Software Assisted Privacy Impact Assessment in Interactive Ubiquitous Computing Systems

Alfredo Pérez Fernández, Guttorm Sindre

Software Assisted Privacy Impact Assessment in Interactive Ubiquitous Computing Systems

Alfredo Pérez Fernández and Guttorm Sindre[1]

Norwegian University of Science and Tecnology, Trondheim NO 7491, Norway,
{alfredo.perez.fernandez,guttorm.sindre}@ntnu.no,
WWW home page: http://www.perezfer.com/

Abstract. Developing ubiquitous computing systems in compliance with the data protection regulation is a difficult task. The European General Data Protection Regulation requests system developers to apply a privacy-by-design methodology and perform privacy impact assessments throughout the whole development life-cycle. Our proposal is a software assisted process framework that facilitates the analysis of privacy implications in ubiquitous computing systems. This software has been evaluated with students and ubicomp experts.

Keywords: privacy, privacy-by-design, privacy-impact-assessment, internet-of-things

1 Introduction

Through history, different advances in technology have tended to facilitate the flow of information of any kind, including personal information. Probably, one of the first and most cited references that confirm this idea is the law review article written by Warren and Brandeis in 1890, *The Right to Privacy*, where they say: "Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life" [32]. Personal computing and the advent of the Internet also supposed a threat to personal privacy since the beginning, as David and Fano mention: "If every significant action is recorded in the mass memory of a community computer system, and programs are available for analyzing them, the daily activities of each individual could become open to scrutiny" [6]. The last great revolution of technology we are witnessing is the Internet of Things (IoT) and it is also raising the concern among the population [4]. The public administration is aiming at maintaining the situation under control by hardening privacy regulations. The new European General Data Protection Regulation (GDPR) (https://www.eugdpr.org/) is applicable from May 25th 2018 and it requires system developers and engineers to acquire a privacy-by-design (PbD) approach during the development and conduct privacy impact assessments (PIA). However, engineers are not given concrete indications of which steps, operations or methods should be used to accomplish that [3, 20, 28]. The focus now is to find strategies to operationalize PbD but there is still a gap between the set of principles and the specific tasks that

need to be performed [15]. This gap is even more pronounced if we look at IoT and ubiquitous computing scenarios, since most of the proposed frameworks are developed with PC-based applications in mind, not necessarily suitable for a situation where much of the user's interaction with information systems will be through surrounding *smart things*. Even though there are a large number of methods, models and frameworks designed to guide developers in the analysis of privacy threats in many different scenarios, we find that not many of these tools have been implemented in software to automate the process. Our contribution is a software implementation of the Privacy Aware Transmission Highway (PATH) [25] process framework. This assistant facilitates the evaluation of privacy risks in interactive ubiquitous computing scenarios. This software has been evaluated with students and experts in ubiquitous computing systems.

In Section 2 of this paper, we analyze the related work on frameworks for supporting privacy-by-design in HCI and ubiquitous computing. Section 3 gives a brief description of our research methodology. We describe our proposed framework and the adaptation to the software platform in Section 4. Section 5 shows the scenarios used for the evaluation of our framework. In Section 6 we show the result of the evaluation. Finally, Section 7 provides some concluding remarks.

2 Related Work

Iachello [11] identified existing privacy frameworks and methods in the field of HCI, grouped as *guidelines* [8, 9, 17], *process frameworks* [7, 12, 31] and *model frameworks* [13, 18, 30]. STRAP [12] and PriFs [31] focus specifically in requirements elicitation combined with goal oriented analysis methods [5, 16]. Spiekermann [29] proposes a framework specific to *RFID* to identify privacy vulnerabilities based on previously defined privacy targets. Inah Omoronyia [21] developed PSatAnalyser a software tool to assist the analysis of smart objects based architectures from a privacy-by-design perspective. The Software Assurance Technology Center (SATC) developed a software tool [33] that made use of natural language processing techniques to analyze the quality of the requirements document based on the structure of the sentences that described the requirements. They compiled a list of quality attributes that could be measured following this approach. Natural language processing has been proposed, as well, to analyze the privacy policies [1] with respect to *vagueness* with the objective of estimating the perception of privacy risks by the users of the system. In February 2018, the Commission Nationale de l'Informatique et des Libertés (CNIL) released a template to conduct a PIA on an IoT based scenario assisting the evaluator through a multiplatform application (https://www.cnil.fr/en/privacy-impact-assessment-pia). Even thought the templates system provides guidance with respect to the type of information that is needed in order to conduct the PIA, the system does not perform any type of automated verification or validation of the entered information, other than checking that there are no empty fields and providing guidance to the user on how this information should be elaborated. There are a number of limitations in the existing solutions, including, lack of evaluation in real-case

scenarios, focus outside of ubiquitous computing scenarios and limited process automation.

3 Research Method

The Design Science Research Methodology (DSRM) [22] has been used to conduct our research, since our motivation was not limited to gaining understanding of privacy-by-design as phenomena, but also to generate an asset that could be used by system engineers to ease the development and analysis of privacy aware ubiquitous computing systems. Our approach is to conduct the six activities of DSRM following the nominal sequential order: *problem identification and motivation*, *specifying the objective of the solution*, *design and development*, *demonstration*, *evaluation* and *communication*.

- **Problem identification and motivation**: Our initial experience, the analysis of the literature and the findings after iterating over the research process indicate that operationalizing PbD cannot be considered a straightforward process and needs to be assisted in some form.
- **Specifying the objective of the solution**: The overall objective of our research is to develop an asset that can be used by system engineers in the analysis and implementation of ubiquitous computing systems. This objective is divided into two, describing a process framework to guide the engineers and implementing a software that facilitates following such a process.
- **Design and development**: The requirements of the process and the software tool are elicited taking real case scenarios as examples. To do that, we contacted a group of experts and system engineers. The requirement for selecting the candidates was that they needed to be involved on the development of a ubiquitous computing project, excluding system engineers and developers of classical computer based applications. The feedback obtained from the evaluations is used to improve the design of the PATH framework.
- **Demonstration**: To demonstrate the usability of the PATH framework and the PATH assistant, we apply it to a case study application of our own, a prototype implementation of a body coupled communication (BCC) [35] device.
- **Evaluation**: The evaluation of the PATH assistant takes place in the form of empirical experiments where experts apply it to their own projects. The PATH framework is also evaluated in comparison with other benchmark frameworks (the results of this study are pending publication).

During the different iterations, we have constrained our research with two main assumptions. First, we consider the reference model proposed by Ziegeldorf [34] as a starting point, if there is a privacy incidence in a ubiquitous computing scenario it has to take place as the result of the interaction between the user and the surrounding *smart things* (which we prefer to call *interaction mechanisms*). And second, the reason why that privacy incidence is caused by an *interaction mechanism* is because of one or more of its attributes.

4 The Software Assistant

As a proof of concept, a software assistant has been implemented following the guidelines of the *Privacy Aware Transmission Highway* (PATH) framework [25]. The PATH framework consists of four phases that are applied iteratively during design and development: goal-oriented analysis (GOA), elaboration and incorporation of the Privacy Related Interaction Vocabulary (PRIV), PRIV based evaluation and iteration.

4.1 Goal-oriented analysis

The PATH assistant starts by requesting the participant to introduce a textual description of the project with the objective of identifying implicit over specifications of the interaction mechanism [26]. The text of the description is compared against a reduced database with preselected interaction mechanisms. This description is compared against the database of interaction mechanisms to identify potential over specifications(Figure 1). If the PATH assistant detects that the

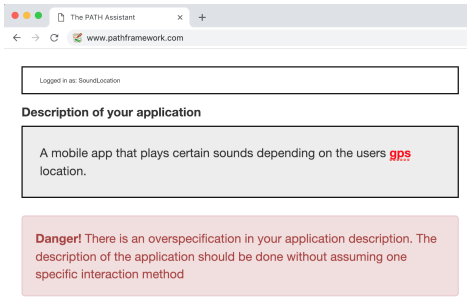

Fig. 1. Screen capture of the automated goa for the high level description of the application

description includes an implicit over specification a warning is displayed with the corresponding term highlighted (In the example given, the expert is introducing an over specification that the application needs to use GPS signals to detect the user's location). The interaction mechanisms are selected separately from a list (or included if they are not present) (Figure 2).

4.2 Privacy Related Interaction Vocabulary

After the GOA phase, a list of attributes for each interaction mechanism are presented to the user (Figure 3). The user selects those that are applicable to the scenario that was described during the GOA phase. If an attribute of an interaction mechanism is identified, it can be added to the list at this point and

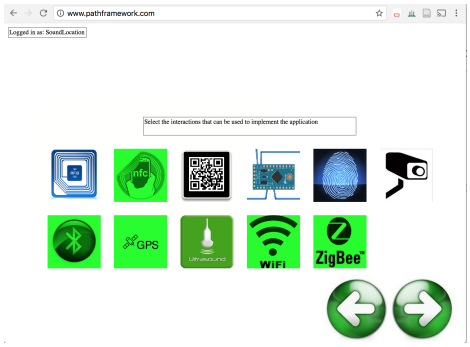

Fig. 2. Selection of the alternative interaction mechanisms

it will be considered for the rest of interaction mechanisms. The attributes of the interaction mechanisms can be ranked from 1 (very low) to 5 (very high) depending on how they are estimated by the user.

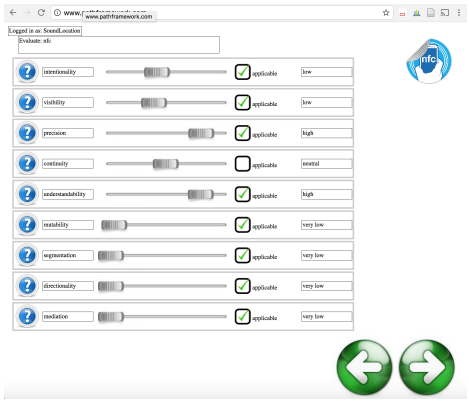

Fig. 3. Estimation of Interaction Mechanism Attributes

4.3 Privacy Impact Assessment

After the value of the attributes have been estimated for each interaction mechanism, another estimation is given for the likelihood of that attribute impacting negatively on users' privacy. The PATH assistant summarizes a chart with the given values and calculated uncertainties that can lead to privacy threats (Figure 4). Those controversial attributes that present more disagreements and uncer-

tainty are more relevant candidates to be investigated in a user evaluation with a prototype.

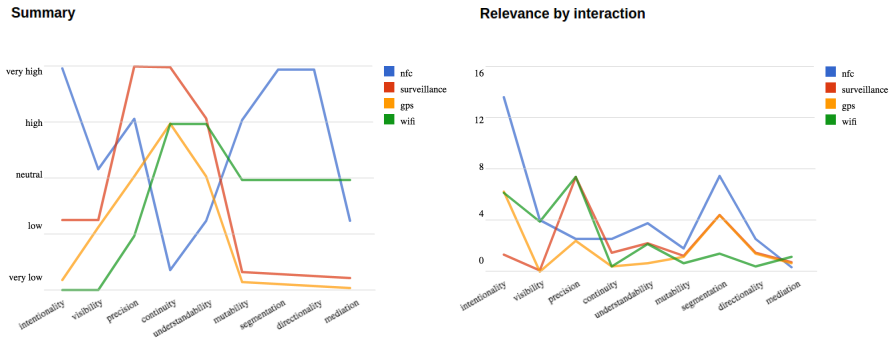

Fig. 4. Overview of attributes estimation (Summary) and uncertain attributes that have an impact on privacy (relevance)

4.4 Iteration

Based on the findings from the PIA, it should be possible to replace or modify the interaction mechanisms or their implementations so that the impact on user's privacy is minimized. After any change is made in the design or the implementation of the system, a new iteration needs to be done to evaluate the new changes.

5 Framework Evaluation Scenarios

The PATH assistant has been evaluated with eight experts in four different scenarios, the Body Coupled Communication Based Shopping, the Adressapark, the Museums Visitor Tracker and the Location Based Sound Player.

5.1 Body Coupled Communication Based Shopping Scenario

The Body Coupled Communication (BCC) [23,35] Based Shopping [14] is a user scenario utilized at Linköping University (LiU) to frame their research. In this scenario, a retail store customer holds a BCC enabled wearable or mobile device that receives random product information when she touches a smart tag situated in the shelf next to the product label (Figure 5, left).

Fig. 5. Left, A customer retrieving product information in the BCC Shopping Scenario (Image facilitated by Acreo, Ri.Se, Linköping University), Right: Scene of a typical use case scenario of the Adressa Park (Image facilitated by Institutt for Elektronikk og Telekommunikasjon, IET)

5.2 Adressaparken

Adressaparken [19] is an interactive media space developed as a collaboration between the municipality of Trondheim (Trondheim kommune), the regional newspaper Adressa (Adresseavisen), and the Norwegian University of Science and Technology (NTNU). An audiovisual storytelling content is projected on the ground while the user walks near the area (Figure 5, right). A set of 12 boxes contains different sensors (temperature, air, light, sun, noise and pollution) and each box holds a Raspberry Pi v2 with an attached night vision infrared camera.

5.3 Museum Visitors Tracker

The Museum Visitors Tracker project was originated as a collaboration between the Technology Transfer Office (TTO) at NTNU and the Science Museum in Trondheim (Vitensenteret). Vitensenteret had already implemented a computer vision system to track visitors and estimate their engagement based on their facial expression (Figure 6). The goal was to implement a privacy-friendly tracking system to obtain statistical information about visitors (age range, gender and preference group) linked to the engagement metrics (time spent) for each exhibition. To achieve this goal, it was planned to evaluate the use of BCC as an interaction mechanism in a similar way as proposed by [24]

5.4 Location Based Sound Player

A *do-it-yourself* (DIY) practitioner started a personal project to conceptualize an augmented noise application [10] similar to the echoes app (https://www.echoes.xyz). The idea of the project was to create a collection of geographically tagged sounds that could be played when the user had visited the associated locations. Since the project was at an early stage only a few high

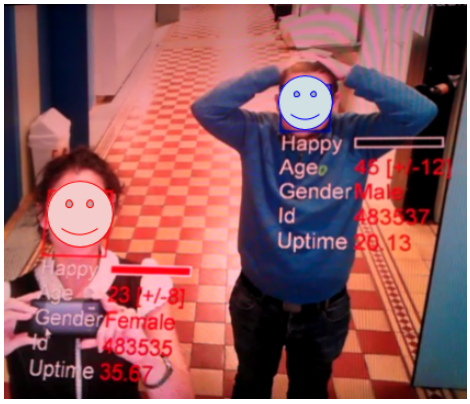

Fig. 6. Example of visitors tracking based on computer vision and facial expression

level decisions were specified. Whether the sounds had to be played immediately without supervision of the user or not was to be decided in later stages of development.

6 Evaluation results

In our evaluation of the GOA phase with the PATH assistant, two of the eight experts introduced a high level description of an application that was detected by the system as over specified, one for the BCC Shopping (for the term BCC) and one for the Location Based Sound Player (for the term GPS). A total of 24 interaction mechanisms were selected by the experts, an average of three each. For the attributes estimations a total of 216 attributes estimations were given by the eight experts. Each estimation ranged from 1 (very low) to 5 (very high). They gave an average estimation of 3.59 with a standard deviation of 1.27 for customer related attributes (CA), intentionality, visibility, precision and understandability, and an average estimation of 2.57 with a standard deviation of 1.63 for non customer related attributes (NCA), continuity, mutability, segmentation, directionality and mediation (Figure 7). From the given attribute estimations, 38 (18%) were accounted as a strong disagreement. We consider a strong disagreement in the estimation of two attributes when two team members give opposed values (high or very high against low or very low). If the team members participate in the same project, the disagreements are considered internal and, if the project is different, the disagreements are considered external. In the Adressa Park four experts participated in the evaluation. Three of them, selected video recording as interaction mechanism, since that was already specified as a requirement for the project. From the nine estimations they assigned to the attributes, five were in disagreement with the estimations provided by the rest of the team (internal disagreements)(Figure 8).

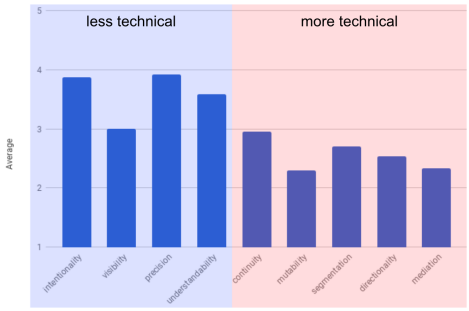

Fig. 7. Overall 216 attribute estimations provided by the experts

Attribute	Expert (W)	Expert (J)	Expert (A)	Expert (Z)
intentionality	4	1	4	0
visibility	1	4	3	0
precision	5	4	4	0
continuity	5	4	3	0
understandability	4	4	5	0
mutability	5	1	1	0
segmentation	5	1	5	0
directionality	1	1	3	0
mediation	5	1	1	0

Fig. 8. Attribute estimations for the video recording interaction mechanism provided by the Addressaparken team.

The experts gave a total of 24 estimations for the impact on privacy. The average for the CAs was 3.6 with standard deviation 1.27. The average for the NCAs was 2.57 with a standard deviation of 1.63 (Figure 9). As can be observed

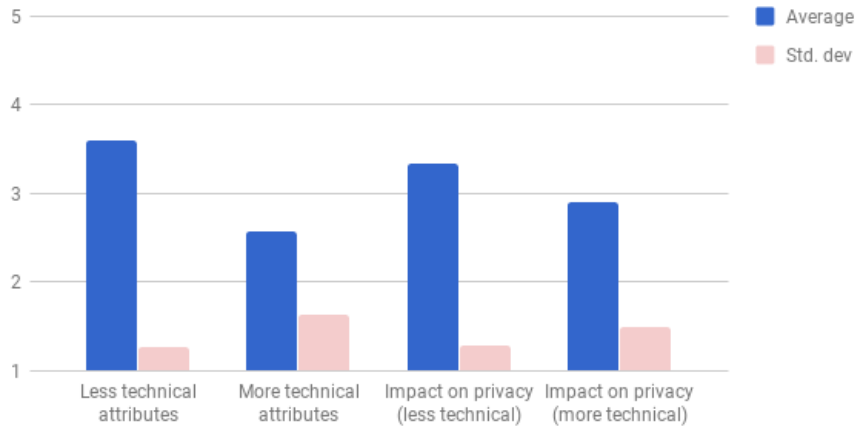

Fig. 9. Average estimated attributes and average estimated privacy impact

from the difference in the standard deviation, experts tend to disagree more when estimating attributes or impact on privacy for NCAs. Another observation is that estimated impact on privacy on CAs is lower than the estimated corresponding attribute, while the behaviour is the opposite for NCAs, where estimated impact on privacy is higher than the corresponding estimated value of the corresponding attribute. Our interpretation of this phenomena is that the practitioners tend to give lower values to attributes that are more difficult to understand. We also conclude that terms they understand less are associated with higher risks.

7 Conclusions and Future work

In this paper we have demonstrated our approach to automate the PIA of ubiquitous computing systems. We have observed that it is possible to assist the process of over specification detection based on the analysis of high level system descriptions. Disagreements in the estimation of interaction mechanisms attributes can be used, as well, as indicators of uncertainties in the PIA of a system. Our prototype, the PATH assistant, only includes a basic search through a small database to identify names of interaction mechanisms. A much better

performance in the automated detection of over specified requirements could be achieved by incorporating a more thorough natural language processing [27]. Elaborating new terms that can be incorporated to the PRIV is a difficult task since the attributes of the interaction mechanisms are not easily perceived and some are not so applicable to the privacy domain. We are considering the incorporation of two new terms in the initial PRIV, *wearability* (when the interaction mechanism is being carried by the user and this could expose the location of the user, i.e. *RFIDs* used in clothes) [2] and *personability* (when the interaction mechanism is tightly associated with a person or group of people, their identities or personal information, i.e. mobile devices are innocuous when bought new but, as they are used, they absorb more and more personal information after the user logs in with different accounts, synchronizes the device with other previous devices or take some pictures) [18]. The current implementation of the PATH assistant is in the process of being adapted as an Asana (https://www.asana.com/) plugin. It makes use of the project description to detect potential over specifications and generates a report based on identified risks for user's privacy caused by different interaction mechanisms. This implementation will be used for future empirical evaluations with experts and students.

References

1. J. Bhatia, T. D. Breaux, J. R. Reidenberg, and T. B. Norton. A theory of vagueness and privacy risk perception. In *Requirements Engineering Conference (RE), 2016 IEEE 24th International*, pages 26–35. IEEE, 2016.
2. M. Bylund, K. Hk, and A. Pommeranz. Pieces of Identity. In *Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges*, NordiCHI '08, pages 427–430, New York, NY, USA, 2008. ACM.
3. A. Cavoukian and C. Staff. Operationalizing Privacy by Design. *Commun. ACM*, 55(9):7–7, Sept. 2012.
4. V. Chellappan and K. M. Sivalingam. Chapter 10 - Security and privacy in the Internet of Things. In R. Buyya and A. Vahid Dastjerdi, editors, *Internet of Things*, pages 183–200. Morgan Kaufmann, Jan. 2016.
5. R. Darimont, E. Delor, P. Massonet, and A. van Lamsweerde. GRAIL/KAOS: an environment for goal-driven requirements engineering. In *Proceedings of the 19th international conference on Software engineering*, pages 612–613. ACM, 1997.
6. E. E. David, Jr. and R. M. Fano. Some Thoughts About the Social Implications of Accessible Computing. In *Proceedings of the November 30December 1, 1965, Fall Joint Computer Conference, Part I*, AFIPS '65 (Fall, part I), pages 243–247, New York, NY, USA, 1965. ACM. event-place: Las Vegas, Nevada.
7. S. A. ElShekeil and S. Laoyookhong. GDPR Privacy by Design. 2017.
8. S. Garfinkel. Adopting fair information practices to low cost RFID systems. In *Privacy in Ubiquitous Computing Workshop*, 2002.
9. R. Gellman. Fair information practices: A basic history. 2017.
10. S. Hastrup. Augmented Noise - Exploring mobile technology design as an enabler of social interaction and spatial awareness. 2017.
11. G. Iachello and G. D. Abowd. Privacy and proportionality: adapting legal evaluation techniques to inform design in ubiquitous computing. In *Proceedings of the*

SIGCHI conference on Human factors in computing systems, pages 91–100. ACM, 2005.
12. C. Jensen, J. Tullio, C. Potts, and E. D. Mynatt. STRAP: a structured analysis framework for privacy. 2005.
13. X. Jiang and J. A. Landay. Modeling privacy control in context-aware systems. *Pervasive Computing, IEEE*, 1(3):59–63, 2002.
14. M. I. Kazim. *Variation-Aware System Design Simulation Methodology for Capacitive BCC Transceivers*. PhD Thesis, Linköping University Electronic Press, 2015.
15. I. Kroener and D. Wright. A strategy for operationalizing privacy by design. *The Information Society*, 30(5):355–365, 2014.
16. A. Kung, F. Kargl, S. Suppan, J. Cuellar, H. C. Phls, A. Kapovits, N. N. McDonnell, and Y. S. Martin. A Privacy Engineering Framework for the Internet of Things. In *Data Protection and Privacy: (In)visibilities and Infrastructures*, Law, Governance and Technology Series, pages 163–202. Springer, Cham, 2017.
17. M. Langheinrich. Privacy by designprinciples of privacy-aware ubiquitous systems. In *Ubicomp 2001: Ubiquitous Computing*, pages 273–291. Springer, 2001.
18. J. T. Lehikoinen, J. Lehikoinen, and P. Huuskonen. Understanding privacy regulation in ubicomp interactions. *Personal and Ubiquitous Computing*, 12(8):543–553, Nov. 2008.
19. W. A. Mansilla and A. Perkis. Multiuse Playspaces: Mediating Expressive Community Places. *IEEE MultiMedia*, 24(1):12–16, Jan. 2017.
20. Y. S. Martín, J. M. d. Alamo, and J. C. Yelmo. Engineering privacy requirements valuable lessons from another realm. In *2014 IEEE 1st International Workshop on Evolving Security and Privacy Requirements Engineering (ESPRE)*, pages 19–24, Aug. 2014.
21. I. Omoronyia. Privacy engineering in dynamic settings. In *Software Engineering Companion (ICSE-C), 2017 IEEE/ACM 39th International Conference on*, pages 297–299. IEEE, 2017.
22. K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee. A design science research methodology for information systems research. *Journal of management information systems*, 24(3):45–77, 2007.
23. A. Pérez Fernández. Towards the Tangible Hyperlink. In *ACHI 2014, The Seventh International Conference on Advances in Computer-Human Interactions*, pages 17–20, 2014.
24. A. Pérez Fernández and G. Sindre. Protecting User Privacy when Sharing Mobile Devices: Research in Progress. In *Norsk informasjonssikkerhetskonferanse (NISK)*, volume 7, Nov. 2014.
25. A. Pérez Fernández and G. Sindre. The privacy aware transmission highway framework. *International Journal of Information Privacy, Security and Integrity*, 3(4):327–350, 2018.
26. A. Pérez Fernández and G. Sindre. Mitigating the Impact on Users Privacy Caused by over Specifications in the Design of IoT Applications. *Sensors, Special Issue Security, Privacy, and Trustworthiness of Sensor Networks and Internet of Things*, 19(19):4318(1–20), Oct. 2019.
27. G. Sizov. Automating Problem Analysis Using Knowledge Extracted from Text. 2017.
28. S. Spiekermann. The challenges of privacy by design. *Communications of the ACM*, 55(7):38–40, 2012.

29. S. Spiekermann. The RFID PIA Developed by Industry, Endorsed by Regulators. In *Privacy Impact Assessment*, Law, Governance and Technology Series, pages 323–346. Springer, Dordrecht, 2012.
30. S. Spiekermann and L. F. Cranor. Engineering privacy. *IEEE Transactions on software engineering*, 35(1):67–82, 2009.
31. K. Thomas, A. K. Bandara, B. A. Price, and B. Nuseibeh. Distilling privacy requirements for mobile applications. In *Proceedings of the 36th International Conference on Software Engineering*, pages 871–882. ACM, 2014.
32. S. D. Warren and L. D. Brandeis. The right to privacy. *Harvard law review*, pages 193–220, 1890.
33. W. M. Wilson, L. H. Rosenberg, and L. E. Hyatt. Automated analysis of requirement specifications. In *Proceedings of the 19th international conference on Software engineering*, pages 161–171. ACM, 1997.
34. J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle. Privacy in the Internet of Things: threats and challenges. *Security and Communication Networks*, 7(12):2728–2742, 2014.
35. T. G. Zimmerman. Personal area networks: near-field intrabody communication. *IBM systems Journal*, 35(3.4):609–617, 1996.