which adds to the ways in which the survivor can be controlled and/or coerced. These works also highlight that the survivors who are being stalked by their partner or ex-partners are of a greatly increased risk of being harmed [10], and that stalking behaviour (whether conducted in cyber space or not) could be viewed as a warning sign of an escalation towards violence. For example, evidence compiled by the US Department of Justice suggests that 81% of women who were stalked by partners were also assaulted by the same partner [16], while the Metropolitan Police found that 40% of domestic violence murders in London were also victims of stalking [11].

Previous studies [3][13] identify that telephones and mobile phones are the most commonly used technology in cases of cyber stalking. Statistics published by the US Department of Justice [3] also show that among 2.4 million victims of cyber stalking in the US, 30.3% of which were stalked by a current or ex-partner, which equates to around 730,000 cases of cyber stalking during 2008. Equally alarmingly, the same report also provides details of the high-tech methods used to monitor the activities of victims, including spyware, video and webcams, listening devices and GPS. In the UK, a survey carried out by the University of Bedfordshire [7] showed that 31.6% of stalkers were either ex-boyfriends (21.2%) or ex-partners (10.4%). Moreover, a report published by the Network for Surviving Stalking and Women's Aid Federation of England [10] details a number of different examples of women being stalked through Facebook, eBay, geotagging, and – most worrying of all – through applications previously loaded onto their smart phone by their abuser which tracks the victims location without their knowledge. However, digital technologies and electronic footprints can also be used to record the actions of abusers, the evidence of which can be used as evidence against them (although this is quite difficult) [15].

1.1 Problem Statement

As we go about our daily lives we are unwittingly leaving *electronic footprints*, which can easily be followed to see what we have been up to. This is because the technologies we use in our everyday lives (such as Internet browsers, mobile phones/smartphones, land lines, and GPS units) maintain records of our activities, which for most of us serve as a convenient aide-mémoire so that we do not have to remember "the number of someone who called yesterday", "the meeting time agreed in an SMS" or "the URL of the website I visited".

In addition to these passive data gathering features, there are also a number of incredibly useful monitoring applications aimed at keeping our children safe, which proactively make use of the data collected from Internet monitoring tools, such as *CheckStick* – `http://www.checkstick.com`, which tells you what your kids are looking at online. In most cases, these tools are very valuable to keep their users safe, but for survivors of domestic violence, these convenient features and monitoring tools can become an instrument of abuse, in that it allows an abuser to track the survivor's activities even when the abuser is not present, and thereby control/restrict the activities of the survivor.

For example, the abuser can control who the survivor can communicate with, monitor what the survivor looks at online, and trace where the survivor travels. All these lead to the intimate partner cyber stalking mentioned earlier.

The reality for a survivor is that any attempt to seek help, either from friends and family or from support organisations, is likely to attract attention and possibly further abuse. This has the effect that although technology is providing more convenient ways for survivors to access the help they require, it is also preventing survivors from accessing those resources. Current Internet browser and mobile phone technologies make it relatively easy for an abuser to review the electronic records that the survivors have collected; conversely it requires a much greater level of technical knowledge and quite a lot of work on the part of the survivors to cover their tracks.

Therefore the technologies that are designed for our convenience inadvertently put survivors at a technological disadvantage. One way to address this problem is by erasing survivor's electronic footprints, but this is not as straightforward as it sounds. Internet browsers and mobile phones will (by default) record their actions, but using a "clear all" approach leaves large gaps in the history, which can also raise abuser's suspicion. To make matters worse for the survivors, each technology stores data in a different way, requiring extra knowledge and effort to effectively remove the traces of their activities. In some cases, data such as mobile phones billing records cannot be altered by the user, leaving them with very limited or even no options.

Key Requirements. The proposed system aims to benefit survivors of domestic violence. These survivors tend to have limited knowledge and experience with technology, and in some cases, English is not their first language [17]. As such, the system needs to be *very easy to use*, with minimal interaction required with its users. In fact, being *invisible* is another key requirement, so that the system does not draw attention from the survivors' abuser. Most – if not all – of the technical activities (such as installing and configuring the proposed system) will be performed by staff at the support centre, with assistance from the authors/developers of the system.

Attacker Model. The main potential attackers will be the survivors' partner (and abuser). They have access to or control of the (shared) computer at home or even the survivors' smartphones. They have sufficient computer knowledge (for example, they know how to check web browser history), but they are not a hacker or an expert in computer security or forensic. They will not monitor the survivors' computer usage all the time (e.g. no key logger or network sniffer will be used). Nonetheless, it is expected that the attacker will be able to take control of the survivor's computer *after* the survivor finishes using it to access domestic violence support websites. Therefore one of the main aims of our proposed solutions (see Section 3) is to remove traces of digital footprints associated with domestic violence support websites from any devices used by the survivors.

1.2 Related Work in Privacy Enhancing Technologies

Issues related to private browsing are not new, and many papers have addressed them to various levels and from different perspectives. Aggarwal et al. [1], considers two

types of attackers threatening private browsing: *local attacker* (family member or other people who has access to the user's machine and might be able to examine its browser's history) and *web attacker* (web sites trying to track and collect data from the user's visit).

Plenty of research has been done in dealing with web and third-party attackers (for example [2][8][20][22]). In our work, however, we are interested in defending against local attackers, which include survivors' abusers in the domestic violence scenario. Portable versions of the popular Internet browsers that allow private or incognito browsing (e.g. Google Chrome [21]) are available to defend against local attackers. There are even more comprehensive solutions such as Tails [23] (which can be deployed as a live USB stick or DVD for preserving anonymity), containing a set of online anonymity tools including Tor [24].

If used correctly, these solutions represent the most effective way of achieving privacy. But they rely on the user being technically savvy or even remembering to turn on and use these features. This is often not the case with survivors of domestic violence. Their knowledge of computer technologies, security and privacy is usually very limited, so it is unreasonable to expect them to be able to use complex features provided by these solutions. Moreover, it is not possible to use USB stick or CD/DVD when accessing information on a mobile phone. Therefore, these solutions – even though they are widely available and provide excellent features for private browsing – constitute only one of the layers of protection that we envisage will be necessary for achieving privacy for these survivors.

2 Case Study: Experience with Survivors

In this paper, we focus on a case study involving survivors who attend a women's support centre for Black and Minority Ethnic (BME), based in the UK (for privacy reasons, we do not state the name of this support centre, instead we refer to it as our "case study"). Data collection was performed using an online survey through several sessions organised by the support centre's staff, in which, groups of survivors as well as women from the control group completed the online survey.

This case study provides us with important insights and experience in the design and development of socio-technical systems where privacy is one of the key features. It also allows us to come up with novel ideas on how new technologies can be used for ensuring privacy. Some of these ideas are still to be implemented, but we are confident that they will contribute positively in improving users' privacy, while being usable and practical at the same time. We are also planning to carry out evaluation of the whole system once it is fully implemented.

The research shows that survivors have two major barriers to successfully accessing the support services that they require [17]:

- locating the support services and the organisations that provide them, and
- fear of provoking further abuse if their abuser discovers that they have been seeking help (hence their reluctance to report the incidents to the police or relatives).

In effect, survivors are being excluded from the socio-technical systems that the rest of us take for granted, largely because they are afraid of using these systems for fear of being found out looking for help. This paper proposes a digital strategy for the social inclusion of survivors. The strategy incorporates several technology-based solutions and a training strategy, which together will help to overcome these barriers. The aims are to publicise domestic violence support services in a way that is most accessible to survivors, while at the same time providing technological solutions that help survivors avoid leaving telltale electronic footprints.

2.1 Method and Implementation

The overall aim of this case study is the *social inclusion of survivors through technology*, and to achieve this, the tasks have been divided into a number of sub-goals:

- Understand how survivors currently relate to technologies that would be useful to them and the technological issues that they face
- Propose a digital strategy to make domestic violence support services more accessible
- Propose a range of technological solutions that help survivors avoid leaving telltale electronic footprints

In order to address these goals, we have worked closely with the staff at the case study's support centre to understand the technological issues faced by survivors. The guidance given by the case study's staff has been invaluable in the development of the technology strategy.

Due to the sensitive nature of the subject, it was not appropriate to use standard user interview techniques to gather the data required. It was felt that an online survey would be a less intrusive way to gather the data, as this could be carried out in the familiar surroundings of the case study's facility with the assistance of its staff, without requiring a member of our research team to be present. The women who completed the survey were selected from women regularly attending services provided by the centre. There were two groups of these women: survivors of domestic violence, and a control group of women who attend the centre for other activities (not related to domestic violence, such as learning English language or new skills). The data were collected from the control group to minimise the influence of other factors common to all women attending this centre, such as socio-economic and/or ethnic group.

The survey collected information relating to the following topics: (i) the location of the computer the women used to access the internet; (ii) websites visited, including any online support services used; (iii) other communications channels such as instant messaging; (iv) type and capabilities of mobile phones used; (v) indication of age range (to eliminate any age related trends); (vi) whether survivors felt they were being monitored; (vii) which support services survivors would like to see implemented.

Two online survey forms were used to collect the data, one for survivors and one for the control group. The control group were not asked questions specifically relating to domestic violence. A multiple-choice format was utilised to obtain the granularity

required and ensure the uniformity of terminology used in responses. The forms can be viewed at:

`http://research.cs.ncl.ac.uk/surveys/survivors-survey.html` and `http://research.cs.ncl.ac.uk/surveys/womens-survey.html`.

There are two competing considerations when deciding which support services would best serve the needs of the survivors:

- preference of the survivors – there is no point providing services which survivors do not want or will not use
- affordability/running costs – it would be counter-productive to create a support service for survivors which has to be withdrawn because it is too expensive to run

Our choices regarding which technologies to use are therefore influenced by these factors and we also endeavour to develop a system that is as easy to use as possible.

2.2 Lessons Learnt

Our contact with the case study commenced in June 2010, and it has involved close collaboration with the staff at the support centre. From this collaboration, we have been able to draw insights and initial conclusions about the issues faced by survivors and the support services used by survivors [17].

One of the major challenges that survivors face is that it requires more effort and more technical knowledge for them to erase their *electronic footprints*, than it does for their abuser to follow them. It is interesting to notice that survivors seem to be aware of the feature of Internet browsers to record the history of the pages they have visited, and that survivors are keen to be able to avoid this. Therefore redressing the balance in favour of the survivor will require a range of measures including redesigned websites, history cleaning technologies and training.

Table 1. Technologies usage of survivors

Category	Survivors	Control
Access to the Internet	71%	100%
Access to computer outside the home (friends, relatives, library)	29%	87%
Used Internet communications such as Skype and IM	43%	87%
Mobile phone usage / ownership	86%	87%

A survey of survivors was performed as part of the case study, to capture first-hand their opinions. A total of 22 women completed the online survey, the results of which have provided valuable insights into – among others – the technology usage of survivors. Table 1 provides a summary of the survivors' usage of technology as compared to that of the control group. The survey results show that the survivors in our sample are 29% less likely than others in their socio-economic/cultural/ethnic group to be regularly using the Internet and the support services it provides. Encouragingly, the survey also shows that mobile phone usage amongst survivors is pretty much equal to

that of the control group of the sample. Although survivors did not express a strong preference for an Internet browser that does not record history, it is felt from our reading of related works mentioned in Section 1 that this will be a valuable tool to implement for survivors. Further information regarding the survey can be found in [17].

3 Proposed Solutions

Our research proposes a number of technology solutions to improve survivors' access to domestic violence support services. We are also focusing on providing a feature for erasing survivors' digital footprints without raising suspicion to their abusers.

3.1 More Accessible Domestic Violence Support Services

There are technologies that can improve the accessibility of support services by ensuring that survivors are not excluded because they do not know that the support service is there. These include *Quick Response* (QR) codes, as well as *Near Field Communication* (NFC) and *Radio Frequency Identification* (RFID) tags.

Fig. 1. QR code example containing a single use URL

QR codes are printed two-dimensional barcodes (Fig. 1) that can encode a URL or a SMS text message and can be read by a smartphone or a laptop with a webcam, allowing easy dissemination of web pages. NFC and RFID tags are a class of wireless information storage device that can store much more data than a QR code and can be read by some smartphones and computers.

With smartphones capable of reading QR codes – and even NFC and RFID tags – becoming more accessible and readily available for survivors to use, we propose embedding information in real world objects using these technologies.

Imagine that you are a close friend or family member of a survivor who is still in an abusive relationship and you want to help by letting her know about online support services which can help, without alerting her partner. What is required is a way of hiding the URL in an everyday object that will not arouse suspicion. QR codes can be printed on self-adhesive labels, making it easy to attach a URL to any real-world object; this could be a poster or flyer advertising the support service or other everyday objects such as a postcard from a friend or on the base of a mug thereby disguising its meaning.

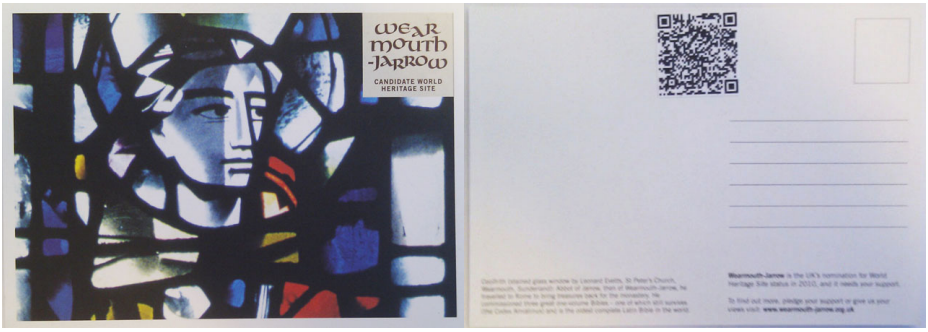


Fig. 2. Postcard with QR code – front (left) and back (right)

An example of a QR code embedded in an innocent-looking postcard can be seen in Fig. 2. The QR code in Fig. 1 can be read with a smartphone and contains a URL which points to a live demo of "single use URL access codes" (see Section 3.2).

QR codes are very cost effective because free software applications can be used to print them, so the only costs are the printing and the sticky labels. This compares very favourably with the cost of NFC, currently around £1.50 per item. NFC tags are much more expensive than QR codes so their use would be limited to applications where the additional functionality they provide is worth the extra expense. NFC objects can carry a great deal more information than QR codes, this provides the opportunity to embed more data. For example, an NFC tag could be used to store a list of all support services in the local area. These tags could then be attached to posters advertising support services, and the survivors can download and view the whole list on their phone without having to connect to the Internet. NFC tags can also give a unique response to each person who accesses the information; this would allow posters to be created that will hand out a different single use URL to each smartphone that accesses the poster.

3.2 Erasing the Digital Footprints of Survivors

Allowing survivors to freely access online resources whilst hiding their activities from their abusers is a complex problem that does not have a single solution. Our approach consists of a number of complementary technologies that provide layers of protection.

Single Use URL Access Codes. Given that a survivor may forget (or even not be aware of how) to use the private browsing feature, or do not know how to clear their history after accessing domestic violence support services online, it is proposed that specific sections of domestic violence websites could incorporate an automatic means of sanititising the browser history of anyone who visits the website.

Our solution hides pages relating to domestic violence support services behind "innocent pages" from a real website that the survivor would quite legitimately use. The website is designed with both innocent pages and domestic violence pages, anyone entering the website without a valid access key would be given innocent pages only.

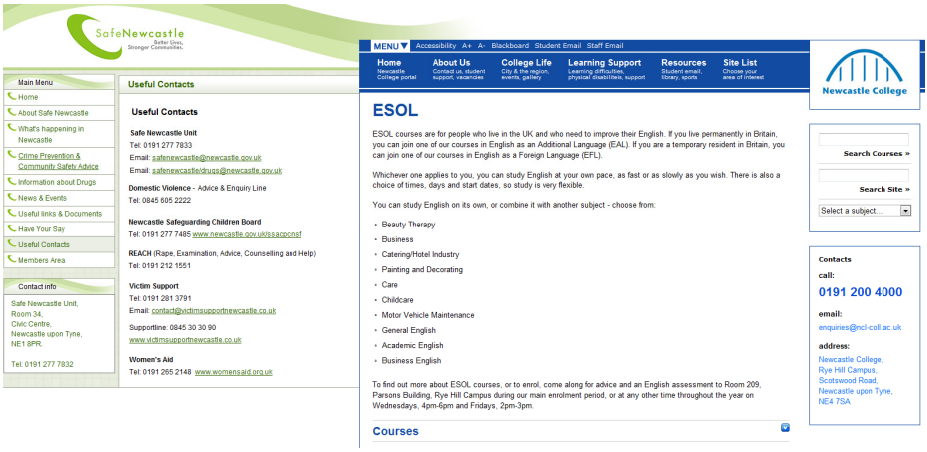


Fig. 3. Domestic violence support directory page (left) and innocent replacement page (right)

This example uses content for ESOL (English for Speakers of Other Languages) courses as the innocent pages, this matches the profile of the women who use the case study's services, many of whom attend ESOL courses to improve their English. Different centres will use different innocent pages, to match the profile of the women attending the different centres. Each access key may only be used *once*; all subsequent attempts to access the domestic violence pages with a used access key will result in innocent pages being presented (an example can be seen in Fig. 3). This stops the abuser from following the browser history to the domestic violence support pages.

The access codes can be distributed in various ways; the method selected should draw the least attention for the survivors who will be using them. Some of the methods we envisage using include: embedding QR codes on postcards, posters, flyers or objects, providing a USB stick containing tools that survivors can use, emailing the URL to survivors, printing the URL on tear-off strips at the bottom of a poster, and sending the URL as a text message.

An algorithm will generate access codes based on the date, so that codes will be valid for a limited time. The access code algorithm will incorporate a checksum to stop random numbers being accepted as valid access codes.

We have implemented a prototype of this solution.

Location-Based Service Advertising. A poster advertising a particular domestic violence support service will be placed in a public location (e.g. bus stops, shopping malls, local shop windows or in the window of the service provider). The poster allows survivors to access online domestic violence service pages and resources on their mobile phone whilst they are at the location of the poster, however once they leave the location, the URL cannot be accessed using the history or back button. This feature will be facilitated by the single use URL mechanism (described previously), which provides the domestic violence support service the first time the URL is used; any subsequent request using the URL code will result in an innocent page being displayed.

Unlike the static Quick Response (QR) codes on postcards/sweets, the poster will give a new single use URL each time a user passes their phone close to the poster. A programmable Near Field Communication (NFC) smart tag is capable of running a small JavaCard program, which will produce a new unique URL for each request. The JavaCard program uses an algorithm to calculate the single use URLs, each URL will be unique and will conform to the validation routine on the web site providing the domestic violence support service.

We will implement this solution soon.

Targeted History Sanitisation Agent. The objective of history sanititation agent is to automatically erase the digital footprints left behind when a user accesses specific support websites, by removing all history entries related to the support websites, including temporary Internet files, browser history entries, and cookies.

The agent will leave intact all other history entries, thereby avoiding making it look like the PC has been cleaned. The agent will automatically download a list of support websites, which will be used to decide which entries to delete; the list will be updated by support centre staff when new support websites go online.

Smartphones are becoming an increasingly popular way for accessing online content. We are therefore developing versions of the history sanitisation agent for Android and iPhone platforms as well. The smartphone agent development has also investigated the ability to automatically cleanse the phone of unwanted entries in the call and SMS history lists. Installation of this agent on smartphones will be carried out at the support centre. It is a bit trickier to deploy the targeted history sanititation agent on the survivors' computer, which tend to be a shared PC at home that their abuser also has access to. We envisage packaging the agent – along with portable anonymous browsers and other privacy tools – into a USB stick or a live CD/DVD that can be distributed to survivors. Training on how to use these tools will be given to survivors by staff at the support centre.

We have implemented a prototype of this solution for Microsoft Windows based PCs supporting various web browsers, as well as for Android smartphones.

Secret Graphical Gateway. The idea is to design and implement an application that will display a set of pictures as the front end of the gateway. When a survivor clicks the correct number of points in the right coordinates on the right picture and in the right sequence (set-up beforehand), the application will direct them to the support services site, otherwise it will do nothing. This way, the application will look innocent and will not raise suspicion to the survivor's abuser. In fact, this gateway application could be disguised as a digital picture viewer. In a sense, this is comparable to graphical password (e.g. [5]), albeit being "invisible" in its nature (without any signposting or obvious interactive feature that might attract attention).

We have implemented a prototype for Android smartphones.

4 Conclusion and Future Work

Through this work, we have demonstrated the need for solutions that will have a significant impact on social inclusion for survivors of domestic violence by improving the accessibility of domestic violence support service and by improving the ability of survivors to avoid leaving electronic footprints when they access these services. The case study shows that existing technologies utilised by survivors unintentionally work contrary to these aims. Given that this situation is unlikely to change and there is a limited budget for this project, we have adopted a strategy that proposes a set of bite-sized solutions, each of which will be relatively quick implement at a modest cost.

We have implemented and tested the single use URL access codes idea, the targeted history sanititation agent, as well as the secret graphical gateway. However, wider deployment and further evaluation of the effectiveness of these solutions are still to be carried out. We will complete the implementation of other novel ideas, including the location-based domestic violence support advertising proposed in Section 3.2, either as proof of concept demonstrations or as fully functional solutions soon.

We plan to continue working with the case study's staff and survivors, to gather additional data and results through questionnaires and "in the wild" deployment of the proposed solutions, including their usability assessment. We will also explore other potential avenues for effective solution by conducting participatory, experience-centred design process with the survivors and the staff at the support centre.

Acknowledgement. We would like to thank the staff at the support centre of the case study, as well as the survivors who participated in our survey for providing invaluable insights into the survivors' story. We also appreciate the anonymous reviewers' feedback and comments, which helped us improve this paper.

References

1. Aggrawal, G., Bursztein, E., Jackson, C., Boneh, D.: An analysis of private browsing modes in modern browsers. In: Proceedings 19th USENIX Security Symposium (2010)
2. Krishnamurthy, B., Malandrino, D., Wills, C.E.: Measuring privacy loss and the impact of privacy protection in web browsing. In: Proceedings of the 3rd Symposium on Usable Privacy and Security (SOUPS 2007), pp. 52–63. ACM, New York (2007)
3. Baum, K., Catalano, S., Rand, M., Rose, K.: Stalking victimization in the US. US Department of Justice National Crime Victimisation Survey (January 2009)
4. Besnard, D., Arief, B.: Computer security impaired by legitimate users. Computers & Security 23(3), 253–264 (2004)
5. Dunphy, P., Yan, J.: Do background images improve "draw a secret" graphical passwords? In: Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS 2007), pp. 36–47. ACM, New York (2007)
6. Logan, T., Walker, R.: Partner stalking: Psychological dominance or business as usual? Trauma Violence Abuse 10(3), 247–270 (2009)
7. Maple, C., Short, E., Brown, A.: Cyberstalking in the United Kingdom: An Analysis of the ECHO Pilot Survey. University of Bedfordshire National Centre for Cyberstalking Research (2011)

8. Mayer, J.R., Mitchell, J.C.: Third-Party Web Tracking: Policy and Technology. In: Proceedings of the IEEE Symposium on Security and Privacy, pp. 413–427 (2012)
9. Mitnick, K., Simon, W.: The art of deception: Controlling the human element of security. Wiley (2002)
10. Perry, J.: Digital stalking: A guide to technology risks for victims. Published jointly by Network for Surviving Stalking and Women's Aid Federation of England (2012)
11. Richards, L.: Findings from the Multi-agency Domestic Violence Murder Reviews in London. Prepared for the ACPO Homicide Working Group, Metropolitan Police (2003)
12. Sasse, M.A., Brostoff, S., Weirich, D.: Transforming the weakest link - a human computer interaction approach to usable effective security. BT Tech. Journal 19, 122–131 (2001)
13. Southworth, C., Dawson, S., Fraser, C., Tucker, S.: A high-tech twist on abuse: Technology, intimate partner stalking and advocacy. Violence Against Women Online Resources (June 2005)
14. Southworth, C., Finn, J., Dawson, S., Fraser, C., Tucker, S.: Intimate Partner Violence, Technology, and Stalking. Violence Against Women 13(8), 842–856 (2007)
15. Spence-Diehl, E.: Stalking and technology: The double edge sword. Technology in Human Services 22(1), 5–18 (2003)
16. Tjaden, P., Thoennes, N.: Stalking in America: Findings form the national violence against women survey. US Dept. of Justice (1998)
17. van Moorsel, A., Emms, M., Rendall, G., Arief, B.: Digital Strategy for the Social Inclusion of Survivors of Domestic Violence. Technical Report CS-TR-1277, School of Computing Science, Newcastle University (September 2011)
18. BBC News Online: Details of 100m Facebook users collected and published, `http://www.bbc.co.uk/news/technology-10796584` (last accessed: November 27, 2012)
19. BBC News Online: Facebook's battle with privacy and profit, `http://news.bbc.co.uk/1/hi/programmes/click_online/8843007.stm` (last accessed: November 27, 2012)
20. Digital Trends: Why Do Not Track may not protect anybody's privacy, `http://www.digitaltrends.com/mobile/why-do-not-track-may-not-protect-anybodys-privacy/` (last accessed: November 27, 2012)
21. Google Chrome: Using the Incognito mode, `https://support.google.com/chrome/bin/answer.py?hl=en-GB&answer=95464&p=cpn_incognito` (last accessed: November 27, 2012)
22. Panopticlick: How Unique – and Trackable – Is Your Browser?, `https://panopticlick.eff.org/` (last accessed: November 27, 2012)
23. Tails: The Amnesic Incognito Life System, `https://tails.boum.org/` (last accessed: November 27, 2012)
24. Tor Project: Anonymity Online, `https://www.torproject.org/` (last accessed: November 27, 2012)

Author Index

Anthonysamy, Pauline 187
Arief, Budi 203

Beckers, Kristian 1
Beyerer, Jürgen 73
Bier, Christoph 73
Birnstill, Pascal 73
Bjones, Ronny 111
Bock, Kirsten 17
Boonstra, Daniel 55

Colombo, Pietro 17
Cook, Blaine 125

Domany, Tamar 17
Dumortier, Jos 157

Emms, Martin 203
Everts, Maarten 55

Faßbender, Stephan 1
Ferrari, Elena 17

Graux, Hans 157
Greenwood, Phil 187

Halpin, Harry 125
Hartman, Alan 17
Heisel, Maritta 1

Kerschbaum, Florian 41
Klitou, Demetrius 86
Kosta, Eleni 157
Krempel, Erik 73
Krontiris, Ioannis 111
Kung, Antonio 177
Kveler, Ksenya 17

Liagkou, Vasiliki 140

Meis, Rene 1
Metakides, George 140
Moorsel, Aad van 203

Paillier, Pascal 111
Pyrgelis, Apostolis 140

Rannenberg, Kai 111
Raptopoulos, Christoforos 140
Rashid, Awais 187

Spirakis, Paul 140
Stamatiou, Yannis C. 140

Vagts, Hauke 73
van Paassen, Ron 55
van Rest, Jeroen 55
van Rijn, Martin 55