UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Six Privacy and Usability Heuristics: from grounded models to validated new heuristics of usable privacy

André de Lima Salgado

Tese de Doutorado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

ICMC USP SÃO CARLOS

André de Lima Salgado

Six Privacy and Usability Heuristics: from grounded models to validated new heuristics of usable privacy

Thesis submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP – in accordance with the requirements of the Computer and Mathematical Sciences Graduate Program, for the degree of Doctor in Science. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Profa. Dra. Renata Pontin de Mattos Fortes

USP – São Carlos
May 2022

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

d278s

de Lima Salgado, André
Six Privacy and Usability Heuristics: from grounded models to validated new heuristics of usable privacy / André de Lima Salgado; orientadora Renata Pontin de Mattos Fortes. -- São Carlos, 2022.
159 p.

Tese (Doutorado - Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2022.

1. privacy. 2. usability. 3. heuristic. 4. evaluation. 5. guideline. I. Pontin de Mattos Fortes, Renata , orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

André de Lima Salgado

Seis heurísticas de privacidade e usabilidade: de modelos fundamentados a novas heurísticas validadas de privacidade usável

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientadora: Profa. Dra. Renata Pontin de Mattos Fortes

USP – São Carlos
Maio de 2022

Ao autor da vida e à você,
dedico.

ACKNOWLEDGEMENTS

I am grateful for the gift of science given by the Lord for the good of all people.

I want to thank my family and friends for their support during this doctoral research, especially my wife, for her steadfast support during the times this thesis was written.

Special thanks to Professor Renata Fortes for her kind and patient supervision during my studies. Also, I would like to thank all my ICMC/USP colleagues, faculty, and administrative staff.

Special thanks to Professor Patrick Hung for his supervision beyond my research internship abroad. Also, I would like to thank all my Ontario Tech University colleagues, faculty, and administrative staff.

This study was supported by the grants 2017/ 15239-0 and 2018/26038-8, São Paulo Research Foundation (FAPESP).

This study was funded in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

“Be Water, My Friend.
Empty your mind.
Be formless, shapeless, like water.
You put water into a cup, it becomes the cup.
You put water into a bottle, it becomes the bottle.
You put it into a teapot, it becomes the teapot.
Now water can flow or it can crash.
Be water, my friend.”
(Bruce Lee)

RESUMO

SALGADO, A. L. **Seis heurísticas de privacidade e usabilidade: de modelos fundamentados a novas heurísticas validadas de privacidade usável**. 2022. 174 p. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

A privacidade usável é mais que um requisito legislativo, é atributo de qualidade do software. Aumentar a transparência de interfaces de políticas de privacidade é um desafio que a ciência da computação deve enfrentar para aumentar a confiança dos usuários na tecnologia. Ao invés de fornecer políticas de privacidade longas e complexas, precisamos projetar interfaces mais usáveis que capacitem pessoas leigas a proteger sua privacidade online. Nesta tese, objetivamos criar critérios de usabilidade para inspecionar tais interfaces. Após uma análise secundária qualitativa, composta por uma revisão da literatura snowballing, análise temática, análise de cluster e avaliação empírica, esta tese cria seis heurísticas de privacidade e usabilidade (push#). Quando aplicadas para avaliar interfaces de políticas de privacidade para leigos, as heurísticas push# aprimoram a utilidade downstream no número de problemas catastróficos descobertos. Também criamos diretrizes preliminares de privacidade e usabilidade (pug#) e modelamos um novo processo para a criação de novos critérios de usabilidade. Além disso, esta tese também fornece: recomendações para a usabilidade dos controles de privacidade dos pais; um mapeamento sistemático de heurísticas de usabilidade para interfaces de políticas de privacidade; modelos de avaliação heurística para avaliadores novatos; um método para aprimorar a usabilidade de políticas de privacidade com análise de card sorting; visão geral das expectativas de privacidade na experiência do usuário em relação a veículos autônomos-conectados; protótipo de interface baseada em gestos para aumentar a privacidade nos sistemas de saúde; uma ontologia preliminar para descobertas de usabilidade; e heurísticas de usabilidade para jogos móveis e jogadores idosos. Discutimos como estudos futuros podem explorar o uso de nossas heurísticas e diretrizes para domínios específicos, como interação humano-robô e interação humano-inteligência artificial. Finalmente, propomos o estudo da usabilidade sugestiva para aprimorar a proteção da privacidade, independentemente das ferramentas de proteção da privacidade

Palavras-chave: Privacidade, Usabilidade, Heurística, Avaliação..

ABSTRACT

SALGADO, A. L. **Six Privacy and Usability Heuristics: from grounded models to validated new heuristics of usable privacy**. 2022. 174 p. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

Usable privacy is rather a legislative requirement than only a software quality. Enhancing the transparency of privacy policy interfaces stands as a challenge that computer science must address to enhance users' trust in technology. Instead of providing long and complex privacy policies, we need to design more usable interfaces that empower laypeople to protect their privacy online. In this thesis, we aimed at creating broad usability criteria for inspecting such interfaces. After a qualitative secondary analysis, composed of snowballing literature review, thematic analysis, cluster analysis, and empirical evaluation, this thesis creates six usable privacy heuristics (push#). When applied to evaluate privacy policy interfaces for laypeople, the push# heuristics enhances the downstream utility on the number of catastrophic problems discovered. We also created preliminary usable privacy guidelines (pug#) and modeled a new process for creating new usability criteria. In addition, this thesis also provides: recommendations for the usability of parental privacy controls; a systematic mapping of usability heuristics for privacy policy interfaces; models of heuristic evaluation for novice evaluators; a method to enhance the usability of privacy policies with card sorting analysis; overview of privacy expectations on user experience regarding connectedautonomous vehicles; gesturebased interface prototype to enhance privacy in health systems; a preliminary ontology for usability findings; and usability heuristics for mobile games and elderly players. We discuss how future studies could explore the employment of our heuristics and guidelines to specific domains, such as human-robot interaction and human-artificial intelligence interaction. Finally, we propose the study of nudging usability to enhance privacy protection regardless of privacyprotection tools.

Keywords: Privacy, Usability, Heuristic, Evaluation.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

Part I

Motivation and Background

CHAPTER

1

INTRODUCTION

1.1 Motivation

As information technology advances towards ubiquitousness, people are often unaware of what information other people, or organizations, have about them (ACQUISTI; BRANDIMARTE; LOEWENSTEIN, 2015; KRUMM, 2018). On the one hand, the coronavirus pandemic increased the opportunities to work, to get medical assistance, education, entertainment, and commerce activities, among other daily activities (VéLIZ, 2021). On the other hand, the multiplicity of connected devices tracking our personal data and posing challenges to our privacy also increased (CUNHA; MENDES; VILELA, 2021; CHAMIKARA *et al.*, 2021; MEHTA *et al.*, 2021).

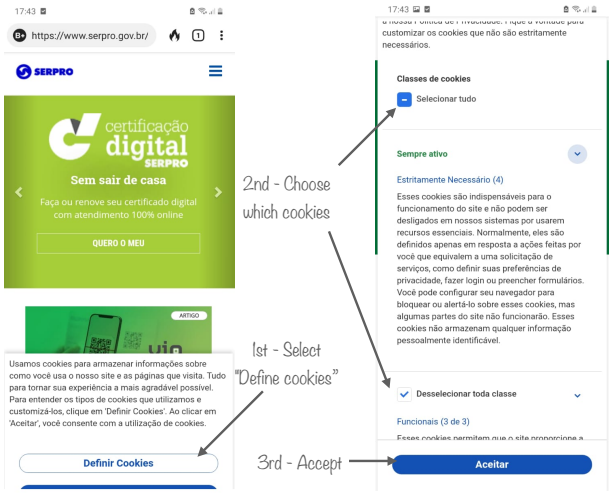
Different Governments have enforced privacy legislation to ensure their citizens control their data. Examples of such regulatory acts are but are not limited to: the European Union General Data Protection Act (GDPR), the California Consumer Privacy Act (CCPA), the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA), the Japanese Act on Protection of Personal Information (APPI) and the Brazilian General Personal Data Protection Law (LGPD)[1]. A central concern among these regulatory acts is to enforce natural persons and organizations to get appropriate authorization before handling others' data. In the information technology context, these acts imply that users must be aware and free to decide whether to share their personal information with other entities (SCHAUB; BALEBAKO; CRANOR, 2017).

Usually, information technology companies display privacy policies in their software applications to make users aware of how the company collects and uses their data (SCHAUB; BALEBAKO; CRANOR, 2017; AïMEUR; LAWANI; DALKIR, 2016). These policies may include privacy choices, settings that let users decide how their personal data is shared (SLEPCHUK; MILNE, 2020; HABIB *et al.*, 2020; GARFINKEL; LIPFORD, 2014). Figure 1 shows an

[1] In Portuguese: *Lei Geral de Proteção de Dados Pessoais*

example of privacy policies that include privacy choices.

Figure 1 – A screenshot from <serpro.gov.br> to exemplify privacy policies that include privacy choices.



Source: Elaborated by the author.

Empowering users with privacy choices inside privacy policies may bring companies closer to complying with legislative acts and also enhance users' trust in the technology (SLEPCHUK; MILNE, 2020). However, privacy policy interfaces are often complex and lack usability (DE; ZEZSCHWITZ, 2016; BERTINO, 2016; OATES *et al.*, 2018; PACI; SQUICCIARINI; ZANNONE, 2018; HABIB *et al.*, 2020). Without usable privacy policy interfaces, the human error vulnerability and its associated threats remain at high risk for companies' information security (MESZAROS; BUCHALCEVOVA, 2017). Besides, information technology companies need to design usable privacy policy interfaces in order to achieve transparency, as required by the different privacy regulations (HABIB *et al.*, 2020).

1.2 Thesis Statement

In this thesis, we sought to create usability heuristics to enhance the design and evaluation of privacy policy interfaces. Although many usability heuristics are in the literature (de Lima Salgado *et al.*, 2020; SALGADO; RODRIGUES; FORTES, 2016; HERMAWATI; LAWSON, 2016), those aiming at the domain of privacy policy interfaces for laypeople remains as a gap. This gap limits the interplay between usability evaluation and information security for privacy. Our goal was to fill out this gap.

Because a set of usability heuristics for the domain of laypeople and privacy policy interfaces remains as a gap in the literature, we can raise the following research questions:

– What are the main characteristics of usability issues that laypeople face interacting with privacy policy tools?

– How effective is the performance of usability heuristics composed of these characteristics?

Our first question has qualitative characteristics. For this reason, we did not speculate any hypotheses for these questions. However, the second question has an associative characteristic, which led us to speculate the following hypothesis:

Employing new usability heuristics focused on the domain of privacy policy interfaces for laypeople enhances the performance of heuristic evaluation in the domain.

The main objective of this thesis is to help designers create more usable privacy policy interfaces for laypeople. We sought to achieve this objective by providing designers with a new set of effective usability criteria, heuristics, to employ in evaluations in such domain. Alternatively, we can describe our objectives as follows:

- *To identify the main characteristics of usability issues that laypeople face when interacting with privacy policy interfaces.*
- *To create usability heuristics from the identified characteristics to better address problems that affect laypeople interaction with privacy policy tools.*
- *To validate the set of usability heuristics.*

After achieving the first three objectives, we expected to reach the following additional objectives:

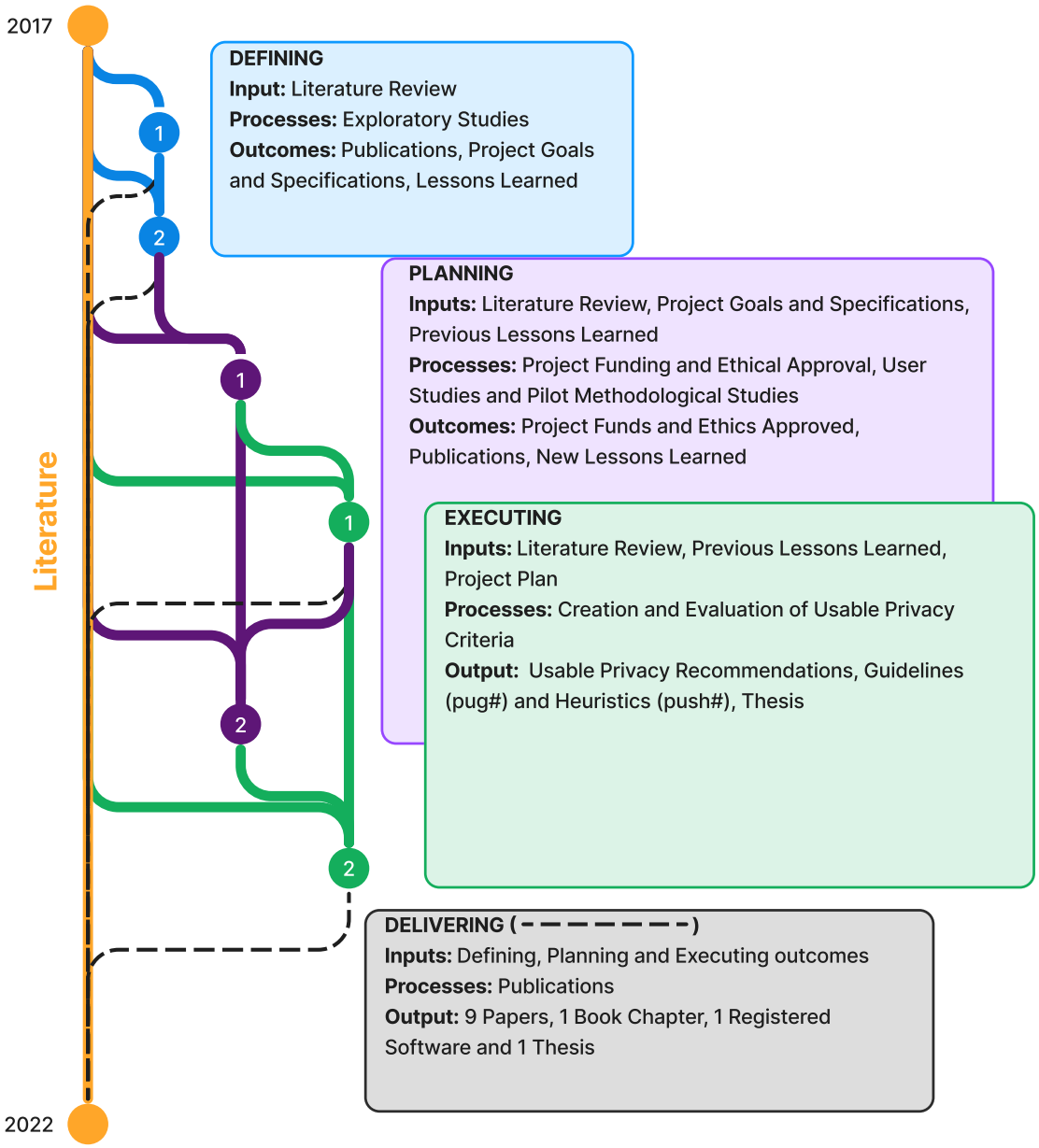
- *To adapt the heuristic evaluation method for information security and privacy professionals with no usability expertise.*
- *To create user-centered design recommendations for the domain of usable privacy.*
- *To understand and describe laypeople's mental models regarding privacy policy interfaces.*

1.3 Study Design and Contributions

We planned a cross-disciplinary method to seek the answers to our questions. To this end, we designed this thesis to employ Human-Computer Interaction (HCI) methods to enhance the usability of privacy policy interfaces. Overall, we designed the thesis among four major processes: *Defining*, *Planning*, *Executing* and *Delivering*. All of these processes receive its main inputs from the literature. Along with multiple comments received from reviewers when

executing the delivering (publication) process, we employed the literature to monitor and control the Thesis‘ project. Figure 2 illustrates the life cycle of this Thesis Design.

Figure 2 – Thesis Design.



Source: Elaborated by the author.

This doctoral research began with the exploratory process of *defining* its goals and specifications. To this end, we conducted initial exploratory studies by collaborating with our colleagues in studies about game and Web accessibility, usability inspection for novice evaluators and methods for comparing usability finding reports. We collaborated in a study to create a simulation tool to support the perception of accessibility issues faced by visually impaired users while browsing the Web (GOMES *et al.*, 2018). This study led us to discuss the importance of supporting software development teams with context information, such as those provided by the simulation tool, to enhance the quality of user interface evaluation outcomes. In parallel, we also conducted a mapping study that explored how accessibility and usability were approached during the digital game development process when designing games for elderly players (SANTOS; SALGADO; FORTES, 2018). The study showed that the interface evaluation process receives special attention during game development aiming for accessibility and usability. Due to the indications of usability evaluation importance from both studies, and the growing interest on privacy research, we decided to investigate usability newcomers' difficulties on understanding different terms that compose the usability, assuming the premise that such understanding is fundamental to perform usability evaluation with quality (SALGADO *et al.*, 2018). The study indicated that the context of use might be the most difficult term for novice professionals to understand when evaluating the usability of an interface. To this extent (as indicated by number 1 at Figure 2), we understood that the goals of this project should be placed on enhancing the evaluation stage of privacy software development. Nonetheless, one of the main challenges in assessing alternatives of usability evaluation refers to comparing the outcomes from such alternatives in order to determine the similarity of usability findings. For this reason, we sought to create a usable privacy findings ontology to evaluate the feasibility of employing in our prominent experiments (SALGADO *et al.*, 2019), as indicated by number 2 at Figure 2. The results indicated that the ontology could support the comparison, although traditional methodologies for comparing usability findings, as those indicated byHartson, Andre and Williges (2001), remain necessary. Together with the motivation presented at section 1.1, we defined our goals and the specifications for the methodology of our mainstream research, also described previously at section 1.2.

At a first iteration of the *planning* process (as indicated by number 1 at Figure 2), we applied for research funding and submitted our methods for. To this end, we used the lessons learned from previous publications, goals and specifications, and updated literature review from the defining process to compose the documents. We have approved the FAPESP doctoral research funding for this research. Meanwhile, we received the ethical committee approval for this doctoral research project [2]. These results led us to the first iteration of the executing stage.

The first iteration of the *executing* process (represented by number 1 at Figure 2) refers to mapping the state-of-the-art on usability heuristics that could be employed in heuristic evaluations

[2] This study was approved by a Research Ethics Committee with CAAE code 69353317.4.0000.5390

of privacy policy interfaces designed for laypeople, first round of experiments comparing different usability heuristics in the domain and pilot studies on creating usable privacy recommendations. The mapping study identified indications from the literature on what usability heuristics could best address usability problems that affect laypeople interaction with privacy policy interfaces. To confirm these findings, we also conducted a confirmatory empirical experiment comparing the state-of-the-art candidate against the traditional usability heuristics of Nielsen (de Lima Salgado *et al.*, 2020). At the end, the first iteration of the *executing* process led us to identify the state-of-the-art on usability heuristics that can be employed to evaluate the usability of privacy policy interfaces designed for people. Nevertheless, we still needed to create our own usable privacy criteria and, to this end, we had to pilot our methodology and investigate potential users characteristics that could be taken into account when creating the criteria. For this reason, we went for a second *planning* stage.

At the second iteration of the *planning* process, as indicated by number 2 at Figure 2, we aimed at understanding users' behavior in the privacy context and piloting the methods of the thesis. To evaluate the privacy context, we chose to investigate different cloud connected devices, as smart toys, Data Glove controlled interfaces and autonomous vehicles. This was done in order to help manage the scope of our project in terms of interactive technologies. To understand users' behavior with privacy protection interfaces, we performed studies as a card sorting experiment that sought to understand users' mental model on information architecture of privacy policy related terms in the context of smart toys. This study considered a smartphone application model of parental privacy control to protect children' privacy with smart toys (SALGADO *et al.*, 2019a). Also, this study resulted in a registered software for online card software, developed by the authors, and a new model of parental privacy control for the domain of smart toys available at <https://github.com/alsalgado/LGPDroid>. For this reason, we present the full text at Chapter 3. We also evaluated the usability of Data Glove interfaces employed in the design of multi-device applications for health treatments (DEMOE *et al.*, 2020). We evaluated the usability of the interface as a promising approach to design future private interactions, since customized hand movements could naturally act as privacy protection measures in human-computer communication. The lessons learned from this study show us promising ideas for future research on privacy by the design of Data Glove interactions, and led us to keep the scope of this thesis on smartphone interactions. In addition, we conducted a pilot study to evaluate preliminary tendencies of users' expectations about privacy on connected-autonomous vehicles (SALGADO *et al.*, 2020). The outcomes of this study showed us indications of the imminent importance of designing privacy controls for the domain of such vehicles. However, after discussing with our colleagues, we decided to not consider such technologies in the scope of this thesis due to feasibility analysis. Finally, still in the second iteration of the *planning* process, we pilot the usability criteria creation process. In Salgado *et al.* (2019b), we created a lean guide for startup practitioners to identify usability heuristics for the design of accessible mobile games for elderly players. The study employed techniques from the grounded theory to

compose such guide and also to create a model of user interactions in the investigated domain. Such techniques were later employed in this doctoral research to create the usability criteria.

After the second round of *planning* process, we went for the second and last *executing* process. This stage, represented by the last number 2 at Figure 2, was aimed at creating and evaluating new usable privacy criteria. Our first study in this process evaluated the interfaces of parental privacy control of different cloud connected toys (smart toys) and discussed usability enhancements that could be done in order to mitigate privacy risks involved (YANKSON; SALGADO; FORTES, 2021). The results are eleven preliminary recommendations to improve the usability in designs of parental privacy controls of smart toys. This study is presented in full text at Chapter 4. The remaining subprocesses aimed at creating and evaluating the new usability criteria. At a first moment, we performed a qualitative secondary analysis of the literature to retrieve a dataset of usability findings discovered by previous studies with potential users of privacy controls that could be conspired as laypeople (average user). This study led us to creating 19 usable privacy guidelines and six usable privacy heuristics. Chapter 6 describes the full study and its outcomes. Thereafter, we conducted an empirical evaluation study to assess the quality of the new six usable privacy heuristics, as described in subsection 6.3.2. This evaluation study tests the hypothesis and supports the conclusions of this Thesis. Overall, this doctoral research delivered nine papers and one book chapter to the literature as is. Besides, it also generated one registered software.

1.4 Organization

This thesis is organized among three parts: *motivation and background*, *preliminary works* and *new privacy and usability criteria*. This chapter introduced this thesis. Chapter 2 presents a review of the literature on usable privacy policy interfaces for laypeople. It also presents terms and definitions that apply to this thesis's context.

The next part of this thesis describes selected preliminary works co-authored by us, as discussed in the previous section. Chapter 3 employs techniques of information architecture evaluation and data science to understand laypeople's mental model about privacy policy interfaces. Meanwhile, Chapter 4 describes a security analysis and usability evaluation of smart toys and describes recommendations for further design of privacy policy interfaces.

The third part of this thesis brings our results and discussions, the creation of new usability criteria: preliminary guidelines and validated heuristics. That part is based on the activities established in the literature to create usability principles (QUIñONES; RUSU, 2017), as follows:

- to identify the state-of-the-art on usability heuristics for privacy policy interfaces designed for laypeople;

- to determine specific features of privacy policy interfaces designed for laypeople;
- to specify the new set of heuristics;
- to validate the new set of heuristics.

In the third part, Chapter 5 presents a systematic mapping study of the literature to identify the state-of-the-art usability heuristics for privacy policy interfaces. Chapter 6 describes a qualitative analysis and methodology to create preliminary usable privacy guidelines (**pug#**) and six usable privacy heuristics (**push#**). After, subsection 6.3.2 presents case studies that empirically evaluate the new usable privacy heuristics. Finally, we present and discuss the conclusions of this thesis in Chapter 7.

CHAPTER

2

BACKGROUND AND RELATED WORK

2.1 The Usability Concept

Although usability may appear as a simple word with an intuitive meaning, different understandings exist in the literature (HERTZUM, 2018; LEWIS, 2014). The word usability is formed by the root word “use” and the suffix “ability”. Such morphology may indicate a simple concept, but it is, instead, a sensitizing concept that still lacks specific empirical instances (HERTZUM, 2018). In the Human-Computer Interaction (HCI) field, an interface with acceptable usability is a usable interface. The adjective usable comes from the capability of being used, and used refers to being employed for a purpose[1]. Therefore, usability may be understood as the capability of an interface being used for a specific purpose. This section provides an overview of the scientific understanding of usability in the HCI literature over the years. On the one hand, this chapter does not aim to provide a definitive concept for usability. On the other hand, the usability concept remains discussed in the literature (HERTZUM, 2018; LEWIS, 2014).

Initial studies in the field described usability as a synonym of easy to use (DEMERS, 1981). This description is due to the importance of designing software that was easy to use by programmers. At the time, easy-to-use software was associated with increased productivity. Therefore, usability was associated with productivity. However, this definition is specific for software programmers. As other professionals (lay users) began to rely on software to perform their daily tasks, usability received another perspective and importance. In this sense, Mayhew (2012) describes usability as the state of being easy to learn or use for any user (including lay users). Mayhew (2012) also refers to usability as the state of an interface being intuitive for users to figure out how to use.

The concept of usability evolved, and authors began to consider other aspects to describe usability. Nielsen (2021) described usability based on the five (5) quality components: *learn-*

[1] <http://www.dictionary.com/browse/usability?s=t>

ability, *efficiency*, *memorability*, *errors* and *satisfaction*. We describe Nielsen's definition for usability because he is one of the authors of the Heuristic Evaluation method, which is the focus of this project. According to Nielsen (2021), these components can be understood as:

> **Learnability:** "*How easy is it for users to accomplish basic tasks the first time they encounter the design?*"
>
> **Efficiency:** "*Once users have learned the design, how quickly can they perform tasks?*"
>
> **Memorability:** "*When users return to the design after a period of not using it, how easily can they reestablish proficiency?*"
>
> **Errors:** "*How many errors do users make, how severe are these errors, and how easily can they recover from the errors?*"
>
> **Satisfaction:** "*How pleasant is it to use the design?*"

Preece, Sharp and Rogers (2015) is among the most popular HCI book in the literature. According to them, usability can be defined in regards to five (5) goals: *effectiveness*, *efficiency*, *safety*, *utility*, *learnability* and *memorability*. Overall, these goals refers to optimizing user interaction with interfaces. According to the authors, these goals are as follows:

- **effectiveness** refers to how good the interface is at doing its purpose and helping users to achieve their goal;
- **efficiency** refers to the way that an interface support users to perform their task, through a "*minimal number of steps*";
- **safety** refers to the interface protecting users from unwanted situations;
- **utility** refers to the interface doing what the it is supposed to do (what users want);
- **learnability** refers to the interface being easy to learn how to use;
- and **memorability** refers to how easy users can remember how to use the interface.

In the privacy and security domain, the software may be considered usable if its users (WHITTEN; TYGAR, 1999, p. 2):

1. are reliably made aware of the security tasks they need to perform;
2. are able to figure out how to successfully perform those tasks;
3. don't make dangerous errors; and
4. are sufficiently comfortable with the interface to continue using it.

Series of international standards (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2010) recognized usability as an attribute of quality and ergonomics of software. These standards describe usability supported by the terms: *user*, *goal*, *effectiveness*, *efficiency*, *satisfaction*, *context of use* and *task*. The description of usability according to these standards is:

> "*the extent to which a system, product or service can be used by specified* ***users*** *to achieve specified* ***goals*** *with* ***effectiveness***, ***efficiency*** *and* ***satisfaction*** *in a specified* ***context of use***"

In addition to the usability description presented, the respective terms that serve as support for such a description are (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a):

User: "*person who interacts with a system, product or service*".

Goal: "*intended outcome*".

Effectiveness: "*accuracy and completeness with which users achieve specified goals*".

Efficiency: "*resources expended in relation to the accuracy and completeness with which users achieve goals*".

Satisfaction: "*freedom from discomfort, and positive attitudes towards the use of the product*".

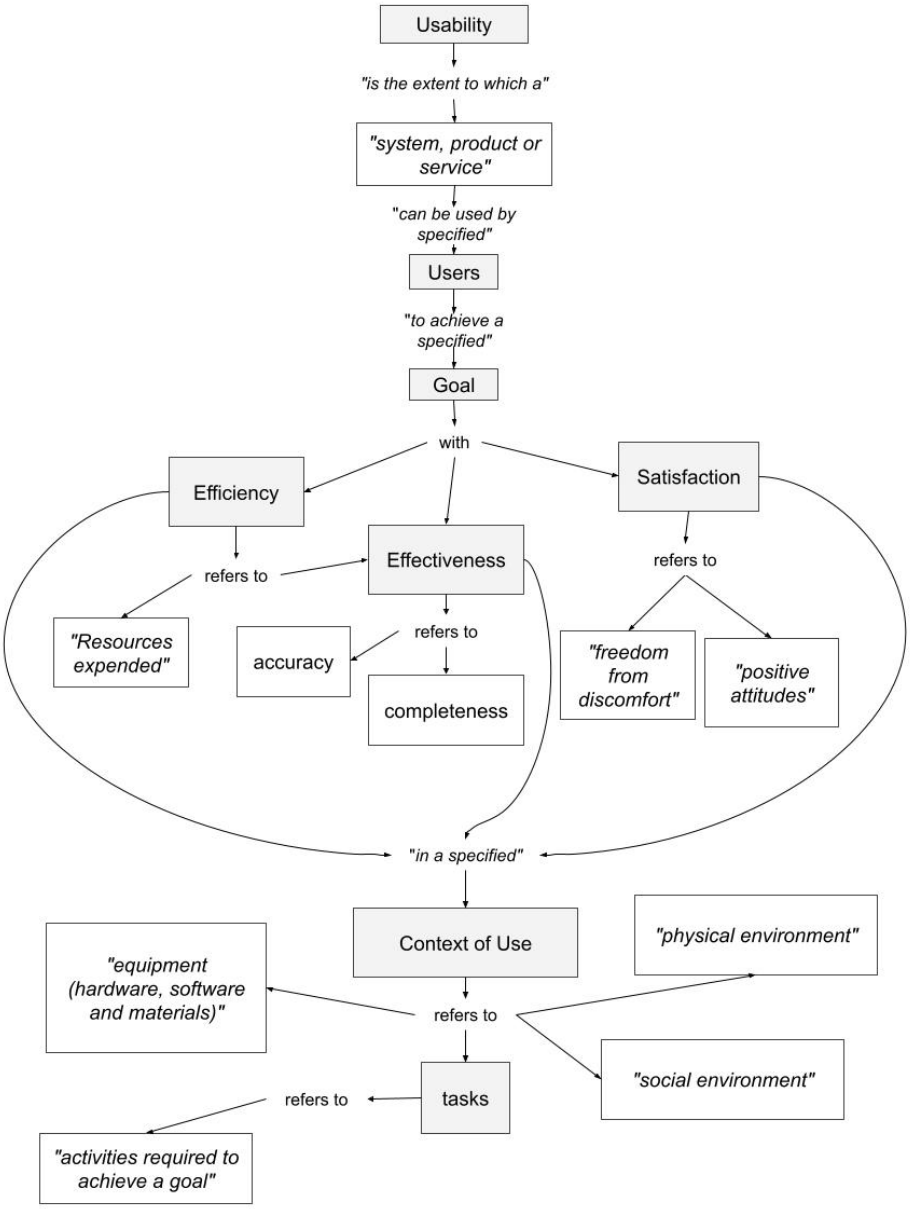
Context of use: "*users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used*".

Task: "*activities required to achieve a goal*".

Figure 3 illustrates the ISO's definition for usability based on its components. Although the ISO's definition is not universally adopted, it represents an international community from the HCI area. In addition, the ISO standards are not restricted to the HCI community; other standards are essential for other areas of computer science as well. As an example, the INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) and INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC) (2011) presents contents for the information privacy area. For this reason, this study adopts the ISO's description of usability.

According to Lewis (2014), usability can be described in respect to two major concepts: *summative usability* and *formative usability*. He argues that summative and formative usability differences are substantial and that a unique definition could not cover both concepts. According to him, summative usability focus on measurements and metrics for usability. Meanwhile, formative usability focuses on diagnosing usability problems and potential solutions.

Figure 3 – Relation among terms used in usability definition.



Source: Salgado *et al.* (2018).

In summary, different definitions of usability may apply for different research purposes. Therefore, researchers should consider (discuss) the pros and cons of each definition to decide which one better applies for their context (HERTZUM, 2018). The focus of this study is to develop a formative usability evaluation approach. For this reason, the following section presents a review of approaches for formative usability evaluation.

2.1.1 Formative Usability Evaluation

The HCI literature shows different approaches to formative usability evaluation. Although the employment of such approaches usually produces different outcomes (LEWIS, 2014), they

are often called evaluation methods. This terminology remains a topic of discussion in the field. Nevertheless, related studies usually refer to usability evaluation methods. For this reason, we also refer to it as a method.

The usability evaluation process should not occur only once during the design of interfaces. On the contrary, it should be done over cycles. Different cycles may require distinct alternatives of evaluation. As an example, initial versions of a prototype may not be adequate to be used by potential users. In these situations, methods that rely on usability specialists may be more appropriate. On the other hand, it is recommended to have end users testing an interface before it is first released. Because of these alternatives, we usually organize usability evaluations between *user-based* and *inspection-based* evaluations (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a; PREECE; SHARP; ROGERS, 2015).

The INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (2016a) classifies usability evaluation as *user-based* or *inspection-based* evaluation. According to it, the user-based type requires the participation of a sample of end-users. On the other hand, the inspection-based requires the judgment of inspectors (typically usability specialists) respecting a determined criterion (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a).

Typically, formative usability evaluation produces a list of diagnosed usability findings (LEWIS, 2014; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). Mitigating the effect of these findings improves the usability of the interface. Usability findings are "*identified usability defect and/or usability problem or positive usability-related attribute*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). Usability defect is a "*product attribute(s) that lead(s) to a mismatch between user intentions and/or user actions and the system attributes and behaviour*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). Usability problem is a "*situation during use resulting in poor effectiveness, efficiency or satisfaction*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). A positive usability-related attribute is the quality of the interface, which has a positive effect on it and requires no solution.

User-based evaluations usually support their diagnostic on observations of real users' interactions. Meanwhile, inspection-based evaluations must refer to a determined criterion to support its diagnostic (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a).

The concepts of user-based and inspection-based evaluations help categorize usability evaluation methods. We can distinguish most of the methods between these two groups. Noteworthy, the Cognitive Walkthrough (CW) can be classified between these groups (user-based and inspection-based). This is because potential users can conduct the CW or usability inspectors (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). For this reason, we discuss the CW in a particular section in this study. We present the other methods according

to their user-based and inspection-based types.

Cognitive Walkthrough

Lewis *et al.* (1990) proposed the first version of CW. Later, different studies proposed consecutive reviews for the CW (POLSON *et al.*, 1992; MAHATODY; SAGAR; KOLSKI, 2010). The CW is centered on four (4) questions related to users' cognitive behavior (MAHATODY; SAGAR; KOLSKI, 2010):

Question 1 - "*Will the user try to achieve the right effect?*"
– This question refers to what users may be thinking when the action begins.

Question 2 - "*Will the user notice that the correct action is available?*"
– This question refers to whether users would be able to locale the command.

Question 3 - "*Will the user associate the correct action with the effect that user is trying to achieve?*"
– This question refers to whether users would be able identify the specific command.

Question 4 - "*If the correct action is performed, will the user see that progress is being made toward solution of the task?*"
– This question refers to users ability to understand the possible given feedback.

To answer these questions, the CW may count on the participation of end-users, usability inspectors, or other professionals. For CW, based on the participation of inspectors or other professionals, they must play the roles of users interacting with a specific interface. For this reason, the organizers must explain to them the users' profile and context of use. Users, inspectors, or others must answer the four questions. Moreover, they need to answer these questions respecting pre-defined tasks, and goals (PREECE; SHARP; ROGERS, 2015; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a).

The CW questions are broad enough to be applied to different domains. Nevertheless, the CW is focused on evaluating the ease of learning (WHARTON *et al.*, 1994), which is only a part of the usability concept (NIELSEN, 2021; PREECE; SHARP; ROGERS, 2015). According to Mahatody, Sagar and Kolski (2010), many variants of the CW method were proposed in the literature. Although, it was out of the scope of this study to describe all of them.

User-based Evaluations

User-based evaluations are often referred to as user testing, or usability testing, because it imposes a situation of test to the interface (PREECE; SHARP; ROGERS, 2015; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). These evaluations may

depend on observation (e.g., laboratory test and ethnography) or on the opinion of users (e.g., questionnaires and interviews) (PREECE; SHARP; ROGERS, 2015; KRUMM, 2016).

Observations are usually conducted in laboratory (PREECE; SHARP; ROGERS, 2015), although some occur in-the-wild (KRUMM, 2016; LINDGAARD, 2015). In regards to laboratory tests, observers may adopt the Think-Aloud technique to obtain important indications of users' cognitive processes. The Think-Aloud technique asks users to verbalize their thoughts, those related to their interaction decisions (ERICSSON; SIMON, 1980; PREECE; SHARP; ROGERS, 2015). Practitioners can record (e.g. audio, video, eye-tracking) users' interactions to post-analyze it (PREECE; SHARP; ROGERS, 2015). These thoughts may indicate their cognitive process and maybe coded by observers according to frameworks as Norman's stages of action (NORMAN, 2013).

Norman's stage of action is a representation of the human cognitive process. It represents seven (7) stages: *goal*, *plan*, *specify*, *perform*, *perceive*, *interpret* and *compare*. These stages represent a cycle. This action cycle can start from the goal, a goal-driven behavior. After deciding on a goal, the user starts to plan its interaction. During the planning stage, users decide which plan to conduct. After deciding the plan, users still have to specify (specify stage) the actions needed to complete such a plan and execute it (perform stage). At this point, the interface will receive users' actions and change or not its status. Whatever the response from the interface (change or not), users will perceive it (perceive stage). Later, users will try to make sense of what just happened (interpret stage). Finally, users compare the interpreted state with their goal (compare stage). The action cycle can also begin after users perceive any change in the interface (from the perceive stage), which may be not a result of their actions (NORMAN, 2013, p. 40-41).

Besides the action framework of Norman (NORMAN, 2013), the literature presents other frameworks. These frameworks are focused on structuring the report of usability findings (YUSOP; GRUNDY; VASA, 2017). The User Action Framework (UAF) frames usability findings supported by Norman's framework. It types usability findings among planning, translation, physical actions, outcome, and assessment categories. As Norman's stages, these categories are also cyclical, according to this respective sequence (VILBERGSDOTTIR; HVANNBERG; LAW, 2014; YUSOP; GRUNDY; VASA, 2017; HARTSON; PYLA, 2012). The literature also presents the Classification of Usability Problems (CUP), the Root Cause Defect Analysis (RCA), the Orthogonal Defect Classification (ODC), and the Usability-Error Ontology (UEO). The CUP, RCA, and ODC aim to provide feedback for developers to help them correct usability findings. These frameworks are domain-free and provide information as the trigger (or root cause) of a usability finding. On the other hand, the UEO domain is specific. The UEO is focused on health systems. It was created after a survey with professionals of such domain (ELKIN *et al.*, 2013). Although these frameworks exist, the UAF has the most recent relevant case studies (VILBERGSDOTTIR; HVANNBERG; LAW, 2014; YUSOP; GRUNDY; VASA, 2017).

It is essential to notice that laboratory-based tests may provide limited outcomes because

it is difficult to simulate the security context of the real-world (JAFERIAN *et al.*, 2014). This remains a limitation of user-based evaluations in the domain of information security and privacy. This fact supports the importance of developing appropriate methods of formative usability inspection for this domain.

Inspection-based Evaluations

Inspection-based evaluations are based on the judgment (inspection) of inspectors. These evaluators are typically usability specialists. They are called to inspect the interface respecting determined criteria. Criteria are necessary to support evaluators' decisions. Popular criteria are guidelines, and heuristics (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). Both guidelines and heuristics are usability principles, but they vary on the degree of specification and quantity.

Guidelines are usually numerous, indicating specific aspects to avoid specific usability issues. On the other hand, heuristics are broad usability principles, rules of thumb and rely on the knowledge of inspectors to diagnostic specific usability findings (Jakob Nielsen, 2018; PREECE; SHARP; ROGERS, 2015). These usability findings can also be reported following the frameworks described in section 2.1.1.

Guidelines Review

During Guidelines Review, evaluators compare the interface's elements with a set of interface guidelines. Guidelines are large sets of usability principles, "*usually 10-200*" (LAZAR; FENG; HOCHHEISER, 2017, p. 269). The Web Content Accessibility Guidelines (WCAG 2.0)[2] is a popular set and focus on usability for users with the widest range of characteristics and capabilities (accessibility) (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a).

The WCAG 2.0 groups its guidelines among the principles: *perceivable*, *operable*, *understandable* and *robust*. The perceivable principle cautions that users must perceive any interface information somehow. The operable principle cautions that any feature must be available through keyboard access. The understandable principle cautions that the content must be readable and understandable by users. Finally, the robust principle cautions that user agents must have machine-readable content, including Assistive Technology.

Each of the WCAG 2.0 principles has a set of guidelines. These guidelines are considered essential to implement the respective principle on the Web. Each guideline has respective success

[2] Retrieved from Web Content Accessibility Guidelines (WCAG) 2.0 Website at <www.w3.org/TR/WCAG20/#guidelines>

criteria, with levels ranging among A, AA, and AAA (further information about WCAG 2.0 success criteria can be found at the WCAG portal[3]).

Reviewing large sets of guidelines may be time-consuming for human evaluators. For this reason, the literature presents automatized tools. These tools implement most of the work of a human evaluator during a guideline review by checking accordances to guidelines. The **Web Accessibility Checker (achecker)**[4] and the **TAW tool**[5] are examples of automatized tools for the WCAG 2.0 guidelines.

Heuristic Evaluation

Heuristic Evaluation (HE) was proposed by Nielsen and Molich (1990a). It involves multiple inspectors that compare the interface against a list of usability heuristics. Although similar to interface guidelines, heuristics are usually small sets of broad described usability principles, about ten heuristics. Therefore, HE takes less time than guidelines review (LAZAR; FENG; HOCHHEISER, 2017). The ten heuristics of Nielsen are traditionally employed in HEs. Nevertheless, Nielsen (1994) argues that domain-specific heuristics may benefit the method. Specific domain principles have been developed to increase the chances that usability of such domains was not overlooked (HERMAWATI; LAWSON, 2016; SALGADO; RODRIGUES; FORTES, 2016; SALGADO; FREIRE, 2014). Chapter 5 maps and discusses usability heuristics for the privacy and security domain. The ten heuristics of Nielsen are entitled as follows (Jakob Nielsen, 2018):

(1.) Visibility of system status.

(2.) Match between system and the real world.

(3.) User control and freedom.

(4.) Consistency and standards.

(5.) Error prevention.

(6.) Recognition rather than recall.

(7.) Flexibility and efficiency of use.

(8.) Aesthetic and minimalist design.

(9.) Help users recognize, diagnose, and recover from errors.

(10.) Help and documentation.

Nielsen (1994) ordered his heuristics according to their explanatory potential. The heuristic's sequence respects the probability of finding a disagreement with a respective heuristic.

[3] WCAG 2.0 success criteria Webpage: <www.w3.org/TR/UNDERSTANDING-WCAG20/conformance.html#uc-levels-head>

[4] <achecker.ca/checker/index.php>

[5] <www.tawdis.net/ingles.html?lang=en>

According to Nielsen (1994), evaluators would find more usability issues related to "*Visibility of system status*" than to "*Help and documentation*".

Regarding the process of HE, Nielsen (2018a) argues that the evaluators decide on their own how to conduct the heuristic evaluation. Although, he suggests an initial pass through the interface to "*get a feel for the flow of the interaction and the general scope of the system*". The second pass would be to diagnose and list the usability findings. Alternatively, Preece, Sharp and Rogers (2015) describe it among three stages: *briefing session*, *evaluation period* and *debriefing session*. During the briefing, inspectors receive all guiding information about how the evaluation should be conducted. During the evaluation, evaluators individually inspect the interface looking for possible violations of any heuristics considered. Finally, at the debriefing, evaluators discuss and review their findings with each other to prepare a final report of usability findings. At this stage, they may also describe potential solutions for each finding. In addition, inspectors might rate a severity for each usability finding. Noteworthy, different inspectors may rate different using the same scale (SAURO, 2014). This overall process is illustrated in Figure 4.

As indicated by Preece, Sharp and Rogers (2015), evaluators must rate a severity level for usability findings at the *debriefing session*. For this subject, different usability severity rating scales are proposed (SAURO, 2018). Nielsen (2018b) suggests the following one:

0 - Not a usability problem: it is not a usability problem at all.

1 - Cosmetic problem: it is only a cosmetic problem. Its correction may be made only if extra time is available in the project timeline.

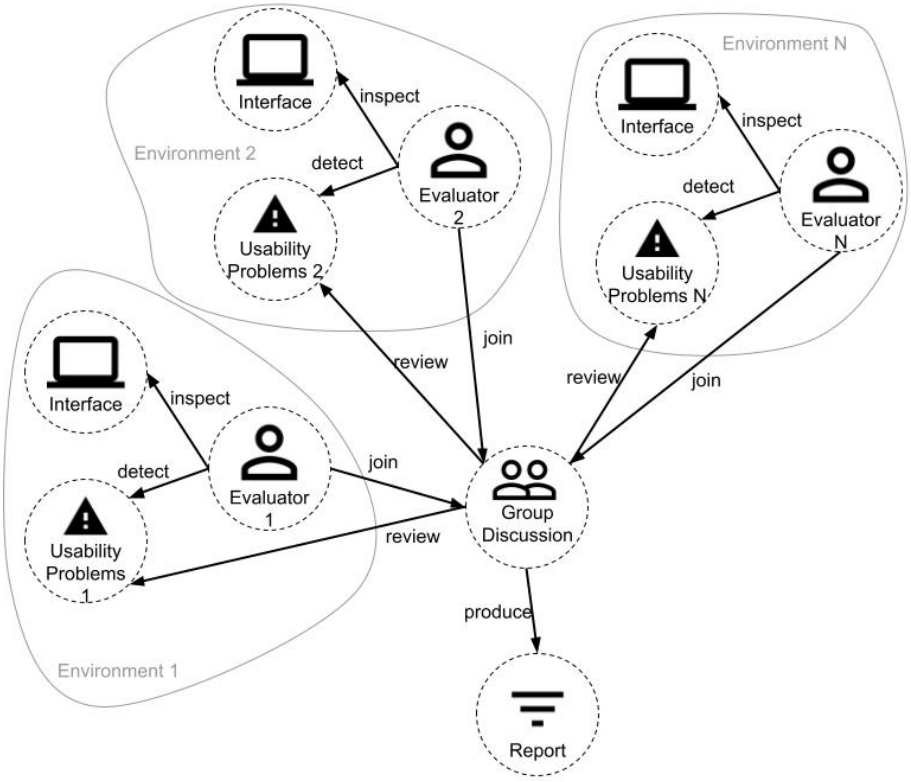
2 - Minor problem: the correction of this kind of problem may receive low priority.

3 - Major problem: the correction of this kind of problem may receive a high priority.

4 - Usability catastrophe: these must be the first problems to be corrected. They must be corrected before the product be released.

To be effective, HE must be conducted by usability experts that are familiar with the heuristics (LAZAR; FENG; HOCHHEISER, 2017). According to Preece, Sharp and Rogers (2015), the number of experts needed for a HE is the key question. Nielsen (2018a) recommend three to five experts to conduct a heuristic evaluation. Because counting on multiple usability experts might be expensive, this number would have a great cost-benefit (revealing about 75% of usability findings). Nevertheless, this cost-benefit analysis remains under discussion in the literature (PREECE; SHARP; ROGERS, 2015; BORSCI *et al.*, 2013). Yet, five inspectors may be considered as a standard in the HCI field (CAINE, 2016).

Figure 4 – Traditional Heuristic Evaluation process.



2.2 Privacy and Information Security

Although privacy "*has emerged as a new critical requirement*" for information security (BERTINO, 2016, p. 401), it is not a new concept. The literature shows evidences of a "*privacy-seeking behavior*" from ancient cultures as Greece and Rome (ACQUISTI; BRANDIMARTE; LOEWENSTEIN, 2015). Referring to information technology, Warren and Brandeis (1890) discussed the right to privacy and its relations with information technologies, represented by photographs and news paper at the time. Warren and Brandeis (1890) argued that privacy should be considered as (WARREN; BRANDEIS, 1890, p. 205):

> ...protection afforded to thoughts, sentiments, and emotions, expressed through the medium of writing or of the arts, so far as it consists in preventing publication, is merely an instance of the enforcement of the more general right of the individual to be let alone.

Since current Web technologies may be considered "*medium of writing*", the definition presented by Warren and Brandeis (1890) remains applicable. However, as discussed by Acquisti, Taylor and Wagman (2016), the literature does not provide a unified theory for privacy because it may have different meanings for different people. Oates *et al.* (2018) also argues that the meaning of privacy may differ according to different peoples' backgrounds because different contexts may have different benefits when choosing between private and public.

The term *privacy* can refer to: "*territorial privacy*", "*privacy of a person*" or "*information privacy*" (KOKOLAKIS, 2017, p. 123). In this thesis, we are interested on *information privacy*, which refers to control over data sharing (KOKOLAKIS, 2017; ACQUISTI; TAYLOR; WAGMAN, 2016). In this context, the ISO/IEC TR 20547 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2020) defines privacy as the "*right of individuals to control or influence what information related to them may be collected and stored and by whom that information may be disclosed*".

As information technology becomes ubiquitous, cloud-connected applications are capable of collecting diverse personal data from their users due to a variate of available sensors (CUNHA; MENDES; VILELA, 2021; CHAMIKARA *et al.*, 2021; MEHTA *et al.*, 2021; KRUMM, 2018). This raised many privacy concerns. As a consequence, in 2016, the European Parliament and The Council of The European Union adopted the General Data Protection Regulation (GDPR) to protect their citizens in regards of the processing of their data (THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, 2016). Besides GDPR, different Governments have enforced privacy legislation to ensure their citizens control their data. Examples of such regulatory acts are, but are not limited to: the California Consumer Privacy Act (CCPA), the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA), the Japanese Act on Protection of Personal Information (APPI), and the Brazilian General Personal Data Protection Law (LGPD)[6]. A central concern among these regulatory acts is to enforce natural persons and organizations to get appropriate authorization before handling others' data. In the information technology context, these acts imply that users must be aware and free to decide whether to share their personal information with other entities (SCHAUB; BALEBAKO; CRANOR, 2017). Therefore, privacy became a critical requirement for information security (BERTINO, 2016), the field of study that refers to " the protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability" [7].

The GDPR guidelines on transparency indicate how to make an online application transparent for users regarding data protection. Among other GDPR principles, transparency refers to how information technology companies communicate their users about their rights under the regulation and how such companies facilitate, for its users, the exercise of their

[6] In Portuguese: *Lei Geral de Proteção de Dados Pessoais*
[7] <https://csrc.nist.gov/glossary/term/information_security>

privacy choices (The European Parliament and The Council of The European Union, 2017). To comply with GDPR and similar regulatory acts, information technology companies display privacy policies at their software applications to let users aware about how the company collects and uses their personal data (SCHAUB; BALEBAKO; CRANOR, 2017; AïMEUR; LAWANI; DALKIR, 2016). These policies may include privacy choices (see Figure 1), letting users decide how their personal data is shared (SLEPCHUK; MILNE, 2020; HABIB *et al.*, 2020; GARFINKEL; LIPFORD, 2014), but privacy policy interface are complex and lack usability (DE; ZEZSCHWITZ, 2016; BERTINO, 2016; OATES *et al.*, 2018; PACI; SQUICCIARINI; ZANNONE, 2018; HABIB *et al.*, 2020). Yet, humans are the weakest link in information security, and a survey on the United Kingdom and the United States indicated that human error is responsible for 88% of data breaches (TESSIAN; HANCOCK, 2020). The field of usable privacy and security aims to solve this issue by enhancing the usability of systems that help users to manage privacy, and security (DE; ZEZSCHWITZ, 2016; GARFINKEL; LIPFORD, 2014).

2.3 Usable Privacy and Security

UPS is the research field aimed to study the usability of systems that help end-users or administrators to manage security, and privacy (DE; ZEZSCHWITZ, 2016; GARFINKEL; LIPFORD, 2014). The interest in UPS field had a rapid development during the past two decades (STILL, 2016; DE; ZEZSCHWITZ, 2016; GARFINKEL; LIPFORD, 2014; CRANOR; BUCHLER, 2014). First, usability and security were only seen as antagonistic and users seen the greatest risk to information security (STILL, 2016). Later, because the range of potential threats increased due to the pervasiveness of data (BERTINO, 2016), laypeople were often required to make security decisions (JANG-JACCARD; NEPAL, 2014) and seen as the "*greatest hope*" for the area (STILL, 2016). Without usable tools, even experts will misconfigure and leave vulnerabilities (SASSE; SMITH, 2016). In such cases, security breaches may be attributed to designers rather than laypeople (WASH; ZURKO, 2017). Nevertheless, even fundamental concepts of laypeople's interaction with privacy tools, as mental models (OATES *et al.*, 2018), remain rare in the literature. The UPS field still needs usable tools for laypeople (BERTINO, 2016); also, it still needs appropriate usability methods for this domain.

Examples of major themes in UPS are but are not limited to: social media privacy, user authentication, anti-phishing efforts, and Web privacy and fair information practice. This research project focus on Web privacy and fair information practice and challenges related to usable privacy policy tools for laypeople. The Web privacy and fair information practice were motivated by the increased opportunities for data collection through the Web. Online stores had unprecedented opportunities to collect and analyze data about their consumers (GARFINKEL; LIPFORD, 2014). For this reason, privacy regulations implied that users must be aware and free to decide whether to share their personal information with other entities (SCHAUB; BALEBAKO; CRANOR, 2017). In this sense, the U.S. Government enforced companies to protect users'

privacy through notice and choice (GARFINKEL; LIPFORD, 2014). Similarly, the General Data Protection Regulation enforced companies in the European Union to provide people with control over their data sharing on the Web (THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, 2016). The Brazilian Government also protects users' privacy on the Web [8]. The following section discusses the usability of privacy policy tools.

Similar to usability, information security also has strong ties with industry standards. Although defining usability remains challenging the literature, defining information security often relates it to information confidentiality, availability, and integrity. As technology evolves, new characteristics may extend these three properties to address the process of information security better (von Solms; van Niekerk, 2013). Thus, it is reasonable to understand the term usable security as making the information security process usable for its users. However, this does not represent the entire field of usable privacy and security.

If usable security is to design usable security processes, we should assume that users' goal is to secure the system. But, "security is usually a secondary goal" for laypeople (WHITTEN; TYGAR, 1999). The unmotivated user property, as defined byWhitten and Tygar (1999), states that:

> Security is usually a secondary goal. People do not generally sit down at their computers wanting to manage their security; rather, they want to send an email, browse web pages, or download software, and they want security in place to protect them while they do those things. It is easy for people to put off learning about security, or to optimistically assume that their security is working, while they focus on their primary goals. Designers of user interfaces for security should not assume that users will be motivated to read manuals or to go looking for security controls that are designed to be unobtrusive. Furthermore, if security is too difficult or annoying, users may give up on it altogether.

For this thesis, we investigate the usability of privacy policy interfaces designed for laypeople to protect the privacy of their children, which is called parental privacy control. Controlling the privacy of their beloved ones is a premise that we assume is motivational for laypeople to use privacy policy interfaces.

2.4 Usability of Privacy Policy Tools

Privacy policy tools are usually composed of interfaces and mechanisms of policy generation, comprehension, configuration (also known as privacy choice), and feedback generation. These mechanisms encompass the main features of privacy policy tools (PACI; SQUICCIARINI; ZANNONE, 2018). According to Paci, Squicciarini and Zannone (2018), most of the

[8] Federal Law Number 12.965 (2014): <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2014/Lei/L12965.htm>

studies about policy generation are focused on mechanisms for automated, or semi-automated, policy generation. They also describe that studies about policy comprehension usually propose new interface designs. This is the same strategy adopted by studies that propose advances on interfaces for policy comprehension. Nevertheless, studies about feedback generation focus on providing feedback about access decision making (PACI; SQUICCIARINI; ZANNONE, 2018). Our focus is on studies about privacy policy comprehension and configuration/choice because, in the literature, these studies are the majority in proposing new interface designs for the domain. In Chapter 6, we review and discuss interfaces proposed by studies in the literature aiming at privacy policy comprehension and configuration. Based on Paci, Squicciarini and Zannone (2018, Table 5), Table 1 summarizes the interfaces identified in the literature, its mechanisms, and the respective study that proposed it.

Table 1 – Overview of interface designs for privacy policy comprehension and configuration.

Mechanism	Interface	Study
Policy Comprehension	Improptu	Rode *et al.* (2006)
	Expandable Grid Windows XP	Reeder *et al.* (2008)
	Expandable Grid P3P	Reeder (2008)
	Nutrition Label	Kelley *et al.* (2009)
	The Eyes Metaphor	Schlegel, Kapadia and Lee (2011)
	The Reflective Policy Assessment (RPA)	Anwar and Fong (2012)
Policy Configuration	Pviz	Mazzia, LeFevre and Adar (2012)
	VeilMe	Wang *et al.* (2015)
	Retinue	Hu, Ahn and Jorgensen (2011)
	Mcontroller	Hu, Ahn and Jorgensen (2013)
	Sigma	Hu, Ahn and Jorgensen (2012)

Source: Elaborated by the author.

2.5 Parental Control Tools

In regards to the privacy of children's data, parents are usually allowed to control data privacy through parental control tools (MCREYNOLDS *et al.*, 2017a; RAFFERTY; FANTINATO; HUNG, 2015). Children lack knowledge about the risks involved with their data privacy (CUNHA; Unicef; others, 2017). Parental controls are a type of privacy policy tool aiming at parents of children not yet prepared to decide by themselves about their data privacy. The UNICEF[9] reinforces, then, that privacy policy must help children and their parents to understand children's data privacy (CUNHA; Unicef; others, 2017, p. 17). Rafferty, Fantinato and Hung (2015) proposed a design of parental control for the domain of smart toys, which allows parents to create and manage privacy rules regarding their children's data. Hung, Tang and Kanev (2017, p. 1) defines smart toys as:

[9] United Nations Children's Fund

> ... a device consisting of a physical toy component that connects to one or more toy computing services to facilitate gameplay in the Cloud through networking and sensory technologies to enhance the functionality of a traditional toy.

Because usability is a challenge for any privacy policy tool, it is a challenge for parental controls that implements privacy policy tools. In a previous study (SALGADO *et al.*, 2017), our research group reviewed usability methods that could enhance the usability of parental controls. With such review, we argued that appropriate usability evaluation methods are still needed to form usability of parental controls. Therefore, we understand that this topic may benefit from researches on the usability of privacy policy tools and vice versa. For this reason, the experiments of this project may be conducted on this topic.

2.6 Final Remarks

In this chapter, we reviewed the main terminologies involved in this study. Also, we reviewed studies that proposed designs of usable privacy policy interfaces. In addition, we analyzed their contributions and clustered results from formative usability evaluations into six themes of user interaction with privacy policy tools: *Fragile Sharing*, *Careful Sharing*, *Checkup*, *Audience Exposure*, *Fast Interaction*, *Fast Interaction* and *Frequent Interaction*. Finally, we discussed the recent topic of usable parental controls, privacy control.

Part II

Selected Preliminary Works

CHAPTER

3

SMART TOYS AND CHILDREN'S PRIVACY: USABLE PRIVACY POLICY INSIGHTS FROM A CARD SORTING EXPERIMENT

This chapter includes the peer-reviewed complete conference paper:

SALGADO, A. de L.; DIAS, F. S.; MATTOS, J. P. R.; FORTES, R. P. de M.; HUNG, P. C. K. Smart toys and children's privacy: usable privacy policy insights from a card sorting experiment. In: **Proceedings of the 37th ACM International Conference on the Design of Communication**. Portland Oregon: ACM, 2019. p. 1–8. ISBN 978-1-4503-6790-5. Available: <https://dl.acm.org/doi/10.1145/3328020.3353951>.

3.1 Abstract

Smart toys are new to the Internet of Things market, and its connectivity to the cloud have raised concerns about children's privacy. Parents and legal guardians have striven to protect the privacy of their owns. However, current approaches for privacy control still lack usability for lay people. In this paper, we have explored the use of Card Sorting to enhance the usability of a privacy control for smart toys. Our goal was to identify and describe benefits of this technique to the design of more usable privacy controls. For this reason, we conducted a case study with voluntarily participants. We chose a parental control model from the literature to be the subject of evaluation for the experiment. Therefore, we extracted 19 units of information from its interface, and put them into cards for the Card Sorting evaluation. After the experiment, we obtained 30 valid responses. From these responses we performed a cluster analysis to understand the best alternative to group privacy related contents. Our contributions include a new model for nutrition

label style mobile parental privacy controls for smart toys, suggestion of Google Material Design icons to be applied as indication for groups of privacy policies and, finally, a six steps process to perform Card Sorting with cluster analysis that does not rely on users' discussions to compose the Information Architecture hierarchy.

3.2 Introduction

Usable privacy and security is an emergent cross-disciplinary field, aimed to study usability of systems that help users to secure their sensitive information (GARFINKEL; LIPFORD, 2014). Once, usability and security were only seen as antagonistic, but the range of privacy threats has increased, and users required to make security decisions about their data (STILL, 2016; JANG-JACCARD; NEPAL, 2014; ALJOHANI; BLUSTEIN; HAWKEY, 2017). Governments have also required that companies enable their citizens to be "aware of" or to control their privacy. For example, the General Data Protection Regulation (GDPR) enforced companies in the European Union to enable people to control their data sharing[1]. Usability has been seen as a bridge towards a more secure Internet of Things (IoT).

Smart toys are new to the IoT market and stands as a promising pedagogical approach. Hung, Tang and Kanev (2017) define smart toy as:

> " ... a device consisting of a physical toy component that connects to one or more toy computing services to facilitate gameplay in the Cloud through networking and sensory technologies to enhance the functionality of a traditional toy."

Because of smart toy's connectivity to the cloud, which has raised concerns about children's privacy, parents have striven to protect the privacy of their kids (RAFFERTY *et al.*, 2017). However, current approaches for privacy control (e.g. user authentication, anti-phishing, email security, social media privacy and privacy policies) still lack usability for lay people (PACI; SQUICCIARINI; ZANNONE, 2018; BERTINO, 2016; OATES *et al.*, 2018; GARFINKEL; LIPFORD, 2014). Among these approaches, privacy policies are popular (SCHAUB; BALEBAKO; CRANOR, 2017). These policies are documents that inform users' consent about information sharing regarding a specific application (GARFINKEL; LIPFORD, 2014; SQUICCIARINI *et al.*, 2015). Nevertheless, these policies are usually long and complex (DE; ZEZSCHWITZ, 2016), and designing usable privacy policy tools for laypeople still challenges practitioners (PACI; SQUICCIARINI; ZANNONE, 2018; BERTINO, 2016; OATES *et al.*, 2018). The literature presents advances to the design of usable privacy policy tools, among which the "nutrition label" metaphor, from Kelley *et al.* (2009), is a promising approach. It stands among the few

[1] <ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en>

approaches in studies about mobile usability of privacy policy tools (VALLEE; SELBY; KRISHNAMURTHI, 2016; KELLEY; CRANOR; SADEH, 2013). Although the literature presents different approaches to enhance usability of such tools, we observed that studies that explore techniques of Information Architecture (IA) for usable privacy tools are still a gap.

In this study, we aimed to identify: (i) usability improvements for smart toy privacy control interfaces, (ii) adaptations for the nutrition label approach for the domain of smart toys and (iii) indications for more usable privacy policy tools.

To achieve our goals, we conducted a case study with voluntarily participants. The participants used an online tool to perform their Card Sorting individually. Later, we conducted a cluster analysis to group participants' answers and generate the IA. To re-design the parental control, we also employed concepts from the Kelley *et al.* (2009)'s nutrition label and Google Material Design guidelines.

The following sections present a literature review on usability, information architecture and privacy policy tools; the methods of this study; the results and discussions; and the conclusions and lessons learned.

3.3 Background

3.3.1 Smart Toys and privacy issues

Smart toys is a recent topic, concerning connected toys which are new to the IoT market and stands as a promising pedagogical approach.

Smart toys are connected to the cloud and, for this reason, they have raised concerns about children's privacy. As consequence, parents have striven to protect the privacy of their kids (RAFFERTY *et al.*, 2017). Although these new devices have raised such a concern, current approaches for smart toy privacy controls (e.g. user authentication, anti-phishing, email security, social media privacy and privacy policies) still lack usability, which becomes particularly important when users are lay people (PACI; SQUICCIARINI; ZANNONE, 2018; BERTINO, 2016; OATES *et al.*, 2018; GARFINKEL; LIPFORD, 2014).

Among the applications for privacy controls, privacy policies are popular (SCHAUB; BALEBAKO; CRANOR, 2017). These policies are documents that inform users' consent about information sharing regarding a specific application (GARFINKEL; LIPFORD, 2014; SQUICCIARINI *et al.*, 2015). Governmental acts, as The Personal Information Protection and Electronic Documents Act (PIPEDA) (Office of the Privacy Commissioner of Canada, 2021), require the use of privacy policies when businesses are handling personal information for their commercial activity. However, these policies are usually long and complex (DE; ZEZSCHWITZ, 2016), and designing usable privacy policy tools for laypeople still challenges practitioners (PACI; SQUICCIARINI; ZANNONE, 2018; BERTINO, 2016; OATES *et al.*, 2018).

The literature presents advances to the design of usable privacy policy tools, among which the "nutrition label" metaphor, from Kelley *et al.* (2009), is a promising approach. It stands among the few approaches in studies about mobile usability of privacy policy tools (VALLEE; SELBY; KRISHNAMURTHI, 2016; KELLEY; CRANOR; SADEH, 2013). Although the literature presents different approaches to enhance usability of such tools, we observed that studies that explore techniques of Information Architecture (IA) for usable privacy tools are still a gap.

Rafferty *et al.* (2017) proposed a model of parental control for smart toys. Their model proposes to provide parents (or legal guardian) with a mobile privacy parental control, with which they would be able to set privacy rules and also check commitment with governmental laws, such as PIPEDA. Analyzing the parental control model presented by Rafferty, Fantinato and Hung (2015), Rafferty *et al.* (2017), we understand that the Nutrition Label model is more appropriate to the context of our research problem. We made that decision because the parental control of Rafferty *et al.* (2017) depends on the use of mobile devices and we believe that the Nutrition Label model is more appropriate to the mobile context (VALLEE; SELBY; KRISHNAMURTHI, 2016; KELLEY; CRANOR; SADEH, 2013).

3.3.2 Usability and information architecture

Usability is defined by ISO/IEC 25066 as "*the degree to which a product can be used by specific users to achieve specific goals with effectiveness, efficiency and satisfaction in a specific context of use*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). In addition, according to Lewis (2014), usability is a sensitive concept, and can be defined as summative or formative. The author shows that summative usability aims to obtain measures and reach objectives of related projects to usability. Therefore, summative usability evaluation methods aim to quantify (assign a grade) to applications usability. According to Lewis, formative usability aims the diagnosis of usability problems. It's understood that the first step to form the usability of an application is to diagnose which are the usability problems exist in it to, then, correct them and, thereafter, form an advance in the application's usability. This study focuses in formative usability, since we understand that the state-of-the-art in researches of parental controls still require enhancements to form a more usable model of parental control. Therefore, we have used Card Sorting technique to evaluate the information architecture and, hence, the usability enhancement diagnosis of information architecture from enhancements in the information architecture.

Information Architecture (IA) models "*information and experiences products to support usability and discovery*" (MORVILLE; ROSENFELD, 2006). Since privacy policies are, usually, long and complex (SCHAUB; BALEBAKO; CRANOR, 2017), IA enhancements in the privacy policies interface may offer precise benefits to the usability of them.

Card Sorting[2] has been employed to generate a good Information Architecture (IA). Such

[2] <www.usability.gov/how-to-and-tools/methods/card-sorting.html>

technique consists in asking groups of users to group named cards (with contents or features of a system) in categories that make sense for them. Card sorting can be open (all users group cards in not named categories and, then, name them), closed (groups of cards with predefined categories and names) or hybrid (groups of cards in predefined categories and names, but changing or creating a category is allowed) (SALGADO; PEREIRA; FREIRE, 2016).

3.3.3 Usability and Privacy Policy Tools

Usable Privacy and Security and (UPS) is the research aiming to study usability of systems that assist users to administrate security and privacy of their data (DE; ZEZSCHWITZ, 2016; GARFINKEL; LIPFORD, 2014). To understand the UPS domain, it is necessary to understand each concept that makes up this field: security, privacy and usability. Information security can be understood as "*preservation of confidentiality, integrity, and availability of information*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016b), and privacy shows up as a critical requirement for information security (BERTINO, 2016). In this context, privacy refers to control over sharing of personal data (KOKOLAKIS, 2017; ACQUISTI; TAYLOR; WAGMAN, 2016; CHO *et al.*, 2018). Usability is "*the degree to which a product can be used by specific users to achieve specific goals with effectiveness, efficiency and satisfaction in a specific context of use*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a).

Initially, usability and security were seen as antagonistic, while users were seen as risks to information security (STILL, 2016). As software applications became more pervasive, privacy risks have increased (JANG-JACCARD; NEPAL, 2014; CHO *et al.*, 2018). As consequence, users were often required to take decisions about security (JANG-JACCARD; NEPAL, 2014) and seen as a great hope of the area (STILL, 2016). UPS researchers seek to develop security and privacy technologies that are usable for a variety of users (DE; ZEZSCHWITZ, 2016).

Whitten and Tygar (1999) have shown five security properties make UPS-related problems difficult to work with traditional user interface design: *The Unmotivated User Property*, *The Abstraction Property*, *The Lack of Feedback Property*, *The Barn Door Property* and *The Weakest Link Property*. In this study, we explore the abstraction property. This property states about the abstraction of security and privacy policies and rules (WHITTEN; TYGAR, 1999). Although the UPS literature has achieved considerable advances, Garfinkel and Lipford (2014) show that these properties remains indicative for researchers in the field. Usability of privacy policy tools remains one of the main themes in UPS (GARFINKEL; LIPFORD, 2014).

To protect citizens' privacy and avoid unauthorized use, the *Federal Trade Commission* of United States of America (USA) demanded that every e-commerce should publish privacy policies related to their practices over the Web (GARFINKEL; LIPFORD, 2014). As large Web companies are located in the USA, this determination had an international impact. Consequently, the importance of privacy policy controls has grown. Despite this importance, users rarely read

such policies, and the usability of these controls in the applications is one of the biggest challenges in the UPS area (GARFINKEL; LIPFORD, 2014; DE; ZEZSCHWITZ, 2016; BERTINO, 2016).

To improve usability of privacy policy tools, studies have explored alternatives to view and configure privacy policies. In this topic, the Expandable Grid (REEDER *et al.*, 2008; REEDER, 2008) has provided great insights for new tools. The Expandable Grid aimed to enhance the visualization and configuration of file permissions for the Windows XP interface. It was a matrix based interface, showing users as the upper axis, files as the vertical axis and permissions (read, write, execute, delete and administrate) as colored squares at the intersection between users and files. Reeder (2008) also proposed an adaptation of the Expandable Grid to be in accordance to the W3C Platform for Privacy Preferences Project (P3P)[3].

Later, Kelley *et al.* (2009) proposed improvements to the Expandable Grid interface, enhancing its usability. They made such improvements by simplifying it with based on the "nutrition label" paradigm. This paradigm was more familiar to users than the previous grid, because of its employment by the food industry. Kelley *et al.* adopted short labels to describe the vertical and upper axis of the P3P Expandable Grid. Meanwhile, they provided longer definitions at an additional screen, the "useful terms" page. They also adopted colored scales together with privacy symbols to indicate how information is collected and used. A legend provided explanation about the meaning of each symbol. According to them, users realized the benefits of the nutrition label paradigm when comparing policies, which was a positive usability related attribute. Although, they observed that users were confused by the symbols and unfamiliar with terms, which represent usability problems with the interface. The nutrition label interface was adapted to fit mobile interfaces by Vallee, Selby and Krishnamurthi (2016).

3.4 Methods

In our study, we aimed to empirically describe: *(i) usability improvements for smart toy privacy control interfaces, (ii) adaptations for the nutrition label approach for the domain of smart toys and (iii) indications for more usable privacy policy tools.* To reach our goals, our process is structured as follows:

1. Perform an open Card Sorting with a sample of potential users.
    - Provide a feasible amount of cards for users (≤ 20).
    - Each card must contain a short and representative content from the privacy policies. Privacy experts may provide these representative contents.
2. Perform cluster analysis with the outcomes from the Card Sorting.
3. Compare the resulted clusters against the nutrition label model (KELLEY *et al.*, 2009).

[3] <www.w3.org/P3P/>

4. Identify which screens can be abstracted with the nutrition label, and which must be created.
5. Transform the nutrition label with the mini-IA process (exemplified in Figure 5).
6. Use the resulted clusters (3) and the transformed nutrition label (6) to prototype the new interfaces.

3.4.1 Participants

To achieve our goals, we first evaluated the IA of the conceptual parental control proposed by Rafferty *et al.* (2017). For this reason, we conducted an open Card Sorting with potential users. We invited 42 participants to voluntarily take part in the Card Sorting, expressing their public opinion on how the information should be grouped. Because we had difficulties to find parents with available time to voluntarily participate, we invited female university students instead. This was done to be in accordance with the Brazilian Institute for Geography and Statistics (IBGE) statistics of motherhood in Brazil[4]. As indicated by the statistics, and to represent the profile of Brazilian mothers, the invited students had ages ranging from 20 to 29. To the best of our knowledge, there are no statistics on percentage of live births by age of the father at birth. Therefore, we did not sample male students in this study. Our sampling method is defined by this feasibility analysis (CAINE, 2016).

3.4.2 Procedures

We carried out a Card Sorting session for each user, so that a user's opinion would not influence other user's opinion. We asked users to group cards according to their own preferences. Each each card contained a different information about parental control. To compose the cards, we retrieved terms from the parental control ofRafferty *et al.* (2017). They summed 19 terms among all the screens of Rafferty *et al.* (2017)'s model. Table 2 indicates the respective terms that we used as cards for the Card Sorting evaluation.To retrieve the terms from Rafferty *et al.* (2017) prototype, we did not considered those terms at the menus. We did not consider those because they already represented information groups (IA), which was in accordance to the designers of the model. This study aimed to understand the conceptual model of IA from potential users' opinion and, for this reason, we could not consider the conceptual model from the parental control designers.

We decided to conduct a Card Sorting experiment because privacy policies are usually long and complex (SCHAUB; BALEBAKO; CRANOR, 2017), and IA enhancements of privacy policy interfaces may benefit usability of privacy policy tools. In addition, to the best of our knowledge, this study innovates by exploring the use of Card Sorting to enhance the design

[4] <www.ibge.gov.br/estatisticas-novoportal/sociais/populacao/9110-estatisticas-do-registro-civil.html?=&t=destaques>

of privacy policies. For this reason, we also collected lessons learned from our experiences to further studies. To reduce the required time to collect card sort responses, and because online Card Sorting tools often limit their features for free plans, we have developed our own online Card Sorting tool using the React Kanban library[5]. Responses were collected as JavaScript Object Notation (JSON)[6] and, later, converted to Comma Separated Values (CSV) to be analyzed using R packages[7].

Table 2 – Terms used for the cards

ID	Information
I1	Child
I2	Child information
I3	Purpose of Access to Child Data
I4	Contact details of parent/guardian
I5	Privacy Policy
I6	Review Privacy Policy
I7	Agreed with the privacy policy service
I8	Show all privacy rules
I9	Description of the Privacy Rule
I10	Create new privacy rule
I11	Review and add privacy rule
I12	Enable privacy rule
I13	Disable Privacy Rule
I14	Receive updates via email
I15	Authorize access to GPS
I16	Mobile service authorized to access children's data
I17	Obligations and Retention of Data
I18	Main Control of Access to Child Data
I19	Choose Recipients to receive child data

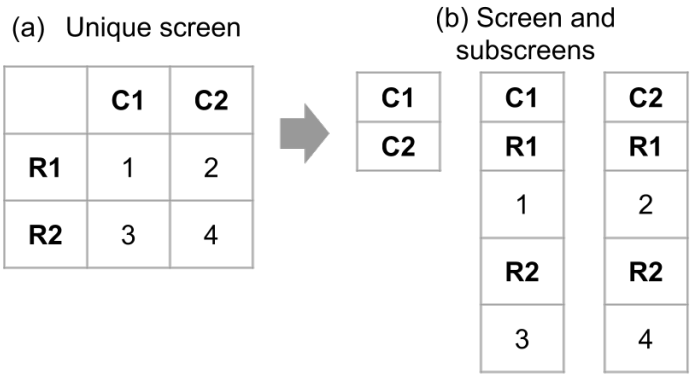
Source: Research data.

In sequence, we adapted the nutrition label interface (KELLEY *et al.*, 2009) using the mini-IA process (NIELSEN; BUDIU, 2013) to compose a Android mobile interface. We understand that the Nutrition Label model is appropriate to visualize and configure privacy policies in mobile devices, because it has been used as basis to support the design of mobile privacy policy tools (VALLEE; SELBY; KRISHNAMURTHI, 2016). In addition, the parental control model of Rafferty *et al.* (2017) depends on the use of mobile devices. The mini-IA process is based on the logical reading order (Western style) to turn the table in a unique column. Figure 5 illustrates the mini-IA process, transforming a tabular interface (columns and rows) into a mobile interface with a unique column with hierarchically displayed information. In addition to the adaptations made to the nutrition label, we also defined interface elements using

[5] GitHub public repository: <github.com/markusenglund/react-kanban>
[6] <www.json.org/>
[7] <stat.ethz.ch/R-manual/R-patched/library/stats/html/hclust.html>

Google Material Design guidelines for mobile devices[8]. Finally, by conducting these steps, we have observed lessons learned in this study to indicate how the (re)design processes might be structured for further practical cases.

Figure 5 – The mini-IA process illustrated. Transforming (a) table columns (C#) and rows (R#) in (b) unique column screens for mobile device interfaces.



The following section shows the results from our methods and provide deeper details on the process to re-design the parental control model.

3.5 Results and Discussions

3.5.1 Card Sorting Experiment

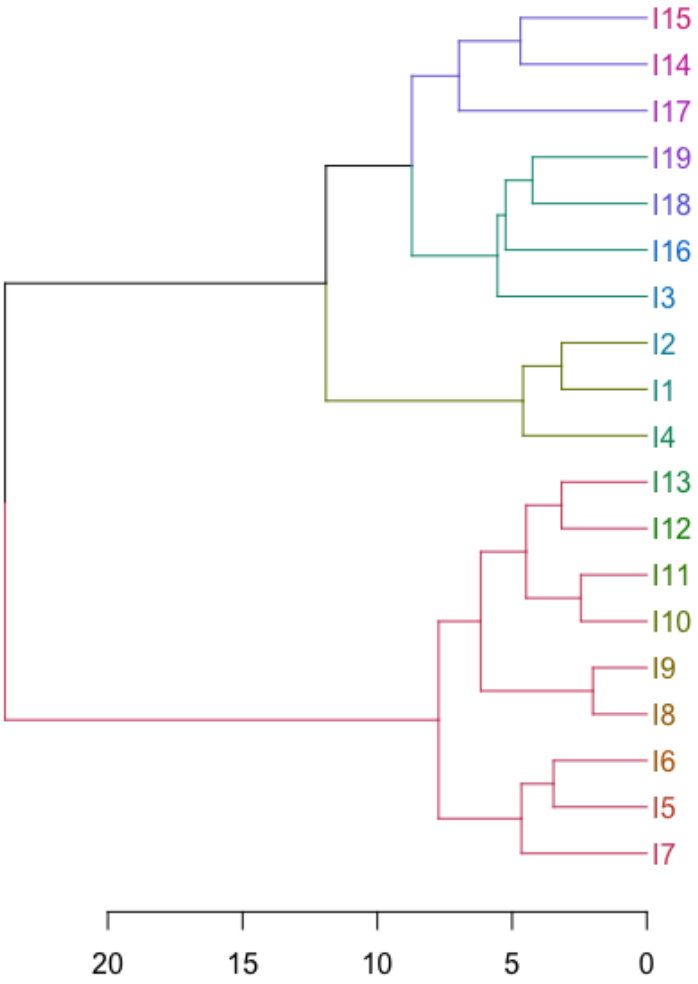
We collected 30 valid responses from the Card Sorting sessions. These responses had, at least, two categories of cards grouped by the participants. From 42 participants, we took out 12 registered answers that were not completed (when participants started the grouping process but did not use all terms). To analyze the results from the Card Sorting, we used the Ward method to group participants' answers based on Euclidean distance (MURTAGH; LEGENDRE, 2014). Euclidean distance is suitable for the case because the results of Card Sorting can be represented in a two dimensions Euclidean space. We decided to employ cluster analysis to evaluate the outcomes from the Card Sorting because we could not have all participants together to discuss the final IA.

The results obtained were analyzed and are presented at Figure 6. As shown at the table, we observed two main ramifications in the resulted dendrogram. The lower branch presents a greater degree of similarity between the terms. It includes terms related to privacy policies and creation of privacy rules. The upper branch presents less similarity between its terms. In

[8] <material.io/design/>

it, a subgroup stands out with greater similarity between the information, represented by the information I1 (Child), I2 (Child information) and I4 (Details of the father, mother or guardian).

Figure 6 – Dendrogram representing the IA as resulted from the Card Sorting and cluster analysis.



The results showed that some features that were presented in more than one screen of the prototype of Rafferty *et al.* (2017) could, according to the observations of this study, be presented on the same screen. Some examples of that are: the layout in different privacy policy display screens (I5); task of reviewing the privacy policy (I6); description of the privacy rule (I9); review and add privacy rule (I11); contact details of the parent or guardian (I4); and information of the child (I2).

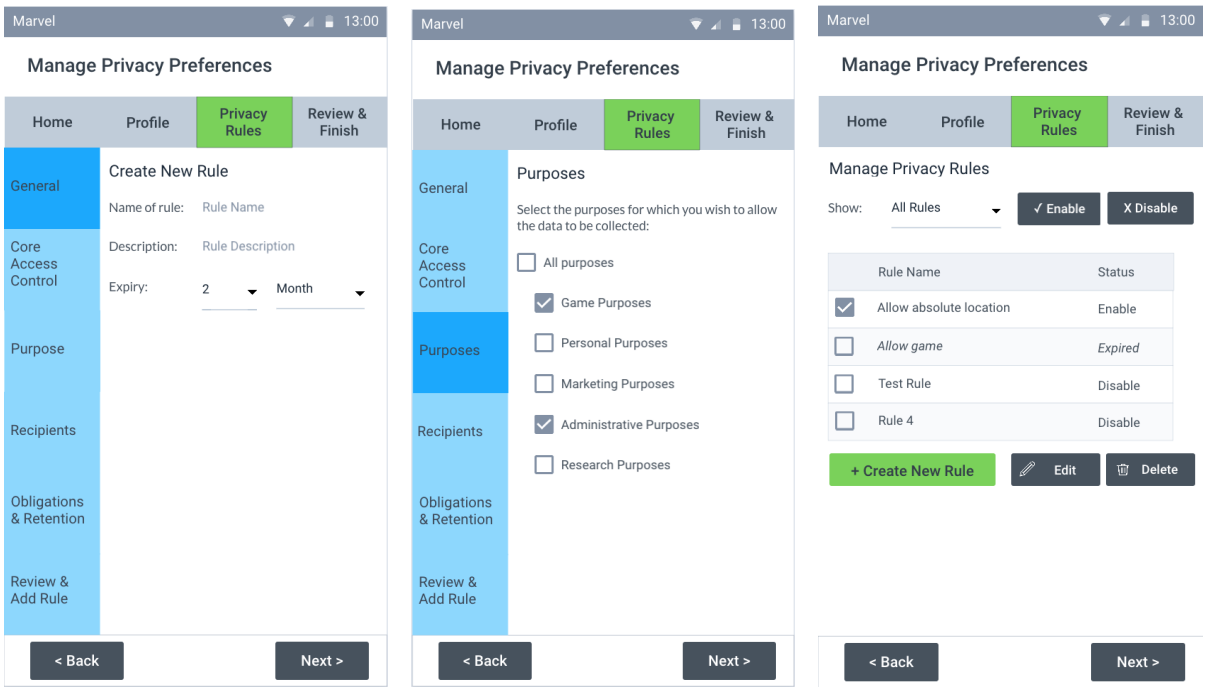
3.5.2 Prototyping the new Parental Control

To prototype our parental control, we first created an Android like version of Rafferty *et al.* (2017)'s model (e.g. see Figure 7). This was done to enable necessary comparisons in

this study. For this version, we performed as few adaptations as possible, so that the model could remain as much original as possible. Thereafter, we adapted the nutrition label model and considered the adaptation results to construct the new prototype.

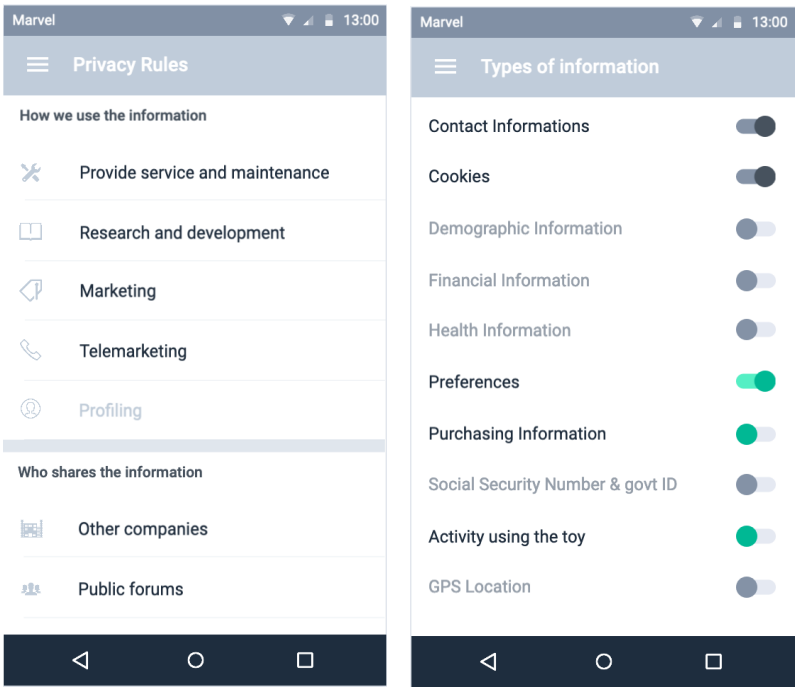
To adapt the nutrition label model, we considered the results from the open Card Sorting and cluster analysis. We compared these results against the nutrition label model. We found that the terms I15 (Authorize access to GPS), I18 (Main control of access to the child's data) and I3 (Purpose of access to the child's data) could be removed, because the nutrition label already contains those terms among its terms and symbols. For this reason, we understood that repeating these terms would represent excessive information in the interface. As indicated by the usability principles of Jakob Nielsen (2018), any excess information competes with relevant information. The results of the removal of excessive information for the contents represented by the nutrition label model can be visualized in the comparison between the interfaces represented at Figure 7 (Model of Rafferty *et al.* (2017)) and at Figure 8 (our prototype).

Figure 7 – Main sections of the "Privacy Rules" category of the prototype proposed by Rafferty *et al.* (2017).



In sequence, we adopted the strategy for mini-IA (NIELSEN; BUDIU, 2013) to allocate the information in appropriate groups, logically displayed according to users' reading style (Western style). Figure 5 illustrates the process of reshaping the (a) nutrition label tabular interface into (b) multiple single-linear interfaces, as recommended by Nielsen and Budiu (2013). We call this the mini-IA process. As shown at Figure 5, we had to divide the table into more than one screen: one root screen containing column's labels only (C#) and multiple sub-screens containing the row's labels (R#) and the values (e.g. 1 to 4). At Figure 8, we show the result of

Figure 8 – Our new prototype.



the mini-IA process of Kelley *et al.* (2009) nutrition label table (see Figure 5 reported in the Kelley *et al.* study). For such, we re-organize the information according to the logical sequence of reading of the Western world (left to right followed by top to bottom), as suggested by Nielsen and Budiu (2013).

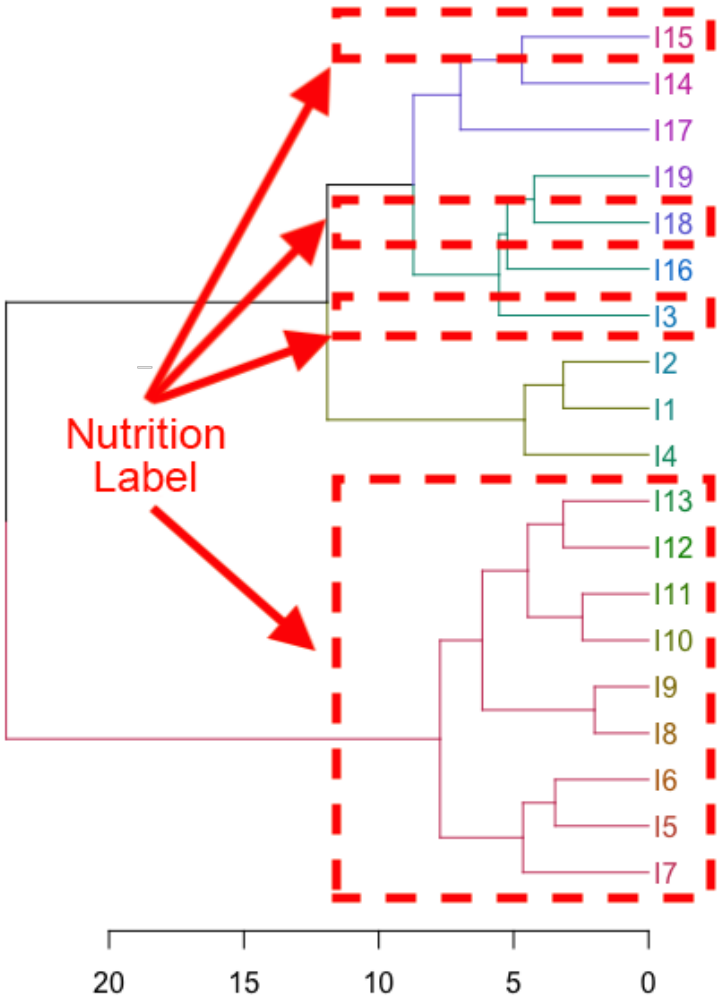
To sum up, we used Google's *Material Design*[9] guidelines and assets from the Marvel App[10] prototyping tool. All prototyping steps were done in the Marvel App tool, considering the interface elements offered by the tool and which are in agreement with the *Material Design* guidelines by Google.

As result, some screens of our prototype are very similar to the screens of Rafferty *et al.* (2017)'s interfaces, despite the use of menus (which in our proposal is hidden waiting for the user's event). Examples are the parent profile screens (see Figure 10) and the children (see Figure 11). On these screens, the navigation buttons have been replaced by the default navigation bar of *Material Design* and by creating a "Continue" button that also follows the standards of *Material Design*. This fact indicates that such screens from Rafferty *et al.* (2017)'s interface may have an appropriate usability for its purposes.

At the end of the re-design process, we reviewed the symbols used at the nutrition label interface. Because the nutrition label employs detailed symbols, which may be difficult to read on mobile screens, we suggested some updates by employing Google Material Design symbols.

[9] <material.io/design/>
[10] We thank for the kind permission of the Marvel App to display images of their assets in our study.

Figure 9 – Comparing the dendrogram clusters with the information grouped at the nutrition label model.



Since the use of *switches* in other Android configuration apps, we used variations of the *switch* to represent the four symbols of nutrition label (’!’, ’OUT’, ’-’ e ’IN’). Figure 12 shows how the definition suggested by Kelley *et al.* (2009) related to the icons chosen by us.

The adoption of the nutrition label model indicated that the creation of privacy rules for each mobile service (as required in the previous parental control model) may be complex for users. The nutrition label model concentrates all the rules in the same interface, and different services must comply with it. This may enhance the efficiency of the parental control by saving users’ time and effort. Because efficiency is a requirement for usability, this finding may also enhance the usability of parental controls.

Figure 10 – Subcategory of "Details of the Parent / Guardian". Rafferty *et al.* (2017)'s model on the left, and our new model on the right.

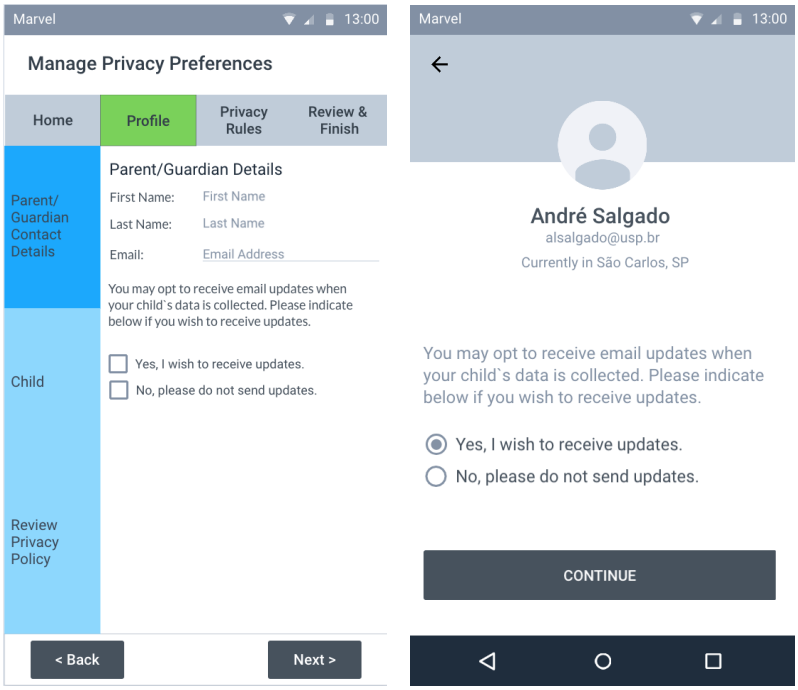


Figure 11 – Subcategory of "Child Information".Rafferty *et al.* (2017)'s model on the left, and our new model on the right.

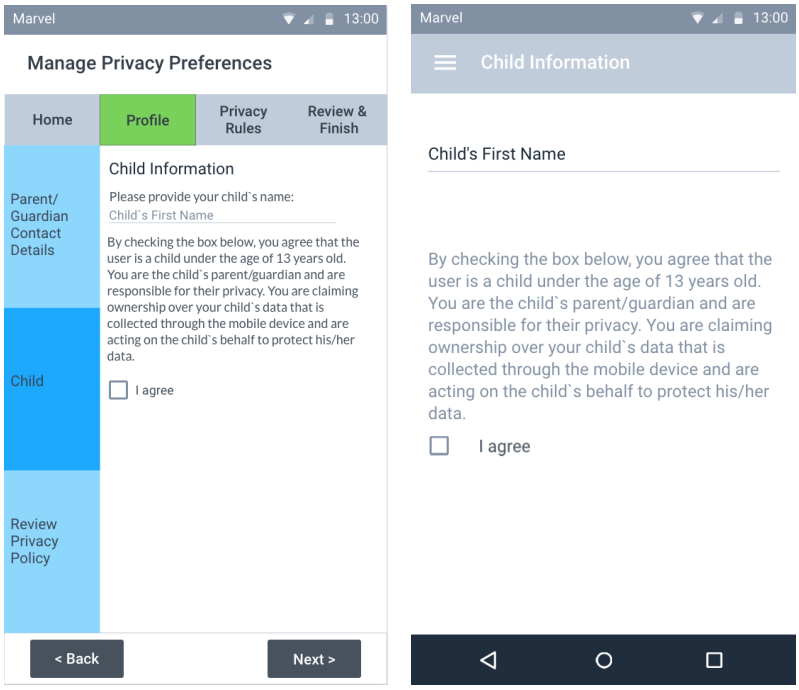
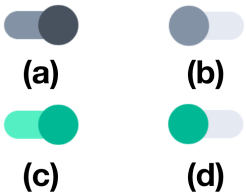


Figure 12 – Icons'labels adapted from Kelley *et al.* (2009): (a) "*We will use the information in this way*"; (b) "*We will not collect or we will not use the information in this way*" ; (c) "*We will use the information in this way unless you opt-out*"; and (d) "*We will not use the information in this way unless you opt-in*" .



3.6 Conclusions

In this study, we aimed to identify and describe: *(i) usability improvements for smart toy privacy control interfaces, (ii) adaptations for the nutrition label approach for the domain of smart toys and (iii) indications for more usable privacy policy tools.*

Our findings provide a linear interface (non tabular) for the nutrition label model, aiming parental control for smart toys. We also suggest the use of material design icons to replace symbols from the original nutrition label interface. In addition, we reduced the amount of information from Rafferty *et al.* (2017) parental control model. Although privacy policies require long and complex information, there is a need to reduce to what is more relevant to the task. Detailed information can be place at secondary screens, as suggested by Kelley *et al.* in the nutritional label model. These modifications assisted us to compose a more efficient parental control in what regards the tasks: *reviewing, reading and creating privacy policies*; and *profile information of parents/guardians and children*. We understand that these were the main improvements that this study has proposed to the parental control model of Rafferty *et al.* (2017).

In this paper, we created a six steps process (see Section 3.4) to perform Card Sorting with cluster analysis in evaluation of IA of mobile privacy policy tools. The results of our method corroborated the IA of the original nutrition label interface, and its groups (clusters) of information. We understand that the main advantage of employing method is that resulted clusters can evolve over time, as new Card Sorting answers may be collected. Thereafter, clusters can be updated by inserting the new answers an executing the cluster analysis again. This is not dependent on discussions with users to define the final IA, but only on data analysis and specialist opinions, which makes the Card Sorting faster to apply. We believe that adopting our process may help praticitioners to design more usable privacy controls. Nevertheless, the effectiveness of our process still needs validation. Future studies can investigate this topic and, also, provided deeper activities for the process. We also suggest, as future studies, to investigate the employment of this process in the design of privacy nudges (ACQUISTI *et al.*, 2017).

Future studies may diagnose usability problems with our prototype and, then, enhance its

quality. Further, we suggest exploring the benefits of using our prototype to generate prototypes for other domains of IoT privacy controls (e.g. connected and autonomous vehicles). Iconography studies are also suggested to better understand the use of icons to replace the symbols suggested by Kelley *et al.* (2009) in their nutrition label. Also, future studies are suggested to investigate the differences on users' performance using our and Rafferty *et al.* (2017)'s model.

CHAPTER

4

RECOMMENDATIONS TO ENHANCE USABILITY AND PRIVACY OF SMART TOYS

4.1 Abstract

The collection of personal information by smart toys causes various privacy concerns. The use of personal information has also been subject to regulatory acts by different governments. For these reasons, smart toy manufacturers need to develop effective privacy controls. However, designing usable privacy controls remains a challenge. In this paper, we sought to identify the main security vulnerabilities involved with smart toys that are related to usability and may

[1] <https://creativecommons.org/licenses/by-nc-nd/4.0/>

impact on users' privacy. To this end, we performed a security analysis and usability heuristic evaluations. After identifying current vulnerabilities, we create a list of design recommendations aiming at enhancing both the usability and privacy of smart toy privacy controls. We also suggest a revised severity scale to help to prioritize the design solutions.

4.2 Introduction

The Internet of Things (IoT) is an ecosystem that is transforming all devices to build a smart society (MOTTA; OLIVEIRA; TRAVASSOS, 2018). These smart devices have benefited consumers in many ways, such as smart thermostats placed in the home and wearable technology to monitor health and fitness (CONTI *et al.*, 2018). IoT has also influenced children's toys that have transformed from simple, stuff toys to Internet-connected toys that can also communicate and interact with children (HUNG; TANG; KANEV, 2017; MCREYNOLDS *et al.*, 2017b). For example, Hello Barbie, an Internet-connected toy from ToyTalk.com and Mattel, operates when the button in the belt buckle is pressed, and it connects the Hello Barbie doll to the Cloud server of ToyTalk.com (Hello Barbie, 2019). CogniToys Dino is another smart toy powered by IBM Watson technology that is Cloud-connected and operates through the Internet. Dino works simply as when the child asks questions by voice, and Dino that is connected to the Internet listens and replies according to the question (COGNITOYS, 2019a). These devices can provide personalized based services to users by collecting data from user contexts such as location, time, and weather. The Elemental Path has described the functionality of CogniToys Dino as that it gathers child personal behavior and preferences such as favorite color, favorite games and provides service according to their age-appropriate content to interact with them (ULANOFF, 2015). However, the collection and use of such sensitive information are subject to regulatory acts, such as the Children's Online Privacy Protection Act (COPPA), from the United States Federal Trade Commission (Federal Trade Commission ("FTC" or "Commission"), 2013), and the General Data Protection Regulation (GDPR), from the European Union (THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, 2016). Also, users (or their legal guardians) may not consent to such devices collecting their personal information. Therefore, smart toy manufacturers are required to implement effective privacy controls to protect the collected information (MCREYNOLDS *et al.*, 2017b; RAFFERTY *et al.*, 2017; GHOSH *et al.*, 2018).

We have seen in recent years, several privacy violations or data breaches, such as the VTech breach that resulted in the disclosure of about 6 million children records (NELSON, 2016). Table 3 presents some well-known children's privacy violations due to ineffective security control or privacy malpractice and their respective related fines, in US dollars, levied by United States Federal Trade Commission (FTC) against the company violating children's privacy.

Table 3 – Some Known Privacy Violation and Fines.

Company	Violation	Year	Fine
ByteDance	COPPA compliance failure with their TikTok app	2019	$5.7 million
Oath	COPPA violation - Online advertising	2018	$5.0 million
inMobi	COPPA violation - location tracking	2016	$950,000

Source: Elaborated by the author.

Although the FTC continues to levy hefty fines against companies violating COPPA, adopting effective privacy controls remains a challenge in the field (PACI; SQUICCIARINI; ZANNONE, 2018). To address this challenge, we sought to review the main security vulnerabilities of some current smart toys and their resulting user privacy concerns and impact. For this reason, we identified security and usability problems that remain present in popular smart toy applications, causing vulnerabilities. To overcome these problems, we present a list of recommendations for further improvements in smart toy technologies.

This paper is organized as follows: Section 2 presents the background on smart toys, information security, privacy, and usability. Section 3 describes the method of this study. Section 4 presents the case studies we performed to reach our goal. Section 5 presents recommendations for future designs of smart toy privacy controls. Finally, Section 6 concludes our paper and discusses future works.

4.3 Background

As shown by Albuquerque et al. (ALBUQUERQUE *et al.*, 2020), although privacy is difficult to define, it relates to the right of people to keep their personal information as secret or not. It generally refers to one's desire to set who has access to them. This is closely related to the concept of confidentiality, which is defined by ISO 27000 as the "*property that information is not made available or disclosed to unauthorized individuals, entities, or processes.*" In this context, confidentiality is an extension of privacy but focuses on how the user's private information is managed to prevent unauthorized users from gaining access. Essential security controls established in ISO 27001 required to maintain confidentiality can be considered as very important, and establishing and protecting privacy.

The literature has diverse approaches focusing on the privacy and security issues of IoT devices, among which smart toys have an increasing interest. Hung et al. identified privacy requirements at the legislative level and privacy laws that are applicable to children's smart toys. They showed that, as the physical safety of a child is mandatory, a framework was needed to attain the privacy of the child by reducing the sensitive data collection and its retention. This included a parent or guardian to control their child-sensitive data (HUNG; FANTINATO; RAFFERTY, 2016). Meanwhile, Rafferty et al. proposed a conceptual model of privacy rule for smart toys,

IoT devices, and mobile services. In the model, parents and legal guardians are owners of child information, which is in accordance with a data privacy act known as COPPA (Children's Online Privacy Protection Act). COPPA allows parents and legal guardians to monitor and regulate the information that is gathered online. Parents must give their consent to rules (access rules) about sharing their child's personal data (RAFFERTY *et al.*, 2017).

McReynolds et al. conducted a survey on parents and children who play with internet-enabled toys. They emphasize the survey on worries that parents and their children have when playing with smart toys, observing that many children were not even aware that their conversations are being recorded. They have pointed out that the toy designers should design toys in a way that it alerts children before recording instead of the red blinking light that is not spotted by children. They have also suggested to toy manufacturers not to keep the recordings of child conversations for a long time and delete in a week or allow parents to delete the recorded conversations permanently. Their study also found that many parents require parental control over the toy, such as the function to turn off the Internet on the toy or to manage its responses to children's questions (MCREYNOLDS *et al.*, 2017b). Yankson, Iqbal and Hung (2020) suggested that toy manufacturers should consider forensic measures while designing internet-enabled toys . Rafferty presented an access control model and framework intended to protect the location of children playing with Internet-connected toys (RAFFERTY, 2015).

Finally, Holloway et al. (HOLLOWAY; GREEN, 2016) show the potential benefits of smart toys (e.g., enthusiasm and enjoyment). Meanwhile, they outline various emerging privacy and security issues found in smart toys. According to them, ToyTalk (responsible for the Hello Barbie) argues that it is not possible to prevent children from providing personal information. Nevertheless, ToyTalk's policies state that if the company comes across any recordings with personal information, the company will delete it. In this sense, Holloway et al. argue that the security protection of smart toys depends on parental choice over parental control. Also, they argue that this may involve other security breaches.

4.3.1 Privacy and Smart Toys

Hung *et al.* (HUNG; TANG; KANEV, 2017) defines smart toy as:

> "a device consisting of a physical toy component that connects to one or more toy computing services to facilitate gameplay in the Cloud through networking and sensory technologies to enhance the functionality of a traditional toy."

Smart toys establish two-way communication with the child (RAFFERTY *et al.*, 2017). The smart toy vendor is able to provide personalized based services through the collection of data from users' context. Smart toys often gather the childs' personal behavior and preferences, such as favorite color, favorite game, and in order to provide age-appropriate content for the

child to interact with the toy. By interacting with smart toys, the toy can gather personalized information about the child. In most cases, the guardian and the child both have no idea of the concept of privacy and how to protect it. Consequently, children reveal their personal information while playing with these toys without the awareness of the dangers of such information revel (RAFFERTY *et al.*, 2017). The personal information used and collected by these connected devices can be hacked; as such, sparking various security and privacy concerns. The concerns become exponential with respect to sensitive personal information about children, as all interactions of a child with the Internet-enabled toy are stored somewhere else on a remote server (RAFFERTY *et al.*, 2017). Because of the challenging nature of privacy and connected devices, some manufacturers of smart toys may not design security and privacy as a top requirement.

The literature shows that some smart toys available on the market remain with security threats. Mattel Hello Barbie, My friend Cayla and i-Que robot are examples of such toys (MILLAR *et al.*, 2017). For instance, parental control is needed for the proper functioning and more security of the toy (RAFFERTY *et al.*, 2017; YANKSON; IQBAL; HUNG, 2020). Hello Barbie is designed to be the child's best friend, talking and sharing secrets (HOLLOWAY; GREEN, 2016; MILLAR *et al.*, 2017). The doll has built-in features that record every conversation between child and Barbie and stores this conversation in a cloud database. This database is also shared with the child's parents, which gives the impression that parents (or legal guardians) have absolute control over the conversations. Hello, Barbie application also includes a feature to share the recordings of children's conversations with the toy on social networks. As a matter of fact, there is a possible threat to sharing the collected data with third parties. This indicates that in both ways, Hello Barbie is not keeping a secret (HOLLOWAY; GREEN, 2016; MILLAR *et al.*, 2017).

Holloway and Green (HOLLOWAY; GREEN, 2016) discuss that security specialists can easily get access to the names of all Wi-Fi networks to which the toy connects, the user account details, and even the sound files of pre-recorded responses of Barbie conversations when the doll is not connected to the Cloud. In 2015, VTech Electronics LLC, a company that develops connected tablets for children, suffered a data breach of almost 6 million children and 4 million parents all over the world (NELSON, 2016). The information included parents' and child names, birthdate, pictures, gender, and account password. VTech failed to protect the Personally Identifiable Information (PII) of parents and their children that they have collected for the use of their connected tablets (SASSE *et al.*, 2016). However, these kinds of data breaches elevate concerns about the privacy of users' data; and rightful question whether these smart toy manufacturers are considerably doing enough to implement security controls necessary to address privacy risks of the collected consumer data. According to the privacy policy of CogniToys Dino (COGNITOYS, 2019b), the Personal Information provided by parents about themselves and their children may include name, home address, contact information, current location, email address. As one can see, this information is privacy sensitive and sufficient to identify users. The policies state that information is only used for the internal purpose, to give

a personalized experience to users, and that the toy company is not going to reveal customers' collected Personal Information to third parties without the consent of users, except as described in their Privacy Policy. This may allow the company to disclose some information to third parties without identifying the identity of the parent or child. For example, to attest that smart toy companies are considering the privacy of children's information in their care, a Ranking Member, Nelson, of the US Senate, requested the security and privacy policies from few famous connected devices and toy companies. He also requested information about how they collect, use, and secure user personal information. The companies in question provided him with the report that reveals the smart toys gather much information, including Personal Identifiable Information of parents and their children. The companies also showed that they have security policies for user data protection applied. However, the security vulnerabilities in Fisher-Price Smart Toy Bear uncovered that they were unsuccessful in protecting and securing customer data. These incidents elevated questions of whether smart toy manufacturing companies are considering the security of consumer data as their top priority (SASSE *et al.*, 2016).

Due to privacy concerns related to smart toys, studies have analyzed the security of these devices in order to identify vulnerabilities. Somerset Recon Inc (Somerset Recon Inc., 2016) analyzed the security of Hello Barbie, one of the first smart toys to become popular in the market. They have identified security vulnerabilities that can be considered as privacy vulnerabilities due to its impact on privacy, as we present in Table 4. A similar security analysis is made available by Pen Test Partners (PARTNERS, 2016) on the Dino smart toy, another popular smart toy in the market. We also included these analyses in Table 4.

Table 4 – Smart toy Vulnerabilities and Privacy Impact.

#	**Information Security Vulnerabilities**	**Privacy Impact**
1	"*Weak passwords*" (Somerset Recon Inc., 2016).	This vulnerability will allow attackers to brute force user account credentials remotely and infiltrate victim user accounts. However, we have found that this issue has been resolved now.
2	"*No Password Brute Force Protections*" (Somerset Recon Inc., 2016).	This vulnerability allows attackers to brute force user mobile app account passwords remotely and infiltrate victim user accounts. An attacker is also able to gain access to audio conversations of the toy with a child as it is accessible through user accounts. However, as we observed, it has been resolved now.

3	"*Hello Barbie device uses unencrypted Wi-Fi network*" (Somerset Recon Inc., 2016).	This vulnerability allows attackers to perform a man in the middle attack by joining open and unencrypted Barbie's Wi-Fi network. However, this Wi-Fi connection is only available in pairing mode by pressing two buttons on the device. There is a possibility that the child might unknowingly press these two buttons and open the Wi-Fi device network. We observed that this vulnerability had not been resolved by now.
	"*Hello Barbie device does not require unique authentication to modify the configuration of the device*" (Somerset Recon Inc., 2016).	This vulnerability could cause the toy to use an account created by the attacker, and in this way, an attacker can listen to audio conversations. An attacker could also gain access to the user account credentials from the toy web application and insert malicious audio conversation files to the victim user account (Somerset Recon Inc., 2016). However, we observed that it had been resolved now.
4	"*Audio files can be accessed without authentication*" (Somerset Recon Inc., 2016).	An attacker can get the URL of an audio conversation that is stored on CloudFront without authentication, and the file is accessible even if the user changes the account password. The problem faced by an attacker while accessing those audio conversation files would be that URL paths to all audio files are random (Somerset Recon Inc., 2016). We observed that this security vulnerability had not been resolved yet.
5	Cross-Site Scripting: The web Interface of the toy, which is available over Wi-Fi and is used in configuration mode, is vulnerable to a few security issues. This includes persistent Cross-Site Scripting (XSS) attack (PARTNERS, 2016).	The web page does not perform input validation or sanitization while entering the SSID, and by submitting the script such as "*<script>alert (1) </script>*" (PARTNERS, 2016), the code gets executed and displays "*1*". An attacker could exploit this vulnerability and perform Persistent XSS and Cross-Site Forgery Request attacks. As we observed, this security vulnerability has not been resolved yet.

6	Use of HTTP for transferring sensitive information: The web interface of the toy is used to add or modify Network SSID. The SSIDs that are in use or to set a new SSID with different priority levels are displayed on the web interface. The users can select any security type and enter a password to connect to SSIDs. This web page uses an unsecured connection HTTP, i.e., <http://192.xxx.x.x>, and it could be easily accessed by the hacker (PARTNERS, 2016).	When the toy is in configuration mode, a hacker can perform malicious activities such as Man in the Middle (MITM), by sniffing the traffic between the user and the toy and stealing any sensitive information. However, it would be better to set login credentials to enter the web interface of the toy. As observed in our study, this security vulnerability has not been resolved yet.

Although security issues are important to identify privacy vulnerabilities, studies have shown that usability also plays an important role in enhancing users' efficacy, efficiency, and satisfaction with different privacy controls (NELSON, 2016; BERTINO, 2016). For this reason, in this paper, we conduct an empirical case study to evaluate the usability of examples of smart toys aiming to identify privacy vulnerabilities.

Usable security issues in smart toys

Recent studies have shown that usability may play an important role in enabling laypeople (as parents/guardians) to effective use of privacy controls (NELSON, 2016), (SASSE *et al.*, 2016). The usability concept is defined by the ISO/TR 9241 as (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018b) as:

> " the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use."

We describe a deeper relationship between usability and privacy controls by first considering the definition of usability. In regards to the usability definition, and considering the context of privacy controls, "*specified goals*" (part of usability definition) are control objectives "*(...) to be achieved as a result of implementing controls*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016b). Because privacy controls are provided and configured by means of user interfaces, poor usability of such interfaces (e.g., because of poor effectiveness) may be seen as a weakness of the privacy control process and can be exploited by a threat. In other words, and considering that the "*weakness of an asset or control (3.14) that can be exploited by one or more threats (3.74)*" is an information security vulnerability, poor usability of privacy controls

may be seen as information security vulnerabilities (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016b; JØSANG *et al.*, 2007).

As the range of privacy threats increases, laypeople are often required to make security decisions (JANG-JACCARD; NEPAL, 2014) by understanding privacy concepts or polices. However, privacy policies are usually long and complex (DE; ZEZSCHWITZ, 2016), and usable tools for laypeople are still needed (NELSON, 2016; SASSE *et al.*, 2016). To design usable tools of any kind, usability evaluations are essential (HORNBÆK, 2010). These methods can be distinguished between those that depend on end-users to be performed (use-based evaluations), and those that depend on inspectors to be performed (inspection-based evaluations) (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a). Among inspections, heuristic evaluation (HE) is popular and allows practitioners to diagnose usability problems on the interface (LAZAR; FENG; HOCHHEISER, 2017).

4.4 Methods

In this study, we sought to identify the main security vulnerabilities involved in the smart toys' context that have an impact on users' privacy. For this reason, we complemented the findings from a literature overview (shown in Section 2.1) with additional security analysis and two empirical usability inspections. The Security analysis stage is a security analysis of one smart toy technology to confirm the findings from the literature and, potentially, identify new issues. This stage can also confirm whether the set of security vulnerabilities is saturated, and no new vulnerabilities are found. The usability evaluation stage comprises the evaluation of two smart toy privacy controls. This is performed to identify human vulnerabilities involved with smart toy privacy controls that are due to usability aspects. To this end, we performed heuristic evaluations. Heuristic evaluations can identify information security vulnerabilities by means of employing usable security heuristics as criteria for the judgment (JØSANG *et al.*, 2007). Heuristic evaluation returns situations when users might face usability problems when setting their privacy controls. This may lead to privacy risks. Our goal is to identify usable privacy issues that can help the literature to understand how to improve the interface of smart toys' privacy controls.

4.4.1 Participants

Both the security analysis and usability evaluation were carried out by experts (four experts). The security analysis was conducted by two security experts (both Ph.D. students). Meanwhile, the usability evaluation was an expert review (Assistant Secretary for Public Affairs, 2013), conducted by two usability experts (a usability researcher, Ph.D., and a Ph.D. student). To conduct the usability evaluation of smart toys' privacy controls, we considered the privacy controls as available on current application markets. Due to time constraints, and because the

literature has previous security analysis on smart toys, we only performed the security analysis on toy Alpha. For the usability evaluation, since the literature still lacks usability evaluation of privacy issues on smart toys, we performed it on both toy Alpha and Beta.

4.4.2 Material

We employed two smart toy technologies as a subject for the experiments. We used the "*Privacy not included*" website from Mozilla (MOZILLA, 2018) to choose both smart toys. To keep the anonymity of the brands and their privacy, we refer to the technologies as toys Alpha and Beta. The smart toy brands mentioned previously in various sections of this work have no direct connection to toy Alpha or Beta used in this section of our work. Toy Alpha is a smart interactive toy that makes conversations with kids. It is connected to a Cloud-based Artificial Intelligence machine for question answering, which operates through Wi-Fi. The setting of Alpha is made available with the free app, which is available to be downloaded for Android or iOS-based phones. For feasibility reasons, the Android-based mobile app of the toys has been used throughout the case study. The Web interface of Beta had input validation errors such as Cross-Site Forgery Request and Persistent XSS. Moreover, it uses unencrypted communication channel HTTP instead of HTTPS to transmit sensitive information. It also allows weak login credentials while creating a user account. We employed toy Alpha for the security analysis and one usability heuristic evaluation.

Beta is an internet-connected toy designed and developed by a traditional toy company and a computing company focused on talking toys. It is aimed to communicate with children, while all conversations between Beta and the child are stored in the Cloud and can be accessed or managed by parents on a dedicated website. Toy Beta can be easily configured with the application available to be downloaded for Android or iOS-based phones on their app store. We employed toy Beta for the second heuristic evaluation.

4.5 Results and Discussion

4.5.1 Security Analysis

The security analysis was based on both mobile and Web/desktop versions of a smart toy privacy control. To complement the analysis, we used the Wireshark (WIRESHARK. . . , 1998) tool to clearly check what happens in the toy connection with the Cloud. The privacy vulnerabilities and their impacts are indicated in Table 5.

Table 5 – Vulnerabilities and related Privacy Impact.

#	**Information Security Vulnerability**	**Privacy Impact**
1	Weak password	This vulnerability will allow attackers to gain access to users' accounts and all private and sensitive information about the user. This security vulnerability has not been resolved.
2	No Password Brute Force Protections	This vulnerability allows an attacker to brute force user password and infiltrate the victim user account and gain access to users' data. This security vulnerability has not been resolved.
3	Use of HTTP on the password reset web page (identified by using Wireshark)	If an attacker sniffs network traffic when the user reset its password, the attacker would be able to access the password reset page and hijack the user's account. As observed in our study, this security vulnerability has not been resolved yet.

Our analysis could only find the three vulnerabilities, as listed in Table 5. Because all of these vulnerabilities were previously identified in the literature, we assumed that the set of vulnerabilities is saturated, and no further analysis is necessary at the moment.

4.5.2 Heuristic Evaluation I

For the first heuristic evaluation, we evaluated the Web browser-based application of privacy control for toy Alpha. We adopted the heuristics of Jaferian et al. (JAFERIAN *et al.*, 2014) as usability criteria to inspect the privacy control. As indicated in Salgado et al. (de Lima Salgado *et al.*, 2020), these are the most appropriate usability heuristics for inspections of parental privacy controls of smart toys. All of the potential usable security vulnerabilities are new (diagnosed in our study) and were not resolved yet. They are described in Table 6.

Table 6 – Alpha - Usability Problem and Reference.

#	**Usability Problem (Information Security Vulnerability)**	**Reference**
1	No alternative audio description: Users can only review the conversation content by listening to the audio files. Users may have to review large audio conversation files to identify a child's privacy breaches. This may be effortful for them.	Heuristic #1 - Visibility of activity status (JAFERIAN *et al.*, 2014)
2	Excessive visibility for recommended audios: Users may only review recommended audio conversation files because they are at the principal page of the Web application.	Heuristic #1 - Visibility of activity status (JAFERIAN *et al.*, 2014)

3	Repetitive security tasks: There is no clear way of identifying which audio conversation files have already been reviewed by users.	Heuristic #2 - History of actions and changes on artifacts (JAFERIAN *et al.*, 2014)
4	Poor visibility of privacy policies after login: Right after login, users are required to set up the child's information and connect the toy. During this task period, there is no indication of privacy policies ("*Provide rules and constraints*" [36]) if they need to review it.	Heuristic #4 - Rules and constraints (JAFERIAN *et al.*, 2014)
5	Lacking audio control: Users cannot control the audio execution ("*analyze historical information*" [36]). If they need to go to a specific part of the audio, they must listen to the entire audio until it.	Heuristic #2 - History of actions and changes on artifacts (JAFERIAN *et al.*, 2014)
6	Excessive deletion: Users unable to delete parts of the file ("*limit the awareness*" [36]) that may contain sensitive information of the audio conversation. Instead, they must delete the entire audio.	Heuristic #1 - Visibility of activity status (JAFERIAN *et al.*, 2014)
7	Poor keyboard navigation: Users may face difficulties to navigate using the keyboard ("allow the incorporation of a workflow").	Heuristic #5 - Planning and dividing work between users (JAFERIAN *et al.*, 2014)

As indicated in Table 6, most (three out of seven) of the usability problems relate to the Heuristic #1—Visibility of activity status (JAFERIAN *et al.*, 2014), followed by Heuristic #2—History of actions and changes on artifacts (JAFERIAN *et al.*, 2014) (two out of seven). To some extent, this was expected because usability heuristics are usually ordered according to its explanatory power (NIELSEN, 1994). Although we could perform the heuristic evaluation to identify the vulnerabilities, rating a severity for the findings was not an easy task. Because all of the issues are related to information security, highly important to the application, we could not rate the severity of problems employing the traditional severity scale as presented by Nielsen (NIELSEN, 1995). For this reason, in Section 5, we recommend the use of a revised severity scale, which we created to address the characteristics of usability problems in privacy control tools.

4.5.3 Heuristic Evaluation II

For the usable security evaluation of Toy Beta, we adopted its free mobile app for iOS devices. As for heuristic evaluation I, We adopted the heuristics of Jaferian et al. (WIRESHARK..., 1998) as usability criteria to inspect the usability of toy Beta privacy control. All of the potential human vulnerabilities are new (diagnosed in our study) and were not resolved yet. They are described in Table 7.

Table 7 – Beta-Usability Problem and Reference.

#	**Usability Problem (Information Security Vulnerability)**	**Reference**
8	Lacking help with password strength: There is no indication of password strength while users are creating it. This is necessary to support the "*freedom to choose different paths that respect the constraints*" (JAFERIAN *et al.*, 2014)	Heuristic #4 - Rules and constraints (JAFERIAN *et al.*, 2014)
9	Lacking indication of password requirements: There is no indication of password requirements (e.g., number of characters) while users are creating it.	Heuristic #1 - Visibility of activity status (JAFERIAN *et al.*, 2014)
10	Privacy Policy on the external website: The app opens its privacy policies in an external website, without providing any advertisement in advance to users.	Heuristic #4 - Rules and constraints (JAFERIAN *et al.*, 2014)
11	Lacking visibility for the privacy policy link: The privacy policy link receives less visibility than account information and the next button. Because this is a sensitive app, privacy policies should receive more visibility.	Heuristic #4 - Rules and constraints (JAFERIAN *et al.*, 2014)
12	Confusing user profile creation: The app does not indicate that the account (being created) belongs to the parents/guardians and not to their children.	Heuristic #5 - Planning and dividing work between users (JAFERIAN *et al.*, 2014)
13	Lacking privacy notice: The app does not inform users when sensitive child information is being sent to the Cloud.	Heuristic #1 - Visibility of activity status (JAFERIAN *et al.*, 2014)
14	Lacking cancelation of information sharing: The app does not provide an option to cancel (undo) information sharing. After users insert children's names and dates of birth, there is no alternative to cancel it before the app sends it to the Cloud.	Heuristic #2 - History of actions and changes on artifacts (JAFERIAN *et al.*, 2014)
15	Lacking information about the connection with mobile Artificial Intelligence (AI) assistant: The app offers a connection with mobile AI assistance, but there is no clear explanation of what information the assistant can access.	Heuristic #1 - Visibility of activity status (JAFERIAN *et al.*, 2014)
16	Menu lacking the option to manage a child's information: The app asks for both parents' and child's information, but there is no indication of where to manage the child's information after its insertion.	Heuristic #2 - History of actions and changes on artifacts (JAFERIAN *et al.*, 2014)

Source: Elaborated by the author.

As one can see, we diagnosed two times more usability problems with the privacy control of toy Beta in comparison with toy Alpha. This fact does not mean that the privacy control of toy Beta is worse than the privacy control of toy Alpha. As we understand, this is due to the fact

that toy Beta provides privacy control with more information about privacy policies. On the one hand, it is important to provide all the necessary information for users about their information privacy. On the other hand, this may implicate more problems related to Heuristic #4—Rules and constraints (JAFERIAN *et al.*, 2014). As indicated by Table 7, most of the problems found were related to the Heuristic #4—Rules and constraints (three out of nine problems) or to the Heuristic #1—Visibility of activity status (three out of nine problems). From these findings, we raise the question if privacy controls with more policy descriptions are prone to more situations that may contradict the Heuristic #4—Rules and constraints. Future research can investigate this topic. As in the heuristic evaluation of toy Alpha, the second most preferred heuristic in this evaluation was also Heuristic #2—History of actions and changes on artifacts. It seems that The first two heuristics of Jaferian et al. (JAFERIAN *et al.*, 2014) are, indeed, those with the highest explanatory power, justifying the order of heuristics.

4.6 Recommendations for the Design of Usable Privacy Controls

In this work, we list nine information security vulnerabilities. Six out of these nine vulnerabilities are retrieved from the literature, while the other three were identified by us in this work. These vulnerabilities are not usability related and motivate us to recommend attention for further development of smart toy privacy controls by means of: (i) do not use HTTP for transferring sensitive information; (ii) validate and sanitize input to avoid Cross-Site Scripting (XSS); (iii) require encrypted Wi-Fi; (iv) protect against remote brute force attacks on users' passwords; and (v) require authentication prior to privacy control.

Although these recommendations are important, we are not the first to reinforce the importance of them, since they are mostly based on the literature. On the contrary, all the 16 human vulnerabilities discussed in this study comes from our study. From these findings, we raised the question if privacy controls with more policy descriptions are prone to more situations that may contradict the Heuristic #4—Rules and constraints. Future research can investigate this topic. From these vulnerabilities, we suggest recommendations to improve the usability of smart toy privacy controls. These recommendations are a result of applying the heuristics of Jaferian et al. (JAFERIAN *et al.*, 2014) in the heuristic evaluations of this study. We present the recommendations in Table 8, along with its sources, which are the usability problems, as numbered (#) in tables 4 and 5, that justify the recommendations.

Table 8 – Recommendations to Enhance Usability and Privacy of Smart Toys.

Recommendation	**Usability Problem (#)**
Provide alternatives to efficiently perceive privacy controls. Users should not be obligated to interact with privacy controls by audio if they find the text more efficient to review information.	#1
Perception of control over the perception of information: The main focus of privacy controls should be on providing the perception of the control instead of providing the perception of the information collected. Users should not perceive excessive information competing with control options.	#2
Apply the Heuristic #2—History of actions and changes on artifacts [36] to provide users with the perception of which information is in accordance with users' control preferences.	#3
Provide privacy policies access at every screen and keep them consistent with the interface design.	#4, #10
Provide flexible controls. Users should be able to opt for fine-grained controls, such as deleting specific sections of the audio.	#5, #6
Provide efficient controls, such as supporting keyboard navigation for experienced users.	#7
Nudge users towards the creation of strong passwords.	#8, #9
Provide privacy notices about ongoing data sharing.	#11, #13
Clearly distinguish settings for children's information from parent's (or legal guardians') information. This is due to the need to provide information about the child, who is the smart toy user, and parents (or legal guardians) for authentication in the privacy control.	#12, #16
Provide clearly indicated alternatives to undo unwanted data sharing. This is to mitigate the consequences of laypeople giving wrong consents. Although this seems impossible, because we cannot affirm that the data has not been seen by anyone else, provide ways to request data deletion from a third party.	#14
Provide efficient control connection with artificial intelligence assistants. Users should know what information the assistant can access, and voice interactions should be human-like conversations.	#15

In the growing market of smart toys, security gets critical as users may be children and novices to the cyber world hidden behind the attractive toys. Because of the sensitive nature of children's personal information, a toy manufacturing company should design smart connected toys with security as a priority. Investment in robust security and continued updates to security measures are critical. Toy manufacturing companies should also apply acceptable data privacy practices such as a collection of only data that is required for the main operations of the smart

toy and to retain collected information for the only limited time that is necessary with valid reasoning. Our recommendations aim to support companies in the design process for better smart toy privacy controls.

In addition to our recommendations, and based on our experience with the case studies, we understand that a new severity rating scale is necessary to fully indicate the severity of usable privacy problems. We need a severity rating scale that represents privacy implications in it, along with usability issues. For this reason, we adapted Nielsen's severity scale (NIELSEN, 1995) to suggest the new usable privacy severity scale:

1. **Cosmetic:** usability problems, not related to policy generation/agreement, that may not stop users from using the interface.
2. **Minor:** usability problems, not related to policy generation/agreement, that may stop users from using the interface
3. **Major:** This leads to generating a wrong policy.
4. **Catastrophe:** leads to agreeing with the wrong policy.

Although new, our severity scale is based on Nielsen's (NIELSEN, 1995) traditional severity scale. This might influence practical aspects of employing our new scale because it keeps the span of four levels and lean descriptions. Nevertheless, future studies are still necessary to validate our severity scale in empirical experiments.

4.7 Conclusions

In this paper, we sought to provide recommendations aiming to enhance the usability and privacy of smart toy privacy controls. To this end, we identified security and usability problems that remain as vulnerabilities in popular smart toy applications. From nine information security vulnerabilities, which include the literature (see Table 4), we recommend five security practices in Section 5. From the 16 usability problems identified in our study, we composed a list of 11 usable security design recommendations to enhance the privacy aspects of smart toys, as presented in Table 8. Our recommendations may be used along different stages of design, from initial requirements to evaluation (testing stage) criteria. In addition, we also create a new severity scale focused on usability problems in the context of privacy policies.

Future studies may diagnose additional problems from similar IoT applications and evolve our recommendations towards a standard. They can also explore the use of our recommendations as criteria for usability inspections in the domain, which may include the revised severity rating scale.

Part III

Creating Privacy and Usability Criteria

CHAPTER

5

USABILITY HEURISTICS ON PARENTAL PRIVACY CONTROLS FOR SMART TOYS: FROM AN EXPLORATORY MAP TO A CONFIRMATORY RESEARCH

In this chapter, the author of this thesis uses his rights in Elsevier's proprietary journals to include the final published version of the article he co-authored. For this reason, hereby, we provide the full acknowledgment of the original article, as requested by Elsevier:

SALGADO, A. d. L.; FORTES, R. P. d. M.; OLIVEIRA, R. R. d.; FREIRE, A. P. Usability heuristics on parental privacy controls for smart toys: From an exploratory map to a confirmatory research. **Electronic Commerce Research and Applications**, v. 42, p. 100984, 2020. ISSN 1567-4223. Available: <https://www.sciencedirect.com/science/article/pii/S1567422320300612>.

The author of this thesis followed the instructions for author rights in Elsevier's proprietary journals given by Elsevier at "*Copyright*" on July 15th, 2021, at: <https://www.elsevier.com/about/policies/copyright#Author-rights>.

5.1 Abstract

In this paper, we aimed to indicate usability heuristics for the design of parental privacy controls for smart toys. During a snowballing mapping process, we examined 589 candidate studies. Our mapping findings draw from 13 included studies and indicate the heuristics for IT Security Management, proposed by Jaferian *et al.*, as the best to address problems that affect laypeople's interaction with privacy policy tools. With the participation of 14 inspectors, we compared the effectiveness of Nielsen's and the IT Security Management heuristics in heuristic evaluations of a parental privacy control model for smart toys. The results show that the IT

Security Management heuristics have better coverage of usability problems (Kruskal-Wallis, p-value ≈ 0.01), which confirms the mapping findings. Future studies can compare these heuristics based on outcomes from test with users as a benchmark. Also, future studies can explore the creation of domain-specific heuristics for parental privacy controls for smart toys.

5.2 Introduction

Toys are "*any product or material designed or clearly intended for use in play by children under 14 years of age*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018a). They have been present in the daily life of human society for thousands of years; currently, they are part of the life of billions of individuals (RAFFERTY *et al.*, 2017). Fostering this market, smart toys that listen and interact with children have recently gained popularity (MAHMOUD *et al.*, 2018; VALENTE; CARDENAS, 2017a; MCREYNOLDS *et al.*, 2017a). According to a study by Juniper Research, sales of smart toys are expected to grow threefold and exceed $15.5 billion dollars by 2022 (Juniper Research, 2017).

Although traditional toys raise little concern for child's privacy in general, smart toys are able to collect users' contextual data (e.g. location and time) and physical activity (e.g. voice). Such data collection is needed by service providers so that smart toys can learn about users' behavior and provide personalized services (RAFFERTY *et al.*, 2017; KAUSHIK; JAIN; SINGH, 2018). Data sharing may have economic advantages (ACQUISTI; TAYLOR; WAGMAN, 2016). Yet, it raises important privacy concerns (RAFFERTY *et al.*, 2017; KAUSHIK; JAIN; SINGH, 2018). According to UNICEF (United Nations Children's Fund), one of the major regulatory mechanisms to protect children's privacy online is requiring parental consent prior to the processing of children's personal data. This includes the US law "*Children's Online Privacy Protection Act*" (COPPA) and the General Data Protection Regulation (GDPR), from the European Union. Countries such as South Africa and Spain have similar provisions, and the UK provides recommendations by the Information Commissioner's Office (Children's Commissioner, 2018). Under these circumstances, parents strive to protect children's privacy, and parental privacy controls are seen as a promising approach to solve the problem of undue exposure of children's information by using smart toys (RAFFERTY *et al.*, 2017).

Parental privacy control is a "*feature in a smart toy for the parents to restrict the content the children can provide to the toy*" (RAFFERTY *et al.*, 2017, p. 1227). As a privacy control, it aims at reducing the likelihood or the consequences of privacy risks, which may include specification of privacy policies (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO); INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC), 2011; PACI; SQUICCIARINI; ZANNONE, 2018; GARFINKEL; LIPFORD, 2014). However, usability of privacy policy specification remains one of the main challenges in usable privacy (GARFINKEL; LIPFORD, 2014; PACI; SQUICCIARINI; ZANNONE, 2018; BERTINO, 2016;

DE; ZEZSCHWITZ, 2016). Without usable tools, even experts are likely to misconfigure and to leave unwanted vulnerabilities in a smart toy (SASSE; SMITH, 2016). More usable privacy tools for laypeople are necessary (BERTINO, 2016), as many parents and legal guardians are not necessarily specialists in IT and on privacy settings in digital systems.

This study attempts to support the design of more usable parental privacy controls by enhancing its evaluation stage. We focused on one of the most popular usability inspection methods (LAZAR; FENG; HOCHHEISER, 2017), (ALONSO-RÍOS; MOSQUEIRA-REY; MORET-BONILLO, 2018), the Heuristic Evaluation method. Since usability heuristics may emerge as new domains arise (e.g. smart toys), domain-focused heuristics have been proposed to avoid that the usability of interactive systems in new domains be overlooked (HERMAWATI; LAWSON, 2016; SALGADO; RODRIGUES; FORTES, 2016; SALGADO; FREIRE, 2014).

In this paper, we sought to systematize existing knowledge on usability heuristics for the domain of lay privacy policy interfaces, and to identify those with the best effectiveness in inspections of parental privacy controls for smart toys. We mapped the literature driven by the question: *What usability heuristics best address usability problems that affect laypeople's interaction with privacy policy interfaces related to smart toys?* Conducting a snowballing procedure (WOHLIN, 2014), we examined 589 candidate studies and performed a systematic analysis on 13 included studies. Furthermore, we conducted an empirical case study with 14 usability inspectors to confirm the findings from the literature mapping.

The remaining of this paper is organized as follows. Section 5.3 presents the background of this study, including terms and definitions and an overview on smart toys and parental privacy controls. In Section 5.4, we describe the research design of this study, indicating how we have joint the mapping study with the empirical case study to answer our question and reach our goal. Section 5.5 details the snowballing procedures of the literature mapping, indicating its cumulative outcomes. In Section 5.6, we present the mapping results and discuss them according to criteria from the literature. We then indicate the most appropriate heuristic set to answer the question, according to the mapping results. Therefore, we describe a case study conducted to confirm the answer. Finally, in Section 5.8 we sum up the conclusions of this study and indicate important topics for future research usable parental privacy controls for smart toys.

5.3 Background

5.3.1 Usable Privacy

Usable privacy is the research field aimed at studying the usability of systems that help end-users or administrators to manage data privacy (DE; ZEZSCHWITZ, 2016; GARFINKEL; LIPFORD, 2014). The ISO/TR 18638 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018c) defines *information privacy* as "*rights and obligations of individuals*

and organizations with respect to the collection, use, retention, disclosure and disposal of personal information". As indicated, Personally Identifiable Information (PII) is a key term in such definition, and is presented by the ISO/IEC 291000 (2011) as:

> Any information that (a) can be used to identify the Personally Identifiable Information (PII) principal to whom such information relates, or (b) is or might be directly or indirectly linked to a PII principal.

Meanwhile, the ISO 9241-11 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2010) defines usability as:

> the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

The interest on usable privacy had a rapid development during the past two decades (STILL, 2016; DE; ZEZSCHWITZ, 2016; GARFINKEL; LIPFORD, 2014; CRANOR; BUCHLER, 2014). First, both industry and academic literature have seen usability and security (including privacy) only as antagonistic, and users seen as the greatest risk to information security (STILL, 2016). Later, because the range of potential threats increased due to the pervasiveness of data (BERTINO, 2016), laypeople were often required to make security decisions (JANG-JACCARD; NEPAL, 2014) and were seen as the "*greatest hope*" for the area (STILL, 2016). Without usable tools, even experts are likely to misconfigure and to leave vulnerabilities in systems (SASSE; SMITH, 2016). In such cases, security breaches may be attributed to designers rather than to laypeople (WASH; ZURKO, 2017). Nevertheless, even fundamental concepts of laypeople's interaction with privacy tools, as mental models (OATES *et al.*, 2018), remain rare in the literature. The usable privacy field still needs usable tools for laypeople (BERTINO, 2016); also, it still needs appropriate usability methods for this domain. In the meantime, governments have been proposing different privacy-related regulations regarding the use of cloud-connected devices, such as smart toys.

To serve as examples, we briefly describe two legislative regulations, comparing them with traditional usability principles, the usability heuristics of Nielsen (NIELSEN, 2018a). The two legislations are the Children's Online Privacy Protection Rule (COPPA), and the General Data Protection Regulation (GDPR). We chose these regulations because they require parental consent regarding children's data privacy.

In 1998, the United States Congress enacted COPPA(Federal Trade Commission ("FTC" or "Commission"), 2013). By the end of 2012, the United States Congress issued an amended Rule to COPPA, which considered new categories of information (particularly used by connected devices, as geolocation, usernames, child's photos and videos, and persistent Web identifiers) to

their definition of personal information[1]. COPPA current version is dated from January 17, 2013 (Federal Trade Commission ("FTC" or "Commission"), 2013). In 2016, the European Parliament and its Council created the GDPR (THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, 2016). Its latest update dates back from April 27th, 2016. The GDPR became applicable in all European Union Member States on May 2018 (UNION; EUROPARAT, 2018, p. 17). By reviewing COPPA and GDPR we can identify usability criteria that impact the design of these technologies based on its similarity with some of the traditional ten usability heuristics from Nielsen. These heuristics are broad usability rules created from sets of known usability problems and their potential solutions (NIELSEN, 1994); they are criteria for usability evaluation (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2010). Examples of relations between regulations and usability criteria are, but not limited to:

Relevance of information: The GDPR principles require that personal data shall be "*adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')*"(THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, 2016). Meanwhile, COPPA requires that a privacy notice "*(...) must contain no unrelated, confusing, or contradictory materials*"(Federal Trade Commission ("FTC" or "Commission"), 2013). From our understanding, these requirements relate to Nielsen's 8th usability heuristic "*Aesthetic and minimalist design*", which implies that "*dialogues should not contain information which is irrelevant or rarely needed (...)*" (NIELSEN, 2018a).

Dealing with human error: GDPR requires that personal data shall be "*accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy')*" (THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION, 2016). Similarly, COPPA states that parents " *(...) may refuse to permit the use, and require the deletion, of the information collected*" and the interface must provide a way in which parents can do so (Federal Trade Commission ("FTC" or "Commission"), 2013). As one can compare, these requirements are related to users' control over information availability and accuracy. We understand that these requirements relate to Nielsen's 3rd usability heuristic "*User control and freedom*", which requires that "*Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue (...)*".

Usable privacy is a relatively new field, and privacy regulations state premisses that might relate to usability evaluation criteria (e.g., usability heuristics). Usability evaluation

[1] <www.ftc.gov/tips-advice/business-center/guidance/complying-coppa-frequently-asked-questions#General%20Questions>

criteria support evaluators' judgment to diagnose usability defects[2] on user interfaces. In regards to usable privacy, these defects may implicate on privacy breaches and vulnerabilities (JØSANG *et al.*, 2007; SASSE; SMITH, 2016). The next session describes the Heuristic Evaluation, which is the traditional method to employ usability heuristics as criteria for usability evaluations.

5.3.2 Heuristic Evaluation

Heuristic Evaluation (HE) is a formative usability evaluation method, which aims at diagnosing usability problems at an interface (LEWIS, 2014). The method was proposed by Nielsen and Molich (NIELSEN; MOLICH, 1990b). It does not involve the participation of potential users; instead, it involves multiple inspectors that compare the interface against a list of usability heuristics (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2010).

Although similar to interface guidelines, heuristics are usually small sets of broadly described usability principles, normally around ten heuristics. Therefore, HEs tend to take less time to perform than guidelines review (LAZAR; FENG; HOCHHEISER, 2017). The ten heuristics of Nielsen and Norman are the most traditional set of heuristics employed in HEs. Their titles and descriptions are available at Nielsen Norman Group website (NIELSEN, 2018a).

Nielsen's revised heuristics (NIELSEN, 1994) are traditional in the field. Nevertheless, Nielsen argues that the use of domain specific heuristics may benefit HEs (NIELSEN, 1994). Specific domain heuristics have been developed to increase the chances that the usability of software of such domains was not overlooked (HERMAWATI; LAWSON, 2016; SALGADO; RODRIGUES; FORTES, 2016; SALGADO; FREIRE, 2014). This is particularly important in the context of Internet of Things (IoT), which raises a variety of new domains to the software industry.

There are different types of usability criteria in the literature. The ISO/IEC 25066 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a) standard shows that usability criteria include user requirements, guidelines, conventions, style guides, task models and standardized principles. Other terms may be also used as synonym of heuristic. Nevertheless, we agree with Lazar *et al.* (LAZAR; FENG; HOCHHEISER, 2017), and understand that a usability criterion may be considered as a heuristic if it is part of a short criteria set (usually no more than ten), and is made to be employed in a HE (shorter in time). In this study, we aimed to find usability heuristics that can be employed in the usability inspection of privacy policy interfaces in the context of smart toys.

[2] As per ISO/IEC 25066 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016a)

5.3.3 Smart Toys, Children's Privacy Protection and Privacy Policies

Smart Toys

Smart toys enhance traditional toys' capabilities with computing services, empowering toys towards ubiquitous computing (ALBUQUERQUE; KELNER, 2018). After an extensive survey of the literature, Albuquerque *et al.* (ALBUQUERQUE *et al.*, 2019) shows that there is still no consensus for smart toy terminology. Besides smart toys, they are also called connected toys, interactive toys, toy computing and Internet of Toys (IoToys) (ALBUQUERQUE *et al.*, 2019). Despite of all terminologies, these toys can listen and interact with children in new ways. They have recently gained popularity in the market (MAHMOUD *et al.*, 2018; VALENTE; CARDENAS, 2017a). In this study, we adopt the term smart toy, which is defined by Hung *et al.* (HUNG; TANG; KANEV, 2017, p. 1) as:

> ... a device consisting of a physical toy component that connects to one or more toy computing services to facilitate game play in the Cloud through networking and sensory technologies to enhance the functionality of a traditional toy.

Examples of smart toys are *CogniToys Dino*, *Cue the Robot* and *Sphero BB-8 Robot*; they can talk with children, according to the Mozilla's website "**privacy not included*" (MOZILLA, 2018). Another example is *Hello Barbie*, one of the first smart toys in the market. It is a cloud-connected doll that can listen to child's questions and answer them with a cloud-based mechanism. Due to child's privacy concerns, a campaign called "*Hell No Barbie*" was carried out in 2015 by an American advocacy group [3]. To clarify this issue, the Children's Commissioner report (Children's Commissioner, 2018) describes that there are many ways in which child's data, collected by smart toys, can reach wrong hands. As an example, they argued that hackers can gain control over smart toys and talk to children through such devices.

Indeed, as companion robots, smart toys challenge the effectiveness of current privacy regulatory mechanisms (BERTOLINI; AIELLO, 2018) and children privacy protection become a core concept for researches in the field (Chu; Apthorpe; Feamster, 2019; YANG; LU; WU, 2018; HUNG; FANTINATO; ROA, 2018; ALBUQUERQUE; KELNER, 2018; VALENTE; CARDENAS, 2017b; HUNG; IQBAL; HUANG, 2016; ALBUQUERQUE; KELNER, 2019).

Children's Privacy Protection

Children's privacy protection refers to enabling "*parents or guardians to be in control of their children's privacy by specifying their privacy preferences for a toy*" (HUNG; FANTINATO; ROA, 2018, p. 1). Also, it is crucial to assume that (RAFFERTY *et al.*, 2017, p. 1227): (i) *children do not understand the concept of privacy and do not know to protect themselves online*, and (ii) *children may disclose personal information to smart toys and not be aware of*

[3] CBC News at: <https://www.cbc.ca/news/business/hello-barbie-1.3292361>

possible consequences and liabilities. In this context, parents and guardians strive to protect children's privacy (RAFFERTY *et al.*, 2017), but there is no universal approach for them to control children's privacy regarding smart toys (STREIFF; DAS; CANNON, 2019; RAFFERTY *et al.*, 2017; XIA *et al.*, 2016). Parental privacy controls stand as a promising approach to provide such control and fill the gap, while privacy policies are an essential part of their mechanism (HUNG; FANTINATO; ROA, 2018; CUNHA; Unicef; others, 2017; UNICEF Innovation, 2019; SALGADO *et al.*, 2017).

Privacy Policies for Smart Toys

Privacy policies are documents that indicate users' or applications' preferences about data privacy (GARFINKEL; LIPFORD, 2014; OATES *et al.*, 2018). Such preferences indicate the users' choices regarding "*(...) how their PII should be processed for a purpose*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO); INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC), 2011). In other words, privacy policies relate data owners to potential data readers (JIANG; LANDAY, 2002), and usually control the access of these readers to the data, according to the owner's preferences (KELLEY *et al.*, 2009).

In the context of smart toys, parental privacy controls are a promising engine to protect children's privacy (HUNG; FANTINATO; ROA, 2018; CUNHA; Unicef; others, 2017; UNICEF Innovation, 2019; SALGADO *et al.*, 2017). Parental privacy controls are features "*for the parents to restrict the content the children can provide to the toy*" (RAFFERTY *et al.*, 2017, p. 1227). As a type of privacy control, they should have mechanisms for policy generation and interfaces for policy comprehension, policy configuration, and feedback (PACI; SQUICCIARINI; ZANNONE, 2018). Moreover, parental privacy controls must have published accurate privacy policies to ensure children's protection (HUNG; FANTINATO; ROA, 2018). For this reason, Yankson *et al.* (YANKSON; IQBAL; HUNG, 2017) proposes a privacy preservation framework that supports context-dependent policies based on eXtensible Markup Language (XML). Later on, Yankson *et al.* (YANKSON *et al.*, 2019) proposes the use of Petri-Nets, a mathematical modeling language, to model, test and verify context-dependent policies.

Privacy policies are usually complicated because they need to represent laws, regulations, and business practices (SCHAUB; BALEBAKO; CRANOR, 2017). Most of the policies only contain a few options to opt-out and lack details on contextual factors, failing at being effective privacy notice and choice (SCHAUB; BALEBAKO; CRANOR, 2017; APTHORPE; VARGHESE; FEAMSTER, 2019). Also, parents may not read these policies nor understand it (KEYMOLEN; HOF, 2019), and they have no means to verify whether the smart toys follow those policies (Chu; Apthorpe; Feamster, 2019). Using *Polisis* (HARKOUS *et al.*, 2018), a Deep Learning Automated Analysis and Presentation of Privacy Policies, one can see that privacy policies related to smart toy companies are still described with generic security statements,

making it difficult for users to understand and chose[4].

Indeed, the usability of privacy policy tools, such as parental privacy controls, challenge the literature (GARFINKEL; LIPFORD, 2014; PACI; SQUICCIARINI; ZANNONE, 2018; BERTINO, 2016; DE; ZEZSCHWITZ, 2016; SALGADO *et al.*, 2017). Although researchers have explored interface alternatives that enhance the usability of privacy policy interfaces (PACI; SQUICCIARINI; ZANNONE, 2018), studies related to parental privacy control for smart toys remain a few. In a previous study of our research group (SALGADO *et al.*, 2017), we discussed how traditional HCI methodologies could be applied to enhance the usability of parental privacy controls. In that study, we also proposed initial updates to the design of Rafferty *et al.*'s (RAFFERTY; FANTINATO; HUNG, 2015) model of parental control aiming to enhance its usability. In a follow-up study (SALGADO *et al.*, 2019a), our group proposed a re-design of Rafferty *et al.*'s parental control model based on Kelley et al.'s (KELLEY *et al.*, 2009) nutrition label model, and a card sort experiment enhanced with cluster analysis. The re-design also contained elements from Google Material Design[5], aiming to enhance usability by interface elements that are familiar to users.

5.4 Research Design

In this paper, we aimed to indicate usability heuristics for the design of parental privacy controls for smart toys. Because parental control for smart toys is a narrow domain, we also evaluated the literature on privacy policy interfaces that are related to smart toys. For this reason, our research design was twofold: we conducted a mapping study (exploratory phase), and an empirical case study (confirmatory phase).

Our mapping study aimed to identify the most appropriate usability heuristics to address usability problems in the domain (broad) of privacy policy interfaces for laypeople. Our empirical case study aimed to confirm the mapping findings in the domain (narrow) of parental privacy controls for smart toys. We performed the case study by empirically comparing the heuristics indicated by the mapping against the traditional usability heuristics of Nielsen. By performing this analysis, we expected to evaluate which set of heuristics was the most appropriated to help professionals in the performance in usability inspections when applied to the specific scenario of parental privacy control.

5.5 Literature Snowballing Procedure

This literature mapping performed in the present study followed the snowballing procedure described at Wohlin (WOHLIN, 2014). The procedure starts with a brief literature search,

[4] We employed <pribot.org/polisis> tool (security tab) to review <ToyTalk.com> and <zenbo.asus.com> privacy policies

[5] <material.io/design/>

which aims at identifying a start set of studies to begin the snowballing. After identifying the start set, the snowballing cycle began. This cycle was divided in two stages: backward and forward. The backward stage evaluates the references indicated at candidate studies, while the forward stage evaluates citations of the candidate studies. These evaluations apply a set of inclusion and exclusion criteria. The snowballing cycle is repeated until saturation, when no new candidate study is identified.

The major advantages of a snowballing procedure is that the backward stage, in most cases, lead straightforward to identify relevant papers. Similarly, using Google Scholar and its citation tracking, the forward stage can be quite informative and helpful to decide about including/excluding a paper (WOHLIN, 2014). These characteristics make the snowballing a process which may be focused on the identification of relevant papers rather than measuring the literature. For this reason, the scope of this study was to identify relevant papers that may answer the research question, and then to conduct a follow-up empirical study to analyze the result.

5.5.1 Defining the Start Set

We began the mapping with seven (7) candidates for the start set (JØSANG *et al.*, 2007; JØSANG; ZOMAI; SURIADI, 2007; NURSE *et al.*, 2011; YERATZIOTIS; POTTAS; GREUNEN, 2012; GARFINKEL; LIPFORD, 2014; JAFERIAN *et al.*, 2014; REALPE *et al.*, 2016). Two researchers independently identified these studies because they could potentially answer the research question. To evaluate these candidates and to perform backward and forward stages, we defined the inclusion/exclusion criteria as indicated at Table 9. For the whole procedure, we first compared candidate papers against the exclusion criterion (E1) and, after, against each of the inclusion criteria (I1-I4). Therefore, we only accepted candidate papers that were not excluded by E1 and were included by all inclusion criteria together (I1 AND I2 AND I3 AND I4). We did not define any time frame as inclusion or exclusion criteria; studies from any publication year could be accepted.

Table 9 – Inclusion and exclusion criteria.

Inclusion Criteria
I1. The abstract provides indication of proposal of new usability heuristics for the privacy and security domain.
I2. The full text shows the list of the proposed usability heuristics.
I3. Published in peer-reviewed journals or conferences, or book chapters with editorial boards.
I4. The study **is not** a work in progress or similar unfinished study.
Exclusion Criteria
E1. The study **is not** available in English.

Source: Research data.

We evaluated the seven candidates to the start set against the inclusion/exclusion criteria.

Because of criterion I2, we did not accept one candidate (Jøsang, Zomai and Suriadi (2007)), which is referred in this paper as C1. All other candidate papers were accepted and formed the start set as follows:

(S1) Jøsang *et al.* (2007)
(S2) Nurse *et al.* (2011)
(S3) Yeratziotis, Pottas and Greunen (2012)
(S4) Garfinkel and Lipford (2014)
(S5) Jaferian *et al.* (2014)
(S6) Realpe *et al.* (2016)

From the start set studies (S1-S6), we began the snowballing cycles. We evaluated their references (backward snowballing), and used a Google Scholar mechanism to identify their citations for further evaluation (forward snowballing). We used Google Scholar to avoid publisher bias, as suggested by Wohlin (2014). We performed these evaluations from August 2nd, 2018 to August 8th, 2018.

5.5.2 Iteration 1

At iteration 1, we performed both backward and forward snowballing with the start set. This iteration involved 347 publications (121 from the backward evaluation and 226 from the forward evaluation). Particularly, S2, S3 and S4 are secondary studies, and already synthesized usability heuristics from previous studies. Because our goal was to identify and systematize such heuristics, we did not perform the backward snowballing for S2, S3 and S4.

Backward snowballing

Table 10 indicates the number of candidate studies in this snowballing phase. The first column (*Study*) indicates the source studies for this backward stage. Column *References* shows the number of references retrieved from the respective study/row. Column *Duplicates* indicate whether we identified, and removed, any duplicate from the references. Columns *I1-E1* show the number of references that did not satisfy any of the inclusion/exclusion criteria. Finally, column *New*[6] shows whether any new study satisfied all inclusion/exclusion criteria and composed our mapping.

As indicated at Table 10, this backward stage resulted in the inclusion of one study: **(S7)** Katsabas, Furnell and Dowland (2005).

The following subsection presents the outcomes from the forward phase of iteration 1.

[6] $New = References - Duplicates - I1 - I2 - I3 - I4 - E1$

Table 10 – Number of candidate studies involved along *iteration 1 backward snowballing* and new mapped studies.

Study	References	Duplicates	*I1*	*I2*	*I3*	*I4*	*E1*	New
S1	21	1 (C1)	20	0	0	0	0	0
S2[1]	-	-	-	-	-	-	-	0
S3[1]	-	-	-	-	-	-	-	0
S4[1]	-	-	-	-	-	-	-	0
S5	74	0	73	1	0	0	0	0
S6	26	2 (S2 and S5)	22	0	1	0	0	1 (S7)
Totals	121	3 (C1, S2 and S5)	115	1	1	0	0	1 (S7)

[1] S2, S3 and S4 are literature reviews and synthesized usability heuristics from their references. Because our goal was to identify and systematize such heuristics, we did not perform the backward iteration for them.

Source: Research data.

Forward snowballing

Table 11 indicates the number of candidate studies for this forward snowballing. It has the same structure as Table 10, despite that it shows a *Citations* column instead of references.

Table 11 – Number of candidate studies involved along *iteration 1 forward snowballing* and new mapped studies.

Study	Citations	Duplicates	*I1*	*I2*	*I3*	*I4*	*E1*	New
S1	69	1 (S2)	61	1	1	0	5	0
S2	49	1 (S6)	46	0	1	0	0	1 (S8)
S3	14	0	10	2	0	0	0	2 (S9 and S10)
S4	44	0	38	2	0	1	3	0
S5	49	4 (S2, S4, S6 and S9)	41	1	0	0	2	1 (S11)
S6	1	0	0	1	0	0	0	0
Totals	226	6	196	7	2	1	10	4 (S8-S11)

Source: Research data.

As indicated at Table 11, we included the following studies from this forward snowballing:

(S8) Fierro and Zapata (2016)

(S9) Gumussoy (2016)

(S10) Yeratziotis, Greunen and Pottas (2011)

(S11) Reynaga, Chiasson and Oorschot (2015)

At this point, we finalized iteration 1. The following section describes iteration 2, performed with the studies included during this iteration (S7-S11).

5.5.3 Iteration 2

We performed iteration 2 with the five studies included during iteration 1 (S7-S11). This iteration evaluated 191 candidate papers (145 from the backward evaluation and 46 from the forward evaluation).

Backward snowballing

Table 12 indicates the number of candidate studies along this snowballing phase. Among the five evaluated studies (**S7-S11**), only S8 returned a new study to the mapping.

Table 12 – Number of candidate studies involved along *iteration 2 backward snowballing* and new mapped studies.

Study	References	Duplicates	*I1*	*I2*	*I3*	*I4*	*E1*	New
S7	14	0	13	0	1	0	0	0
S8	20	1 (S2)	17	0	1	0	0	1 (S12)
S9	37	1 (S3)	35	0	1	0	0	0
S10	17	0	17	0	0	0	0	0
S11	57	0	55	0	2	0	0	0
Totals	145	2 (S2 and S3)	137	0	5	0	0	1 (S12)

Source: Research data.

This backward snowballing added one study to the mapping, denoted as: **(S12)** Paz *et al.* (2014).

The following section presents the forward snowballing of iteration 2.

Forward snowballing

Table 13 indicates the number of studies involved along this snowballing phase. Similarly to iteration 2 backward, only S8 returned a new study to the mapping.

Table 13 – Number of candidate studies involved along *iteration 2 forward snowballing* and new mapped studies.

Study	Citations	Duplicates	*I1*	*I2*	*I3*	*I4*	*E1*	New
S7	15	1 (S6)	14	0	0	0	0	0
S8	2	0	0	0	0	0	1	1 (S13)
S9	14	0	11	0	0	0	3	0
S10	6	0	6	0	0	0	0	0
S11	9	0	7	1	0	0	1	0
Totals	46	1 (S6)	0	0	0	0	0	1 (S13)

Source: Research data.

This forward snowballing added one study to the mapping, denoted as: **(S13)** Díaz and Río (2018).

The next section presents the third iteration, performed after the outcomes of this one.

5.5.4 Iteration 3

We performed iteration 3 with the two new studies included after iteration 2 (S12 and S13). This iteration evaluated 61 candidate papers (42 from backward evaluation and 19 from forward evaluation).

Backward snowballing

Table 14 shows the number of studies involved at this backward snowballing. As indicated, this backward snowballing did not include any new study to the mapping.

Table 14 – Number of candidate studies involved along *iteration 3 backward snowballing* and new mapped studies.

Study	References	Duplicates	*I1*	*I2*	*I3*	*I4*	*E1*	New
S12	8	0	8	0	0	0	0	0
S13	34	2 (S8 and S12)	32	0	0	0	0	0
Totals	42	2 (S8 and S12)	40	0	0	0	0	0

Source: Research data.

Forward snowballing

Table 15 indicates the number of studies involved at iteration 3 forward snowballing. As shown, no study was included after this forward snowballing. Because both backward and forward stages of iteration 3 did not include any new study to this mapping, the snowballing procedure was finished. The following section presents the data extraction and the analysis.

Table 15 – Number of candidate studies involved along *iteration 3 forward snowballing* and new mapped studies.

Study	Citations	Duplicates	*I1*	*I2*	*I3*	*I4*	*E1*	New
S12	19	2 (S8 and S13)	13	0	0	0	4	0
S13	0	-	-	-	-	-	-	0
Totals	19	2 (S8 and S13)	0	0	0	0	0	0

Source: Research data.

At this phase, we concluded the collection of new studies for our mapping. The following section present and discuss the mapping results.

5.6 Results from the Mapping Study

To indicate usability heuristics for the design of parental privacy controls for smart toys, we mapped 13 studies from the literature (S1-S13). Each of these studies suggests usability criteria to employ in HEs of security and privacy systems. Our results include studies from 2005 up to 2018. Because we did not limit a time frame for this mapping, these results suggest that the first study to propose usability principles for the security and privacy domain dates back to 2005. Therefore, it indicates that this research topic may have 14 years, which would represent an average of almost one mapped study per year. Nevertheless, eight out of the 14 studies dates from 2014 to 2019, which may indicate an increased interest on the topic. Figure 13 indicates the mapped studies along the timeline of this research topic.

Figure 13 – Timeline of usability principles for security and privacy tools.

Source: ©Elsevier's Electronic Commerce Research and Applications, 2020

We extracted data from the 13 studies and evaluated against the themes proposed by Hermawati and Lawson (2016, p. 35) to assess domain specific usability heuristics. The themes (T#) are: "*Adequacy of the domain*" (T1), "*Creation process*" (T2), "*Validation*" (T3), "*Adequacy of heuristics' description*" (T4) and "*Effectiveness*" (T5). These five themes guided our discussion of the results, and the evaluation of the research question.

5.6.1 Adequacy of the domain

Among the 13 included studies, we identified heuristics for eight different domains: *Usable privacy and security*, *Banking software*, *Access control*, *Information Technology Security Management (ITSM)*, *User authentication*, *Online Health Social Network (OHSN)*, *captchas on smartphones (Completely Automated Public Turing test to tell Computers and Humans Apart*[7]*)*

[7] Definition retrieved from: <dictionary.cambridge.org/dictionary/english/captcha>

and *transactional websites*. Table 16 relates these domains with their respective studies. As indicated, none of the studies proposed heuristics for lay privacy policy and smart toys.

Table 16 – Domains identified from mapped studies.

Domain	**Study**
Usable privacy and security	S1, S2, S3 and S7
Banking software	S8, S9 and S13
Access control	S4
Information Technology Security Management (ITSM)	S5
User authentication	S6
Online Health Social Network (OHSN)	S10
Captchas	S11
Transactional websites	S12

Source: Research data.

In a broad sense, we found generic usability principles for the cyber-security domain (S1, S2, S3 and S7) and for cyber-security subdomains (S4, S5, S6, S8, S9, S10, S11, S12 and S13). Among the usability principles for cyber-security subdomains, S4 and S5 have closest focus on the domain of privacy controls. S4 proposes heuristics for access controls, which require users to view and author privacy policies and is close related to our domain of interest (GARFINKEL; LIPFORD, 2014). Meanwhile, S5 proposes heuristics for ITSM. ITSM involve different user profiles, such as software developers, security auditors and lay users. Therefore, these heuristics may also be appropriate for the domain of lay privacy policy interfaces and smart toys. With regards to this theme, it may indicate that S4 and S5 sets are more appropriate to evaluate usability of privacy policy interfaces.

5.6.2 Heuristics Definition process

Among the 13 studies (S1-S13), we identified ten different definition methodologies for the sets of heuristics. The method proposed by Rusu *et al.* (2011) is predominant among these studies (S6, S8, S12 and S13). This method comprises six steps to establish usability heuristics: exploratory, descriptive, correlational, explicative, validation and refinement. It is important to notice that S6 only adopted the first four (4) of these steps (REALPE *et al.*, 2016), neither validation nor refinement steps were performed.

Meanwhile, the method proposed by Yeratziotis, Pottas and Greunen (2011) is the second most referred. Both **S3** and **S10** adopted this methodology. The method of Yeratziotis *et al.* is a three-phase process. Phase 1 is to design high-level heuristics, phase 2 is the validation of the high-level heuristics and phase 3 the application of these high-level heuristics. These are not cascade phases, and researchers can go back and forth among them. It is just not possible to go straight from phase 1 (design) to phase 3 (application). This method is not only similar to Rusu *et al.* (2011)'s, but also more recent. Such similarity indicates that there is a need for a standardized

methodology to create usability heuristics in the field. Yet, it reinforces the recommendations of Hermawati and Lawson (2016) that the creation of usability heuristics should contemplate validation and refinement stages.

Six studies defined their own method to create their heuristics. Three of them (**S5**, **S7** and **S9**) conducted some type of validation during the process. In **S5**, Jaferian *et al.* defined their own method based on grounded theory techniques (JAFERIAN *et al.*, 2014). They performed a top-down methodology to justify, support and combine design guidance into new usability heuristics. They also performed empirical validation and refinement stages.

At **S7**, Katsabas, Furnell and Dowland (2005) adapted Johnston, Eloff and Labuschagne (2003)'s usability criteria together with Nielsen's heuristics to compose their own. During this process, they also considered aspects of the cyber-security domain and the first principles of interaction design (TOGNAZZINI, 2014). This process was, in some extent, subjective and centered on the author's knowledge about the cyber-security domain.

In study **S9**, Gumussoy (2016) adapted the heuristics of Muller *et al.* (1998). She evaluated the coverage of these heuristics against a database of three banking software projects. She also performed a cluster analysis to group heuristics according to its coverage of usability problems by severity.

The remaining three studies (**S1**, **S2** and **S4**) defined their own method to create the heuristics, without validation stages. **S1** reports limited information about how their principles were created (JØSANG *et al.*, 2007), but informed that such principles were created from the Kerckhoffs' principles for identity management (KERCKHOFFS, 1883). **S2** reviewed the literature on usability recommendations for cyber-security systems and consolidated a set of 19 guidelines (NURSE *et al.*, 2011). Nurse *et al.* (2011) grouped similar recommendations into unique guidelines, and renamed such guidelines in accordance to the recommendation content. **S4** also summarized previous literature to compose their set of guidelines (GARFINKEL; LIPFORD, 2014). Their process is described as a subjective analysis of the authors about the literature reviewed. After, they summarized lessons learned from the literature as guidelines to create usable access control mechanisms.

To sum up, nine out of the 13 studies (S3, S5, S7, S8, S9, S10, S11, S12 and S13) performed some type of validation of their principles during its creation. As recommended by Hermawati and Lawson (2016), and supported by our results, including validation procedures during the creation of usability principles may enhance its quality. For this reason, these nine studies stand as more appropriate than the others to be employed in further HEs in the privacy and security domain. The following section presents further description of validation processes among the mapped studies.

5.6.3 Validation

Nine of the studies reported some type of validation (S3, S5 and S7-S13) when proposing their new heuristics. Five of them (S5, S8, S11, S12 and S13) empirically compared their heuristics against Nielsen's. The other four (S3, S7, S9 and S10) performed alternative validation methods. We understand that those validation procedures that empirically compared new heuristics against Nielsen's are more appropriate to our propose. The following subsections describe each of these methods.

Comparison with Nielsen's heuristics

Studies S5, S8, S11, S12 and S13 compared their heuristics with Nielsen's heuristics. At **S5**, Jaferian *et al.* (2014) designed a between-subjects study with 28 participants (inspectors), equally divided between two groups. Both groups evaluated the same interface, from one identity management system. Most of the participants performed remote evaluation. All of them had, at least, HCI and computer security background, and previous experience with HE.

In **S8**, Fierro and Zapata (2016) describe a between group comparison. They do not inform the number of inspectors in each group, neither inspector background. In their case study, the subject for evaluation was one international bank website.

In **S11**, Reynaga, Chiasson and Oorschot (2015) conducted a between group study with 18 participants. They were divided in two groups of nine, each group performing HEs with different heuristics. One group employed Nielsen's heuristics, while the other employed the new heuristics for captchas on mobile phones. The evaluators were HCI experts, also familiar with computer security. Reynaga, Chiasson and Oorschot (2015) chose four captchas images to be evaluated during the HEs. These captchas were chosen because they represent the main captchas categories.

At **S12**, Paz *et al.* (2014) organized four HEs with different groups of inspectors. The study does not indicate the number of inspectors in each group. Nevertheless, three groups employed Nielsen's heuristics, while the other one employed the new heuristics. They evaluated one exemplar transactional website. The study does not provide information about inspectors' background.

S13 reports a validation process comparing their new heuristics against Nielsen's. At such study, Díaz and Río (2018) designed a between group comparison; one group applied Nielsen's heuristics while the other applied the new ones. A baking website was subject of inspection. The study does not inform the number of inspectors in each group, neither the inspectors' background.

Alternative validation methods

This section overview studies that did not compared their heuristics against Nielsen's (S3, S7, S9 and S10). Among these studies, S3 demonstrated the application of their heuristics, while the others (S7, S9 and S10) performed some kind of questionnaire-based validation. None of them described conducting a HE with the proposed heuristics.

At **S3**, Yeratziotis, Pottas and Greunen (2012) demonstrated the application of their heuristics against a set of examples. They did not count on any external evaluators to perform the validation. Also, they used one interface of Online Health Social Networking for the demonstrations.

During **S7**, Katsabas, Furnell and Dowland (2005) applied their principles in a questionnaire-based evaluation (not a HE). Inspectors were required to related each principle to a five-point scale ("*Application diverges completely from the guideline*" to "*Application completely follows the guideline in all possible sections*"). No description of inspectors' background was given. Ten interfaces were subject of evaluation. According to the authors, these interfaces could give an overall mix of both security-specific tools.

In **S9**, Gumussoy (2016) counted on three usability experts to rate the severity of 266 known usability problems. These problems came from three banking software projects. After, the same experts indicated how well each of the new principles describe each of the problems. To perform this stage, the experts should answer a six-point scale questionnaire. This research design aimed to reveal the interaction between severity level of usability problems and each of the new principles. Also, it allowed the author to cluster their principles according to the severity of problems that each of the principles is related to.

Finally, at **S10**, Yeratziotis, Greunen and Pottas (2011) conducted a study with six evaluators. All evaluators were postgraduate students in the field of Information and Communication Technology. The evaluators were requested to inspect two online health social networks. They were provided with scenarios for the evaluators. After, the evaluators applied the new heuristics and rated a five-point scale questionnaire ("*Very Good*" to "*Very Poor*") to indicate the usability of the interface.

5.6.4 Adequacy of heuristics' description

This mapping identified 278 usability heuristics. Seven studies proposed sets with no more than ten heuristics (S4 = 5; S13 = 6; S5 = 7; S10 = 7; S11 = 7; S1 = 8; S7 = 10). Meanwhile, six studies proposed sets with more than ten rules (S3 = 13; S9 = 13; S8 = 15; S12 = 15; S2 = 19; S6 = 153). Although S6 presents a set of 153 principles, only 75 are considered by the authors as related to usability (REALPE *et al.*, 2016). Despite of the number of heuristics, studies also varied on how to describe them. Among the studies, we identified the three description styles:

Succinct: the description is composed by a title (or succinct description. This is the case of checkpoints and quick tips.

Traditional: the description is composed by a title and one or two paragraphs. We called this as traditional because it is the closest to the description of Nielsen's traditional heuristics.

Lengthy: the description is composed by a title and multiple checklist items (YERATZIOTIS; POTTAS; GREUNEN, 2012) or multiple usability criteria (GUMUSSOY, 2016). This style increases the total number of rules to be considered during a HE, which may be not desired.

Table 17 indicates each study and relates them to their respective description style and the number (N) of criteria. We identified three sets with succinct descriptions (S1, S4 and S6), eight with traditional descriptions (S2, S5, S7, S8, S10, S11, S12 and S13) and two sets with lengthy descriptions (S3 and S9). Because S3 and S9 have lengthy descriptions, they have an increased number of criteria. This would make the number of criteria in S3 increases from 13 to 86, and from 13 to 51 in S9. We understand that such large number is due to the lack of standardization on heuristic description styles. However, large number of heuristics may be inappropriate to be employed in HEs, which is often intended to be fast, compared to guideline reviews (LAZAR; FENG; HOCHHEISER, 2017).

Table 17 – Studies, description style and number of principles (N).

Study	**Description**	***N***
S1	Succinct	8
S2	Traditional	19
S3	Lengthy	86
S4	Succinct	5
S5	Traditional	7
S6	Succinct	75
S7	Traditional	10
S8	Traditional	15
S9	Lengthy	51
S10	Traditional	7
S11	Traditional	7
S12	Traditional	15
S13	Traditional	6

Source: Research data.

As indicated by Lazar, Feng and Hochheiser (2017), we understand that heuristic sets with no more than ten (10) heuristics are more appropriate to be employed in HEs, especially considering that inspectors are meant to memorize them and to apply them in the process, and not have to constantly go back to their descriptions as one would do in a review of guidelines.

With this characteristic, seven studies stand as more appropriate heuristic descriptions (S1, S4, S5, S7, S10, S11 and S13).

5.6.5 Effectiveness of the heuristics

Five studies (S5, S8, S11, S12 and S13) measured and discussed the effectiveness of their new criteria. Some studies adopted the traditional effectiveness metric, as showed by Hartson, Andre and Williges (2001), while others adopted alternative measures. S5 adopted the traditional effectiveness metric, and S11 provided enough information to calculate it as well. **S5** found that their heuristics had higher *f-measure* compared with Nielsen's heuristics (S5 = 0.80; Nielsen's = 0.72; $\alpha = 0.5$) (JAFERIAN *et al.*, 2014). **S11** argues that their heuristics resulted in more unique usability problems and less false positives than Nielsen's (REYNAGA; CHIASSON; OORSCHOT, 2015). Based on the values they reported, we calculated the *f-measure* of their heuristics. We found that it had higher *f-measure* than Nielsen's heuristics (S11 = 0.82; Nielsen's = 0.73; $\alpha = 0.5$). These results are close to the results shown by S5.

On the other hand, S8, S12 and S13 adopted alternative measures. **S8** argues that their new criteria resulted on the uncovering of 46% of unique problems, while Nielsen's resulted in 34% (FIERRO; ZAPATA, 2016). **S12** presents a case study where their new heuristics found more usability problems than Nielsen's. However, they could not make any inference about the observed advantage. **S13** shows that Nielsen's heuristics were more efficient (discovery of usability problems by time) them their new heuristics (DÍAZ; RÍO, 2018). Although, they argued that their heuristics resulted in a higher number of usability problems (56%) than Nielsen's (40%). Also, they showed that their heuristics resulted in identification of more severe problems.

In summary, the mapped studies indicated that Nielsen's heuristics have an effectiveness about 0.72 and 0.73 for HEs in the privacy and security domain. Meanwhile, they indicated that domain-focused heuristics had effectiveness rates of about 0.80 and 0.83 for HEs in the privacy and security domain.

5.7 Comparison Between Domain-Specific and General Usability Heuristics on the Inspection of Parental Privacy Control of Smart Toys

To confirm the findings from the mapping study with preliminary empirical evidence, we conducted a case study with 20 participants [8]. We sought to compare the diagnosis of usability problems between Nielsen's heuristics and the ITSM heuristics.

[8] This study was approved by a Research Ethics Committee with CAAE code 69353317.4.0000.5390

To this end, we invited 20 novice usability inspectors based on feasibility analysis (CAINE, 2016). They were all undergraduate students in Computer Science that, at the time this study was conducted, had just completed their Human-Computer Interaction course. We decided to invite such novice inspectors because experts are rare to find and have limited available time to participate as voluntary in our study. Studying the performance of novice inspectors may help the field to enhance their performance in HEs, which may benefit the organizations that relies on their performance. Therefore, the quasi-experimental design (between subjects) is defined as presented at Table 18.

Table 18 – Planned quasi-experimental design (between subjects).

Nielsen's condition	***ITSM condition***
10 participants	10 participants

Source: Research data.

Based on the parental control model proposed by Rafferty *et al.* (RAFFERTY; FANTINATO; HUNG, 2015), we designed a prototype be analyzed at the inspections. To populate the prototype with real world information, we used the information from an online Brazilian toy store [9]. We made the prototype in Portuguese, because the participants were all native speakers of Portuguese (Brazilians). All participants had one hour to complete the inspection and deliver their individual list of problems. They were not required to perform group discussion after the HE. To ensure that participants would find a minimum number of usability problems to diagnose, we introduced known usability problems (incorrect words and disabled fields) to the prototype ("*seeding known usability problems*" (HARTSON; ANDRE; WILLIGES, 2001, p. 384)).

Although the 20 participants voluntarily accepted to participate in our study and began the inspection, six of them decided to quit their participation before completing the inspection. For this reason, we did not include their inspections' results. This fact left us with the feasible quasi-experimental design as presented at Table 19.

Overall, the six participants that used Nielsen's heuristics diagnosed 24 usability problems ($Min. = 5; Median = 7; Mean = 9.667; sd \approx 5.0; Max. = 17$). After a Shapiro-Wilk normality test ($p - value \approx 0.038$), we found that these results may have been drawn from a normal distribution. Meanwhile, the eight participants that used the ITSM heuristics diagnosed 29 usability problems ($Min. = 8; Median = 11.50; Mean = 13.12; sd \approx 4.36; Max. = 19$); and a Shapiro-Wilk normality test ($p - value \approx 0.13$) has shown that we cannot assume, based on the observations of this study, that the ITSM would be normally distributed.

To compare both conditions, we produced a standard usability problem set by the union of usability problem sets obtained from the empirical HEs ("*Union of usability problem sets over UEMs*[10] *being compared*" (HARTSON; ANDRE; WILLIGES, 2001, p. 384)). To this end, we

[9] We have the permission of *Clube Reborn & Toys* to use their content for research purposes.
[10] UEMs: Usability Evaluation Methods

Table 19 – Feasible quasi-experimental design (between subjects).

Condition	Participants (p#)	Diagnosed Problems (N)
Nielsen	p1	5
Nielsen	p2	12
Nielsen	p3	7
Nielsen	p4	7
Nielsen	p5	7
Nielsen	p6	6
Jaferian *et al.*	p7	11
Jaferian *et al.*	p8	13
Jaferian *et al.*	p9	8
Jaferian *et al.*	p10	8
Jaferian *et al.*	p11	16
Jaferian *et al.*	p12	13
Jaferian *et al.*	p13	9
Jaferian *et al.*	p14	8

Source: Research data.

applied a Kruskal-Wallis rank sum test to compare both distributions according to the number of diagnosed usability problems. The results show that the ITSM heuristics had a significantly higher coverage of problems in the diagnosis of usability problems (Kruskal-Wallis chi-squared $= 6.1381$, df $= 1$, p-value ≈ 0.01), which is in accordance with the findings of our mapping study. Nevertheless, future studies may compare these heuristic sets using outcomes from test with users as a benchmark set. Test with users may also reveal usability problems that were not diagnosed by the inspectors, and indicate if there is a lack of coverage of problems among the heuristics. In case a lack of coverage is verified, new heuristics might be necessary to cover them.

To calculate the f-measure of each condition, we used the formulas as indicated by Hartson, Andre and Williges (2001, p. 390-394):

$$Thoroughness = \frac{number\ of\ diagnosed\ problems}{size\ of\ the\ benchmark\ set} \tag{5.1}$$

$$F - measure = \frac{1}{\alpha(\frac{1}{Validity}) + (1-\alpha)(\frac{1}{Thoroughness})} \tag{5.2}$$

In sequence, we aggregated the usability problems diagnosed by both groups (37 problems) to become the benchmark set. Because we adopted the union of both conditions as the benchmark set, the validity measure for both conditions was equal to one (1), making the f-measure equal to the thoroughness. The results also show that the Jaferian *et al.*'s condition influenced a higher f-measure (≈ 0.78) compared to Nielsen's condition (≈ 0.65). Therefore, these results also confirms the findings of our map, evidencing that the ITSM heuristics are the most appropriate for lay privacy policy interfaces in the domain of smart toys.

After the HEs, we asked the inspectors about their opinion on the method. All inspectors affirmed that they had difficulties to assign the heuristic to problems during the inspection, which is similar to the feedback from inspectors in the study of Jaferian *et al.* (2014), when inspectors let to assign the heuristics after identifying the problems. For this reason, and because we limited the inspection time on one hour, none of the participants assigned heuristics to their diagnosed problems.

Although inspectors that used the ITSM heuristics diagnosed a higher number of usability problems, there were eight (8) usability problems identified only by inspectors that used Nielsen's heuristics. These problems refer to general usability aspects, which are not specific related to smart toys' privacy. Meanwhile, among the 13 problems that were only diagnosed by those inspectors using the ITSM heuristics, there are both general usability problems and privacy related problems. Moreover, inspectors that employed the ITSM heuristics found all the eight (8) usability problems related to smart toys' privacy, while those that employed Nielsen's heuristics found only three of them. The full list of usability problems, with the indication of which group diagnosed them, is presented at de Lima Salgado *et al.* (2020).

Table 20 presents the usability problems related to smart toys' privacy. For each of them, we analyzed which of the the ITSM heuristics is most appropriate to address the respective problem. Interestingly, the problem "*The application does not follow conventions for user authentication*" was diagnosed by inspectors that employed the ITSM heuristics. However, the ITSM heuristics do not cover consistency related problems. This fact emphasizes the need for a consistency heuristic among the ITSM heuristics, which was also observed during their study ("*The need for a consistency heuristic was indicated by PI2, PI14, and PI6.*" (JAFERIAN *et al.*, 2014, p. 341)).

Table 20 – Coverage of ITSM heuristics on observed usability problems that are related to smart toys' privacy

Jaferian *et al.*'s Heuristic	**Usability Problems**
Rules and constraints	- At the "*Obligations and Retention*" screen, it is only possible to select PIPEDA and stated purpose - At the "*Obligations and Retention*" screen, it is only possible to select PIPEDA and stated purpose - At the "*Review and add rule*" screen, there is no option to disagree with the policy - At the "*Child Information*"screen, users can only go ahead if they they agree with the terms
Visibility of activity status	- At the "*Review and add rule*", the privacy rule description is still complex to understand - At the "*Review Privacy Policy*", at the Clube Reborn policies, the unique contact information is a WhatsApp number (external app).

Planning and dividing work between users	- The application does not provide feedback on what user is using the application (father/mother/guardian)
None	- The application does not follow conventions for user authentication

Source: Elaborated by the author.

Finally, as indicated at Table 20, only three out of the seven ITSM heuristics were necessary to refer to usability problems related to smart toys' privacy, while one of these problems could not be referred by their heuristics. This may indicate the need for creating new heuristics to cover all usability problems related to smart toys' privacy (e.g. privacy consistency and standards). Future studies may explore this gap.

5.8 Conclusions

This study aimed to answer the question: *What usability heuristics best address usability problems that affect laypeople interaction with privacy policy interfaces related to smart toys?* To answer this question, we performed a snowballing mapping study, evaluating 589 publications (despite potential duplicates) among three snowballing iterations (337+191+61), which resulted in 13 mapped studies (S1-S13).

Table 21 summarizes the results of our map and their adequacy with the evaluated themes (T#). We indicate studies that are adequate to the particular theme with a "x". Among the mapped studies, no one proposed heuristics for our domain (T1) of interest (privacy policy interfaces for smart toys). Yet, two studies (S4 and S5) suggested heuristics for broad domains that include lay privacy policy interfaces for generic security devices. Nine studies[11] describe the creation of their heuristics (T2) with some validation process (T3). Five of them[12] empirically compared their heuristics against Nielsen's, which may better support its employment on HEs (T4). Finally, two studies (S5 and S11) reported traditional metrics that indicate the effectiveness of their proposed heuristics (T5).

As shown in Table 21, the heuristics indicated in **S5** stand as the most appropriate to answer our question. To confirm this finding, we conducted a case study comparing the heuristics proposed in S5 (ITSM heuristics) against the traditional heuristics of Nielsen. The results of the case study confirmed that the ITSM heuristics have a greater and significant impact on the diagnosis of usability problems in lay privacy controls for smart toys (Kruskal-Wallis, p-value ≈ 0.01). For this reason, we conclude that, to the extent to which this study covered, the ITSM heuristics (JAFERIAN *et al.*, 2014) best-addressed problems that affect laypeople's interaction with parental privacy controls for smart toys, and can be pointed as the answer to our research

[11] S3, S5, S7, S8, S9, S10, S11, S12, and S13
[12] S5, S8, S11, S12, and S13

Table 21 – Summary of adequacies between studies and themes (T#).

Study	**T1**	**T2**	**T3**	**T4**	**T5**
S1				x	
S2					
S3		x			
S4	x			x	
S5	**x**	**x**	**x**	**x**	**x**
S6					
S7		x		x	
S8		x	x		
S9		x			
S10		x		x	
S11		x	x	x	x
S12		x	x		
S13		x	x	x	

Source: Elaborated by the author.

question. Nevertheless, future studies may explore the creation of new heuristics for lay privacy policy controls for smart toys, and compare against Nielsen's and the ITSM heuristics.

Based on the findings of this study, we suggest as future studies to compare Nielsen's and the ITSM heuristic sets using outcomes from tests with users as a source of a benchmark set. We also recommend exploring the creation of new domain-specific heuristics for parental privacy controls for smart toys, which may include aspects of consistency that are not described among the ITSM heuristics.

CHAPTER

6

SPECIFYING AND EVALUATING THE NEW HEURISTICS

This chapter is composed of three consecutive studies. Its goal was to create and evaluate usability heuristics for usability inspections of privacy policy interfaces for laypeople. Our method is based on the activities established in the literature to create usability principles (QUIñONES; RUSU, 2017):

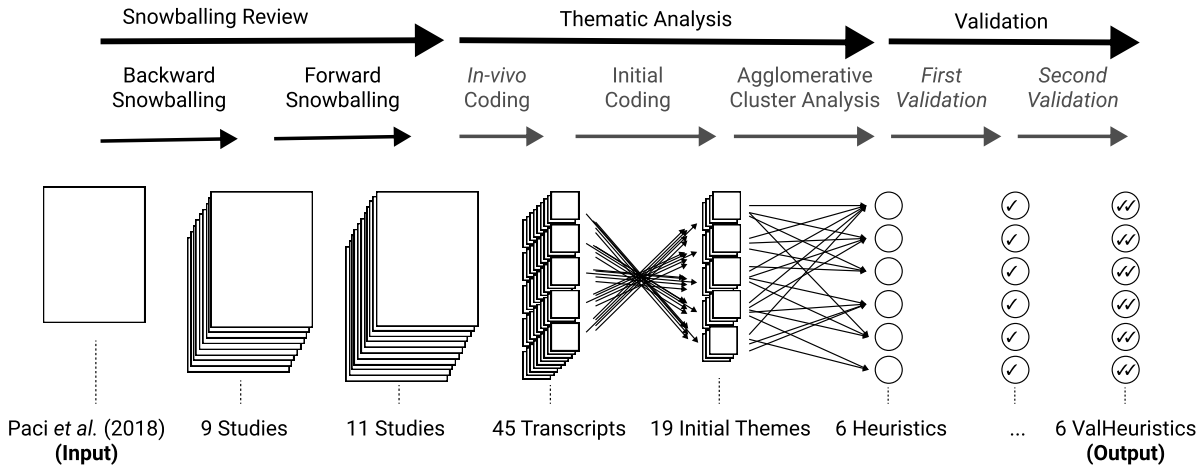
- To identify the state-of-the-art on usability heuristics for privacy policy interfaces designed for laypeople.
- To determine specific features of privacy policy interfaces designed for laypeople.
- To specify the new set of heuristics.
- To validate the new set of heuristics.

To identify the state-of-the-art usability heuristics for privacy policy interfaces designed for laypeople, we systematically mapped the literature as presented in the previous Chapter 5. This chapter determines specific features of privacy policy interfaces designed for laypeople, specifies the new heuristics, and evaluates them. To this end, we performed a qualitative secondary analysis (SILVERMAN, 2016) by reusing data from the literature to retrieve these issues. We reviewed studies that report usability findings from formative evaluations of privacy policy interfaces. In the end, we create six usable privacy heuristics by clustering the usability findings and evaluating the outcomes.

Figure 14 presents a summary of this chapter by indicating the steps of the process. We start by mapping the literature on usability heuristics for privacy policy interfaces through a snowballing review. The snowballing review is separated into two processes, backward and

forward snowballing. In total, we collected 11 studies. After, we conducted a thematic analysis of these 11 studies. We searched for empirical usability test results indicating positive or negative usability attributes for each. We began by transcribing in-vivo codes, usability findings described in the texts. Forty-five transcriptions were collected. After, we performed an initial coding process to find emerging themes among the transcripts. We found 19 initial themes at this stage, later suggested as an initial set of usable privacy guidelines. In sequence, we clustered the 19 initial themes into **six new privacy and usability heuristics** called push#. Finally, we evaluated the push# heuristics through a first and a second quasi-experiment study comparing the new heuristics against the state-of-the-art usability heuristics for privacy controls.

Figure 14 – Qualitative Secondary Analysis Process.



Source: Elaborated by the author.

The remaining chapter describes each of the steps of the qualitative secondary analysis: the snowballing review, the thematic analysis, and the evaluation. Along with them, we present and discuss preliminary and final results.

6.1 Snowballing review

As indicated at Figure 14, we composed the benchmark set of usability issues by collecting usability findings from the literature. First, we performed a snowballing review of the literature to collect studies reporting usability findings from a formative evaluation of privacy policy interfaces. At this step, our start set was the nine studies referred by Paci, Squicciarini and Zannone (2018) in their survey related to policy comprehension and configuration. These studies relate to the following interfaces:

(a.) Improptu (RODE *et al.*, 2006).
(b.) Expandable Grid for Windows XP (REEDER *et al.*, 2008).
(c.) Eyes Metaphor (SCHLEGEL; KAPADIA; LEE, 2011).

(d.) Reflective Policy Assessment (RPA) (ANWAR; FONG, 2012).
(e.) PViz (MAZZIA; LEFEVRE; ADAR, 2012).
(f.) VeilMe (WANG *et al.*, 2015).
(g.) Retinue (HU; AHN; JORGENSEN, 2011).
(h.) MController (HU; AHN; JORGENSEN, 2013).
(i.) Sigma (HU; AHN; JORGENSEN, 2012).

From the start set, we performed the snowballing process (WOHLIN, 2014) to collect more studies from the start set [1]. For the snowballing, we only performed the forward process. We did not perform the backward snowballing to find new studies because Paci, Squicciarini and Zannone (2018) already surveyed the previous studies. For this reason, we skipped the backward discovery and performed the analysis on those surveyed studies. Overall, from the whole snowballing process, we retrieved two additional studies reporting formative usability findings, which represent the following interfaces:

(a.) Nutrition Label (KELLEY *et al.*, 2009).
(b.) Expandable Grid for P3P (REEDER, 2008).

In total, we collected 11 studies that diagnosed usability findings related to privacy policy interfaces. Therefore, the next step was to collect transcripts containing the usability findings. We retrieved 45 transcripts from the full texts, which formed the benchmark set of usability findings. Annex A presents the 45 transcripts. For this chapter, we organized these studies as the quasi-experimental design described at Table 22, which presents the interfaces, the number of participants involved in formative usability evaluation (if any), the number of transcripts in the full text, and the reference for the study.

Table 22 – Quasi-experimental design for the qualitative secondary analysis.

Interface	**Participants (n)**	**Transcripts (n)**	**Reference**
Improptu	24	16	(RODE *et al.*, 2006)
Expandable Grid (XP)	36	3	(REEDER *et al.*, 2008)
Eyes Metaphor	41	2	(SCHLEGEL; KAPADIA; LEE, 2011)
Reflective Policy Assessment	36	4	(ANWAR; FONG, 2012)
PViz	20	5	(MAZZIA; LEFEVRE; ADAR, 2012)
VeilMe	124	5	(WANG *et al.*, 2015)
Retinue	0	0	(HU; AHN; JORGENSEN, 2011)
MController	0	0	(HU; AHN; JORGENSEN, 2013)
Sigma	0	0	(HU; AHN; JORGENSEN, 2012)
Expandable Grid (P3P)	12	5	(REEDER, 2008)
Nutrition Label	24	5	(KELLEY *et al.*, 2009)
TOTAL	**317**	**45**	

[1] Searches performed on November 1st, 2018.

6.1.1 Discussing snowballing outcomes

This section presents interfaces for policy comprehension and configuration, features, and the main usability findings transcribed from the snowballing process.

Impromptu

Rode *et al.* (2006) designed the Impromptu prototype. Their goal was to explore and enhance the visualization of system status (shared files) and its configuration. Impromptu aimed to provide users with collective visibility of actions from all other users. In this sense, they designed a visualization panel with concentric pie charts (see an example at <https://bit.ly/2LFkLNh>). Different pie slices represent different users; labeled dots indicated shared files. Different concentric regions represented different privacy permissions (according to the distance from the center of the pie). Files shown closer to the center of the pie were those with more sharing permissions. On the other hand, files shown far from the center were those with fewer permissions.

In regards to visualization of system status, the design of Impromptu received positive feedback from users: "*the rings and blink around file icons indicate what is open, Permits you to see what others are doing, a clear indication of which files belong to who, concentric spheres representing levels of privacy, clear who is logging in, a clear indication of who is looking at what file, clear indication of who is accessing your own files and good visualization of different levels of access*". In regards to configuration, Impromptu received the following positive feedback: "*easy to share files, easy to set permissions, easy to modify files, doesn't require technical knowledge of permissions, private level is intuitive and one can show or hide easily*" (RODE *et al.*, 2006, p. 5-6).

Although Impromptu received positive feedback, Rode *et al.* (2006) observed two main usability problems. First, users confused the visualization of their actions (read, edit or copy) with others' actions. Second, users "*were confused as to whether the ring indicated the current state of the file (ownership) or whether it represented a past edit of the file*".

Expandable Grid and Nutrition Label

Reeder (2008) and Reeder *et al.* (2008) proposed the Expandable Grid interface. Screenshots of the expandable grid versions are available at <https://bit.ly/2Pgl7fA>. The expandable grids aimed to enhance both the visualization and configuration of file permissions of Windows XP. It was a tabular-based interface at first. It represents users with the upper axis, files with the vertical axis, and permissions (read, write, execute, delete and administrate) as colored squares at the intersection between users and files.

In addition to the Expandable Grid for Windows XP, Reeder (2008) proposed a version of the Expandable Grid for visualization of P3P (Platform for Privacy Preferences Project). The

P3P, created by the World Wide Web Consortium (W3C), enabled websites to express their privacy preferences in both human and machine-readable formats. P3P is currently discontinued. Nevertheless, the P3P website remains online. It describes data practice through statements, which were composed by a "*PURPOSE element, a RECIPIENT element, a RETENTION element, a DATA-GROUP element, and optionally a CONSEQUENCE element*" [2]. From such statements, the P3P user agents could compare privacy preferences from service providers against the users' privacy preferences (SUN; HUANG; KE, 2014). The P3P Expandable Grid shows data categories along the vertical axis and the recipient and purpose (privacy practice) at the upper horizontal axis.

Comparing the contributions of both grids, Reeder (2008) found that the Expandable Grid approach to policy-authoring interface design was preferable to offline applications. In addition, Reeder (2008) observed that users faced the following usability problems:

- Users had difficulty to read rotated text.
- Users had difficulties to answer questions for medium and long P3P policies because each policy had multiple P3P statements.
- Users had difficulty to find policy metadata that was out of the Grid.
- Users seemed confused with P3P concepts and terminology.
- Users had difficulty to understand two dimensions in one axis (upper axis).
- Users slip the mouse into a wrong column of the matrix.
- Users became confused with resources with similar names.
- Users seemed to have difficulties to find a place to start looking at in the grid.
- Users seemed confused with the icons adopted.
- Users had difficulties to find relevant information in the P3P data hierarchy.
- Some users did not interpret that the grid was expandable.
- Policies were not very clear at all.

Later, the P3P Expandable Grid was studied by Kelley *et al.* (2009). They improved its usability by employing a nutrition label metaphor familiar to users. Kelley *et al.* (2009) argue that the P3P statements basically consists of multiple information triples of *data, purpose, recipient*. For this reason, the Nutrition Label Expandable Grid reduced the amount of information displayed at the P3P Expandable Grid employing statement categorizations. Longer definitions

[2] Retrieved from P3P website at: <https://www.w3.org/TR/P3P11/#Policies>

were shown on a "useful terms" page. Kelley *et al.* (2009) adopted colors and symbols (!, −, OUT and IN) to simplify the visualization of permissions. A legend provided explanation about the meaning of each symbol.

According to Kelley *et al.* (2009), users realized the benefits of the nutrition label paradigm for comparison of policies, which was a positive usability-related attribute. Although, they observed that users were confused by the symbols and "*completely unfamiliar with the terms*", which represent usability problems with the interface.

The Eyes Metaphor

Schlegel, Kapadia and Lee (2011) proposed the Eye Metaphor interface. It aimed to summarize the visualization of users' exposure and provide effective control on mobile platforms. They adopted the metaphor of a pair of eyes to suggest whether any information is being collected by someone else. The eye's size suggests the information exposed (bigger the eyes, more information exposed). Users could identify other users by their relationship (e.g., family, friend), not by their identity.

According to the authors, the eyes metaphor was more intuitive than a detailed information interface used compared with the control group. They observed that static disclosure policies (detailed information interfaces) are insufficient to mitigate exposure threats. They also observed that specific relationships (e.g., siblings/parents/relatives instead of family) might play an essential role in enhancing users' sense of control.

The Reflective Policy Assessment (RPA)

Anwar and Fong (2012) designed a prototype for Reflective Policy Assessment (RPA), also called Mirror-Looking Metaphor (PACI; SQUICCIARINI; ZANNONE, 2018). Such interface paradigm is also referred as mirror-looking metaphor (PACI; SQUICCIARINI; ZANNONE, 2018). To achieve such a paradigm, Anwar and Fong (2012) employed graphs to represent the neighborhoods. RPA implies that users can visualize their information from the viewpoint of another user in their virtual neighborhood, reflecting the policy for information exposure assessments. An undirected graph represents the neighborhood, and nodes represent users. Clicking on nodes allows users to access the information exposure from the respective user (represented by the selected node).

According to Anwar and Fong (2012), distance policy (the graph visualization) helps users to think about strangers' access to their files, which is a positive usability-related attribute. Meanwhile, they found that their tool would be more useful if the nodes were colored, each representing an access scenario, indicating a usability defect.

PViz

Mazzia, LeFevre and Adar (2012) designed the PViz prototype. They aimed to align the interface visualization with users' mental models. According to them, users' mental models of privacy involve groups and subgroups of their friends. Therefore, PViz aimed to represent such groups with a bubble chart. Bubbles with different sizes represented groups with a different number of friends. Each bubble represented specific information (detailed at the side axis) and the group of users who could access it. In addition, the color of the bubbles indicated the level of visibility of the respective information.

According to Mazzia, LeFevre and Adar (2012), in comparison with Facebook's privacy control interfaces, PViz was easier to use to get a general idea of who could see what information. Nonetheless, they observed two usability problems with PViz. First, it was was hard to see what data an entire group had access to. Second, it was hard to see the audience view and settings menu percentages.

VeilMe

Wang *et al.* (2015) proposed the VeilMe interface. The VeilMe interface is based on the Twitter profile design. Also, it has a space to show recent tweets at the side of privacy settings. VeilMe allows users to configure their privacy preferences according to their personality portraits. Their portraits are automatically calculated by VeilMe and shown in a hierarchy of traits, among three columns. Each column represents a group of connections (other users). The authors call this the social distance metaphor.

The social distance metaphor represents the information exposure distance from users' information to their social groups (WANG *et al.*, 2015). By grabbing a knob icon, users can set the distance percentage for the groups: Close Colleagues, Distant Colleagues, and Public. VeilMe calculates and defines which users compose each group according to the distance percentage. Besides setting the distance for each group, users can set each characteristic (trait) level of obfuscation in their portrait. Users can also set the level of obfuscation for groups of traits. In these cases, VeilMe applies the consequences to all traits in the respective group.

According to Wang *et al.* (2015), users said it was easier to set privacy preferences for closer groups. They also found it easier to set privacy preferences for groups of traits instead of setting for each trait. These findings may be considered as positive usability-related attributes. On the other hand, VeilMe's finer-grained traits seemed difficult for users to set their constraints. In addition, users complained that VeilMe did not provide enough transparency and information to assess the quality of their settings.

Retinue

Hu, Ahn and Jorgensen (2011) proposed the Retinue prototype for collaborative control of photo sharing. The Retinue aims to resolve privacy conflicts in collaborative privacy controls. Retinue's main interface is a photo gallery inspired by Facebook. This gallery has four (4) tabs: Owner, Tagged, Contributed, and Disseminated. At these tabs, users can view their photos (Owner tab), photos they were tagged (Tagged tab), photos they contributed (Contributed tab), and photos they disseminated (Disseminated tab). All users can set their privacy preferences over photo sharing.

The Retinue also indicates the trade-off between privacy risks and sharing loss, Hu, Ahn and Jorgensen (2011) adopted speedometer metaphors at the configuration interface. One speedometer to indicate the privacy risks level, and one to indicate the sharing loss levels.

Hu, Ahn and Jorgensen (2011) did not report any formative usability evaluation of the Retinue's interface. Instead, they performed a summative evaluation. According to them, they only measured users' basic opinions of particular features. They compared Facebook's privacy controls against Retinue in regards to users' self-understanding of likeability, understanding, and control. According to them, users preferred Retinue for all the aspects covered. This may indicate a positive usability aspect of Retinue. However, it is impossible to understand which usability-related attributes made Retinue the preferred users.

MController

Hu, Ahn and Jorgensen (2013) created the MController. It aims to enable multiple users to control the privacy of collaboratively created data. This prototype is visually similar to the Retinue (HU; AHN; JORGENSEN, 2011), but has additional features (e.g., to solve conflicts). MController's main interface is a photo gallery inspired by Facebook. This gallery has four (4) tabs presented at the Retinue interface: Owner, Tagged, Contributed, and Disseminated. At these tabs, users can view their photos (Owner tab), photos they were tagged (Tagged tab), photos they contributed (Contributed tab), and photos they disseminated (Disseminated tab). All users can set their privacy preferences over photo sharing. However, owners have access to additional controls in contrast to the Retinue prototype.

At the MController's interface, data owners can resolve sharing conflicts manually. In this case, owners receive an alert when a new conflict occurs. Conflict resolution is automatic by default. In addition, Hu, Ahn and Jorgensen (2013) employed the speedometer metaphor to indicate the sensitivity score of photo permissions. The MController calculates the sensitivity score based on the settings chosen by users. Because Hu, Ahn and Jorgensen (2013) do not report any formative usability evaluation with MController's interface, it is difficult to know which usability problems it resolves.

Sigma

Hu, Ahn and Jorgensen (2012) designed the Sigma prototype. It aims to allow users to control the privacy of their photos on Google+ collaboratively. Its main interface allows users to sort their connections in different circles and set a trust level for individuals or circles. If users set the trust level to circles, all connections in such circles receive the respective trust level. Hu, Ahn and Jorgensen (2012) reported summative usability evaluation of Sigma, but did not report any formative usability evaluation. For this reason, we cannot retrieve solved/unsolved usability problems from such a study.

6.2 Thematic Analysis

To specify our new heuristics, we performed a thematic analysis (CLARKE; BRAUN, 2014; BRAUN; CLARKE, 2006) by retrieving 45 transcripts identified in the literature to associate them among higher-level themes. The transcripts are *in-vivo* codes, which means that they are users' feedback described in the literature. To be included, the *in-vivo* code should contain an empirical indication of user behavior with privacy policy tools. For each transcript, we coded initial themes (called the initial coding process) by retrieving user behaviors indicated in the transcripts. Two researchers did this coding process independently. They performed the process until theoretical saturation when no new themes emerged. Appendix A presents the relation among transcripts, their reference, and initial themes. In total, we coded 19 initial themes, as shown in Table 23.

Table 23 – The 19 initial themes resulted from the initial coding process.

	Theme
1	**Perception of exposure** Users aim to understand the extent to which their data is exposed.
2	**Accessing frequently** Users frequently check their privacy choices.
3	**Mitigating exposure threats** Users are acting to mitigate existing privacy threats.
4	**Assessing personality analytics** Users aim to understand the quality of their privacy choices.
5	**Checking audience view** Users aim to understand who has their data (audience).
6	**Sharing information** Users are performing the tasks required to share their data.

continued on the next page

Table 23 – The 19 initial themes from the initial coding process. (continued from the previous page)

	Theme
7	**Assessing own protection by testing other's protection** Users assess others' privacy choices and the consequences before deciding their own privacy choices.
8	**Deducting how the interface works** Users creating the mental model of the privacy policy interface.
9	**Mapping exposure and information** Users mapping the level of exposure with respective types of personal data.
10	**Being alert** Users being vigilant while protecting their data.
11	**Being careful about sharing** Users taking precautions before sharing their data.
12	**Making slips with tabular interfaces** Users making slips when performing privacy choices in tabular privacy policy interfaces.
13	**Complaining about rotated text** Users complain about the readability of rotated text in privacy policies.
14	**Preferring broad than fine-grained settings** Users prefer to choose privacy policies with fewer settings.
15	**Confusing symbols** Users being confused by symbols that are not familiar to them.
16	**Comparing policies** Users are comparing different privacy policies in the same interface.
17	**Understanding policies** Users reading privacy policies to understand them.
18	**Exploring the interface** Users randomly interact with privacy policy interfaces to understand its conceptual model.
19	**Looking for information** Users searching for specific information.

Source: Elaborated by the author.

The 19 initial themes represent the main characteristics of usability issues that laypeople face interacting with privacy policy tools, as retrieved from empirical data available in the literature. After defining the initial themes, we revised them to suggest respective usable privacy recommendations for future designs of privacy policy tools. The recommendations are suggested design requirements to be considered in the future development of privacy policy tools. The following recommendations could be seen as a preliminary set of *Privacy and Usability Guidelines (pug's)* , as listed in Table 24.

Table 24 – The Preliminary Privacy and Usability Guidelines (**pug#**)

	Guideline
pug#1	**Perception of Data Exposure** The interface should facilitate users to perceive the extent to which their data is exposed.
pug#2	**Support Frequent Access** The interface should provide efficient ways for users to check their privacy choices frequently.
pug#3	**Support Controlling Exposure Threats** The interface should provide precise controls for users to mitigate privacy threats.
pug#4	**Provide Privacy Choice Analytics** The interface should provide feedback about the quality of users' privacy choices.
pug#5	**Provide Audience View Feedback** The interface should inform users about the audience who access their data.
pug#6	**Efficient Sharing** The interface should provide efficient ways for users to share their data safely.
pug#7	**Provide Examples of Privacy Choices** The interface should allow users to compare their privacy choices with others' before sharing their data.
pug#8	**Good Conceptual Model** The interface should be easy to understand its conceptual model at first glance. This recommendation is based on Norman's design principle: conceptual model.
pug#9	**Mapping Exposure to Data** The interface should provide a precise mapping between the level of exposure and the respective personal data types.
pug#10	**Privacy Alerts** The interface should notify users about any privacy occurrence related to their data.
pug#11	**The Privacy Watch Dog** The interface should continuously help users watch out before sharing their data.
pug#12	**Avoid tabular interfaces** Interfaces with multi-columns by multi-rows tables are more likely to make users slip.
pug#13	**Avoid Rotated Text** Avoid displaying rotated text when describing privacy policies. Rotated text confuses the reading of policies.
pug#14	**Provide Flexible Settings** Provide flexible privacy settings, allowing users to choose between general or fine-grained settings to perform their privacy choices.

continued on the next page

Table 24 – The Preliminary Privacy and Usability Guidelines (**pug#**). (continued from the previous page)

	Guideline
pug#15	**Appropriate Use of Symbols** Symbols may help users to understand privacy policies rapidly. However, symbols must be familiar to users.
pug#16	**Comparing Policies** The interface should provide an efficient way for users to differentiate the privacy policies.
pug#17	**Understanding Policies** The interface should provide readable and easy-to-understand privacy policies. Cut the clutter.
pug#18	**Learnable Interface** The interface should allow users to explore its features without damaging their data protection.
pug#19	**Provide Search Engine** The interface should provide a way for users to search for specific information.

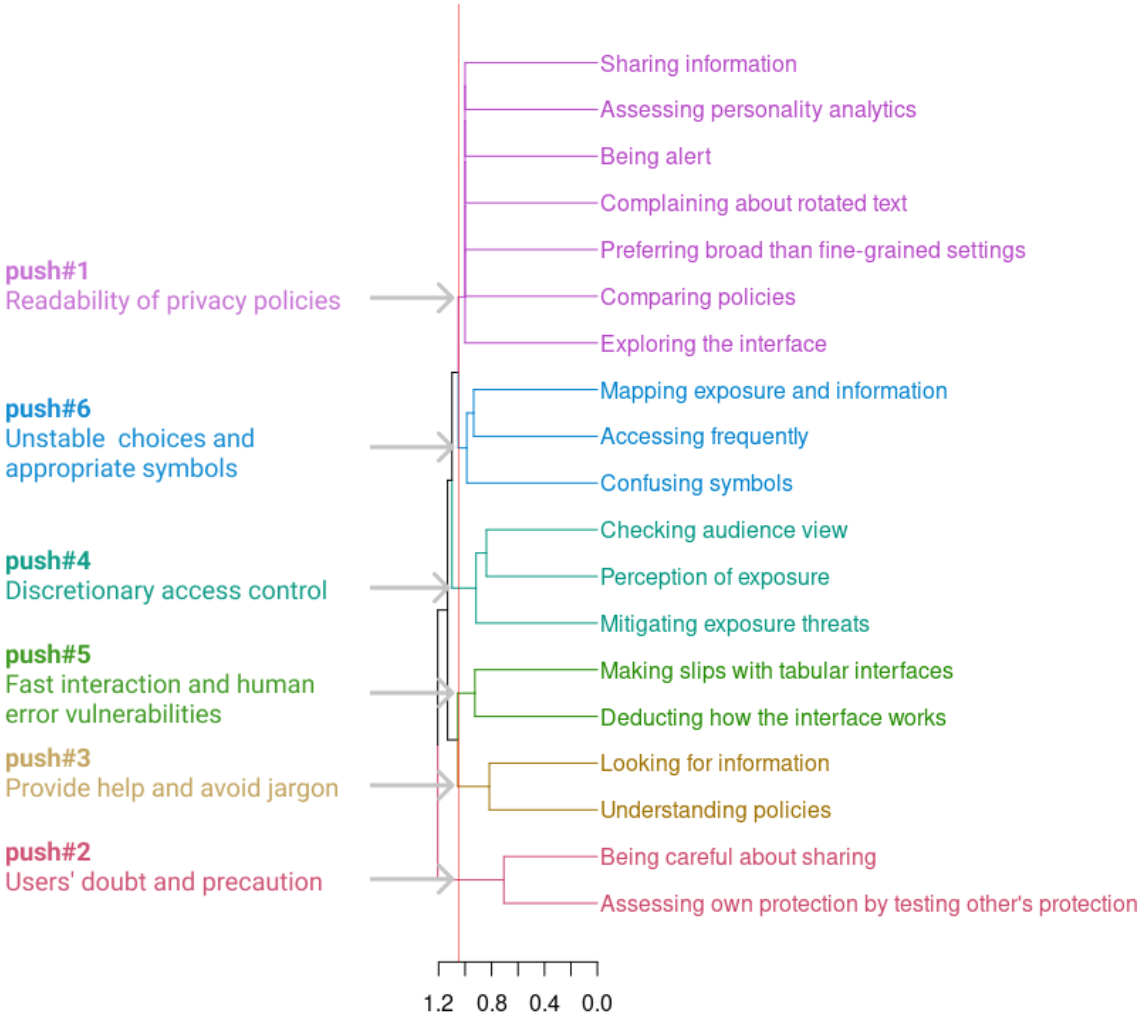
Source: Elaborated by the author.

6.2.1 Discussing thematic analysis outcomes

After we achieved the saturation of initial themes, we revised the 45 transcripts to register occurrences of the 19 initial themes among them. We registered the initial themes and their occurrences among transcripts (the *in-vivo* codes) in a dataset (see Appendix A). After that, we performed a cluster analysis (HAIR *et al.*, 2010) with the dataset to identify higher-level themes that could indicate how to compose the new heuristics from the transcripts.

The cluster analysis led us to identify distinct usability findings with similar attributes. We used the 19 pug# as objects and their occurrences as attributes to perform the cluster analysis. Therefore, each object was a vector of binary data indicating the presence (1) or absence (0) of themes among user observations (transcripts). We clustered the objects using Ward's method and Jaccard distance. We applied Ward's method because it minimizes the within-group (within-cluster) dispersion (MURTAGH; LEGENDRE, 2014). Although the Euclidean distance is indicated for Ward's method (MURTAGH; LEGENDRE, 2014), we employed Jaccard distance not to weight negative matches (0,0) (CHOI; CHA; TAPPERT, 2010). Finally, we sought a maximum of ten clusters to later correspond to a maximum of ten heuristics and not be confused with usability guidelines. As displayed at Figure 15, we got six as the greatest number of clusters under ten from the cluster analysis results. From these six cluster, we composed our usable privacy heuristics. We called them as *Privacy and Usability Heuristics (push)*, as described in the following section. According to their explanatory power among data, we numbered the push# heuristics from 1 (highest) to 6 (lowest).

Figure 15 – Outcome from Cluster Analysis indicating the six heuristics.



Source: Elaborated by the author.

By creating our six heuristics, we were able to test this thesis's hypothesis seeking to answer the second question. Our privacy and usability heuristics for laypeople's heuristic evaluations of privacy policy interfaces follow. The Privacy and Usability Heuristics (push#) are described in Table 25.

Table 25 – The Six Privacy and Usability Heuristics (**push#**)

	Heuristic
push#1	**Readability of privacy policies.** The readability of privacy policies is crucial for users to understand how they share their data. Users may want to access personalized privacy analysis to understand the risks of sharing their data. While users set their privacy choices, they become vigilant and start to explore the interface and assess policies. They may also prefer interfaces with fewer privacy choices instead of complex settings.
push#2	**Users' doubt and precaution.** Users may assess the consequences of others' privacy choices before deciding about sharing their own.
push#3	**Provide help and avoid jargon.** Users may search for specific information, such as terms and definitions, or seek for help to understand privacy policies. Avoid jargon.
push#4	**Discretionary access control.** Users may want to know the extent in which their personal data is being shared. They may also want to know who access their personal data. After that, they may want to restrict the access to their data.
push#5	**Fast interaction and human error vulnerabilities.** Users may seek for fast interactions with privacy policies. To this end, they deduce how the interface works. In these cases, they need to quickly understand how the privacy choice settings work (conceptual model) and human error is very likely to occur.
push#6	**Unstable choices and appropriate symbols.** Users may change their privacy choices over time. A good policy-choice mapping is desirable in these situations. Employing appropriate symbols enhance the mapping.

Source: Elaborated by the author.

The following section presents evaluation studies comparing the push# heuristics against the state-of-the-art empirically.

6.3 Evaluation

To evaluate the push# heuristics, we conducted quasi-experiment studies with between-group designs and compared the new heuristics against the state-of-the-art usability heuristics for privacy controls (JAFERIAN *et al.*, 2014; de Lima Salgado *et al.*, 2020). Therefore, we invited start-up professionals to voluntarily perform separate inspections[3]. Participants were randomly

[3] The Research Ethics Committee, CAAE code 69353317.4.0000.5390, approves the experimental design of this project.

assigned to the control group, which used Jaferian et al.'s heuristics (JAFERIAN *et al.*, 2014), or to the treatment group, which used the push# heuristics.

We followed the recommendations of Caine (2016) to define the sample size for each group of usability evaluators. As shown by Caine, we should seek for a statistical power with Cohen's $d = 0.5$, $\alpha = 0.05$, and $\beta = 0.85$. From these requirements, and assuming parametric distributions for a two-sample t-test power calculation with an alternative hypothesis greater than the null hypothesis, the minimum required sample in each group would be four participants ($n \approx 4$).

We chose Rafferty, Fantinato and Hung (2015) parental privacy control model to be the subject of evaluation. We prototyped the model to be accessed online by a browser simulating a mobile device. It is the same prototype employed in our mapping study that confirmed the Jaferian et al.'s heuristics (JAFERIAN *et al.*, 2014) as the state-of-the-art on usability heuristics for privacy controls (de Lima Salgado *et al.*, 2020).

Finally, we compared the performance of both sets of heuristics by measuring their downstream utility by the number of usable privacy problems reported by participants from each group. **Downstream utility** is a performance measure indicated by Hartson, Andre and Williges (2001). It refers to the performance of a usability evaluation method on generating outputs that add value in the design change process. **Downstream utility** is particularly important for privacy interfaces because we want to find out issues in the interface that are not under privacy legislation and must be redesigned. Table 29 presents the list of usability problems and their respective severity ratings. To measure the downstream utility among these problems, we separated them among the severity scale for usable privacy problems proposed in Yankson, Salgado and Fortes (2021). The referred scale helps us identify usability problems that lead users to agree with the wrong policy (catastrophe) or lead users to generate wrong privacy policies (major). Their scale also offers two lower levels, referring to usability problems unrelated to privacy policy generation/agreement. However, we cannot ensure that redesigns would occur for those levels. Also, we aimed to evaluate the new heuristics by their possible advantage in finding privacy-related problems. Otherwise, we would not be evaluating the new heuristics to the domain they are supposed to be applicable.

6.3.1 Discussing first evaluation outcomes

We invited 100 human-computer interaction undergraduate students to participate in our experiment for our first study. All of them were novice evaluators and had just seen how to conduct a heuristic evaluation for the first time. The participation was voluntary, and 29 out of the 100 invited students accepted to take part in our experiment. Therefore, the feasible sample for this experiment was 27 participants, structured as shown in Table 26.

Table 26 – Experiment design for the first study.

	Control Group	**Treatment Group**
Heuristic set	ITSM Heuristics	Six usable privacy heuristics
Participants (n)	14	13

Source: Research data.

We planned our experiment to be remote and online, as required by the social distancing scenario when this experiment was conducted (2020). We asked each participant to answer a questionnaire with 40 questions checking "*true or false*" about the existence of usability problems in the parental privacy control model. The questionnaires were randomly generated from a dataset of 40 actual usability problems and 40 false usability problems. We generated the dataset from a sum of 40 true usability problems, reported in (de Lima Salgado *et al.*, 2020) and complemented by usability problems indicated by experts. After, we generated the 40 false usability problems by denying the existence of the true problems. To goal of this first study was not to evaluate the new heuristics but to overview their potential prior to conducting the actual experiment.

After the 27 participants answered the questionnaire, we calculated the number of correct answers for each group. Table 27 shows these numbers along with the number of correct answers by a participant.

Table 27 – Results from the pilot study.

Participant	**Group 0 - Control / 1 - Treatment**	**Number of correct answers**
1	0	21
2	0	23
3	0	20
4	0	21
5	0	26
6	0	27
7	0	22
8	0	20
9	0	33
10	0	21
11	0	26
12	0	28
13	0	28
14	0	30
15	1	35

continued on the next page

Table 27 – Results from the pilot study. (continued from the previous page)

Participant	Group 0 - Control / 1 - Treatment	Number of correct answers
16	1	21
17	1	23
18	1	25
19	1	22
20	1	30
21	1	28
22	1	32
23	1	26
24	1	33
25	1	25
26	1	27
27	1	29

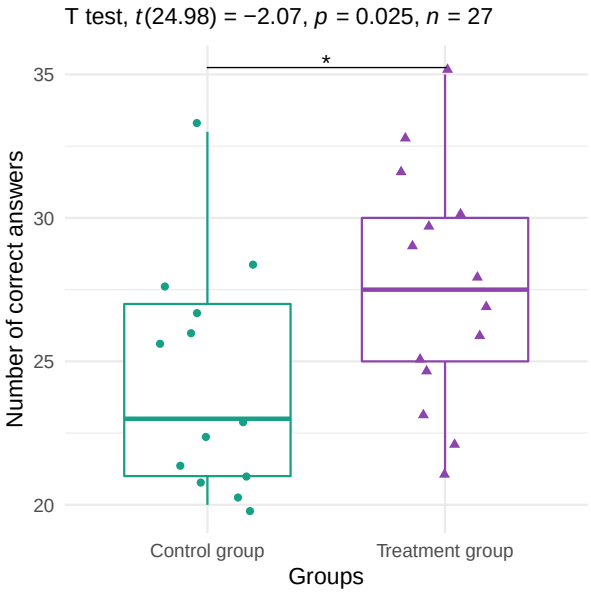
Source: Research data.

After Shapiro-Wilk normality tests, neither the control ($p-value = 0.15$) nor the treatment ($p-value = 0.92$) distribution were significantly different from normal distribution. For this reason, we assumed them as normal distributions. Therefore, we compared both groups using a parametric comparison, the Student's t-test, with an alternative hypothesis as true location shift is less than 0. The test results showed a significant difference between groups ($p-value = 0.025$). This result might be coherent with our alternative hypothesis. However, the first experiment is sufficient to evaluate push# heuristics. Participants were not performing a usability inspection. For this reason, we planned a follow-up experiment to evaluate the new heuristics from usability inspection' data empirically.

6.3.2 Discussing second evaluation outcomes

We invited 60 professionals from a cluster of software development start-ups located in Lavras, Minas Gerais, Brazil, for our evaluation study. The participation was voluntary, and 32 out of the 60 invited professionals accepted to take part in our study. Therefore, the feasible sample for our between-group experimental design was of 32 participants, structured as shown in Table 28. The control group employed the state-of-the-art, and the treatment group employed the six new push# heuristics. We randomly assigned the participants to the groups. The experience of the control group on user interface design projects ranged from less than a year to six whole years of experience ($M = 0.5, SD \approx 2.10$). The experience of the treatment group on user interface design projects ranged from less than a year to four whole years of experience ($M = 0.0, SD \approx 1.54$). The difference of experience between groups is not significant.

Figure 16 – Student's t test results for the first experiment.



Source: Elaborated by the author.

Table 28 – Experiment design for the evaluation study.

	Control Group	**Treatment Group**
Heuristic set	ITSM Heuristics	6 push#
Participants (n)	16	16

Source: Research data.

This evaluation study was also performed remotely due to social distancing constraints. At this time, we asked each participant to inspect the parental privacy control model by employing one of the heuristic sets according to their group. All participants had 40 minutes to perform the inspection. Afterward, the participants reported the diagnosed usability problems in a spreadsheet indicating the problem's description and the affected heuristic. In total, both groups diagnosed 31 usability problems in the interface.

After evaluating the 31 problems, we noticed that nine could lead users to agree with the wrong policy (catastrophe). Meanwhile, 14 could lead users to generate wrong privacy policies (major). The results include four minor problems, which do not relate to policy generation/agreement but may stop users from using the interface. They also include four cosmetic problems, which do not relate to policy generation/agreement and may not stop users from using the interface. The results are shown in Table 29[4].

[4] The full dataset with the respective severity ratings can be found in Appendix B

Table 29 – Descriptions of diagnosed usability problems.

ID	Usability problem description	Severity
P1	There is no verification of consent.	Catastrophe
P2	The interface does not provide users with privacy rules' log.	Minor
P3	In the profile screen, users cannot select "*Não..*" (No) after selecting "*Sim...*" (Yes)	Minor
P4	In the privacy rules screen, when users select only one checkbox (sub option), all of them become checked.	Major
P5	In the privacy rules screen for obligations and retention, only one out of two check-boxes is working properly.	Major
P6	In the review and add rule screen, the privacy policies inform that users can deny the policies, but there is no option for denying.	Catastrophe
P7	The design does not ensure that users read all of the privacy policies before agreeing with it.	Catastrophe
P8	In the purposes screen, the purposes' descriptions are not sufficiently clear for users when users are consenting with the options.	Catastrophe
P9	In the recipients screen, the information about who will have access to the data is insufficient for a transparent rule.	Major
P10	The privacy policies are ambiguous and unclear.	Catastrophe
P11	Users cannot search for specific information.	Minor
P12	The design does is not sufficiently clear on how the collect data will be shared.	Catastrophe
P13	The screen for parent/guardian details has confusing information.	Minor
P14	The privacy policies screen does not allow users to continue the interaction without agreeing with the policies.	Catastrophe
P15	The design does not allow users to have a fast and simple interaction towards their goals.	Major
P16	The design employs jargon.	Major
P17	The design does not employ symbols to help users in understanding the policies.	Major
P18	The design should be in Portuguese, but one of the screens is in English.	Major
P19	The design's color palette (blue and bold black fonts) is not appropriate for readability.	Cosmetic
P20	The design does not allow users to select "anyone" as recipients.	Major
P21	The design does not offer help for users.	Major
P22	In the core access control screen, the design asks users to select among "operations" (plural), while there is only one option of operation.	Cosmetic
P23	The design does not have a pattern for icons, buttons and writing.	Cosmetic
P24	The design does not inform users about potential consequences of privacy choices.	Catastrophe
P25	The design does not offer users an explanation about their rights according to the respective regulation/Law.	Catastrophe
P26	In the core access control screen, the design provides insufficient explanation about the checkbox options.	Major

continued on the next page

Table 29 – Descriptions of diagnosed usability problems. (continued from the previous page)

ID	Usability problem description	Severity
P27	In the core access control screen, users cannot proceed the creation of a rule after checking the relative location.	Major
P28	In the purposes screen, the purposes' descriptions are not sufficiently clear for users when users are choosing the options.	Major
P29	In the purposes screen, users cannot proceed the interaction depending on the combination of chosen options.	Major
P30	In the obligations and retention screen, the search box is not working.	Cosmetic
P31	The design does not allow users to cancel the creation of a privacy rule.	Major

Source: Elaborated by the author.

Figure 17 shows the number of times that participants referred to each heuristic. As one can see, the general reference to heuristics from the state-of-the-art was lower than the reference to the six new heuristics. Also, none of the control group participants referred to heuristic #6 from the state-of-the-art. This difference suggests that the 6 push# heuristics are easier to apply in privacy policy interfaces for laypeople. Nonetheless, these numbers are not sufficient to evaluate the new heuristics. The following sections present a deeper comparison among the group's performance by assessing the number of diagnosed problems according to their severity.

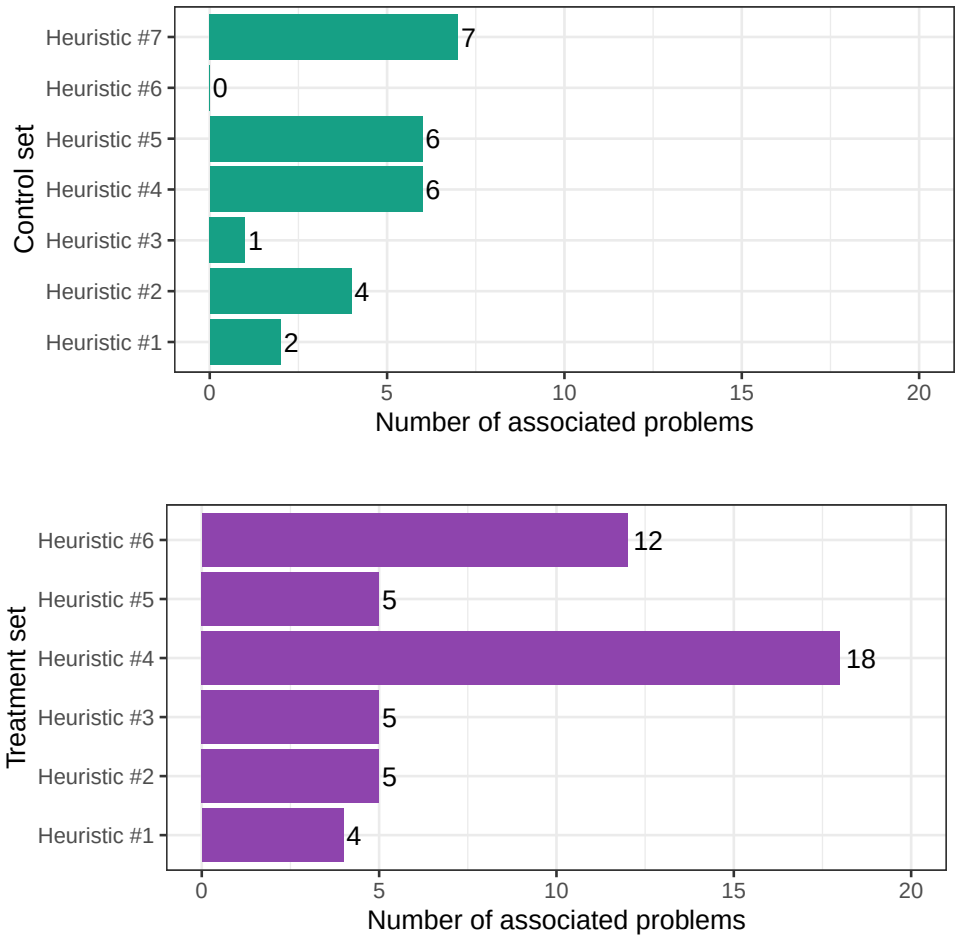
Performance measures

To assess the performance of each set of heuristics, we measured the downstream utility of the usability problems reported in the experiment. We measured the downstream utility by counting the number of catastrophic and major usability problems reported by each group in the experiment. We also compared the overall number of usability problems reported and the number of minor and cosmetic problems reported by each group.

Downstream utility on catastrophic problems

We compared the downstream utility on the number of catastrophic problems diagnosed by participants in each group. In the control group, participants discovered a minimum of zero and maximum of two catastrophic usability problems ($\bar{X} = 0.375, M = 0.00, S \approx 0.72$). Meanwhile, in the treatment group, participants discovered a minimum of zero and maximum of three catastrophic usability problems ($\bar{X} = 0.9375, M = 1.00, S \approx 1.00$). After a normality test, both the results' distribution from the control group ($p - value = 9.986e-06$) and from the treatment group ($p - value \approx 0.0065$) significantly differed from a normal distribution. Therefore, as indicated by Lazar, Feng and Hochheiser (2017), we employed the Wilcoxon signed-rank test to statistically compare the downstream utility on the number of catastrophic problems

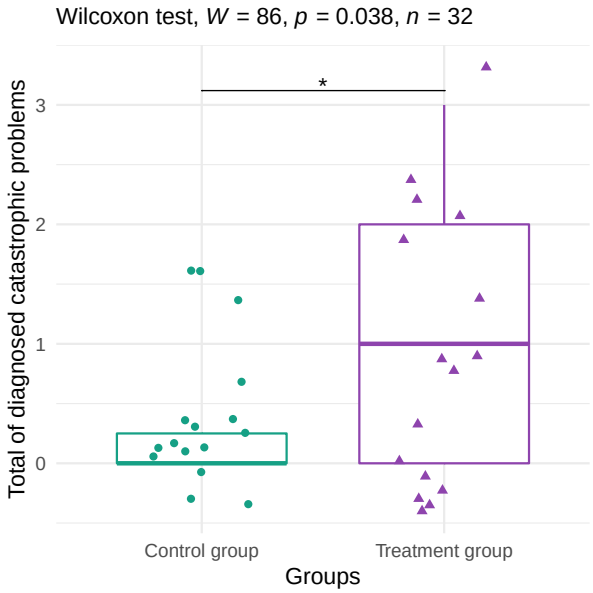
Figure 17 – Distribution of diagnosed usability problems by heuristics.



Source: Elaborated by the author.

between groups. The results are shown in Figure 18 with jitter, boxplots, and the respective statistical results. For the downstream utility on catastrophic problems, evaluators' performance using the push# heuristics was significantly greater than the performance of evaluators using the state-of-the-art ($p - value = 0.038$). In addition, the data analysis showed a moderate effect size (0.317).

Figure 18 – Results of the Wilcoxon signed-rank test.



Source: Elaborated by the author.

For instance, the results of push# heuristics' performance on catastrophic problems evidence its benefits over the state-of-the-art. For this reason, we accept the thesis hypothesis for the scope of catastrophic problems.

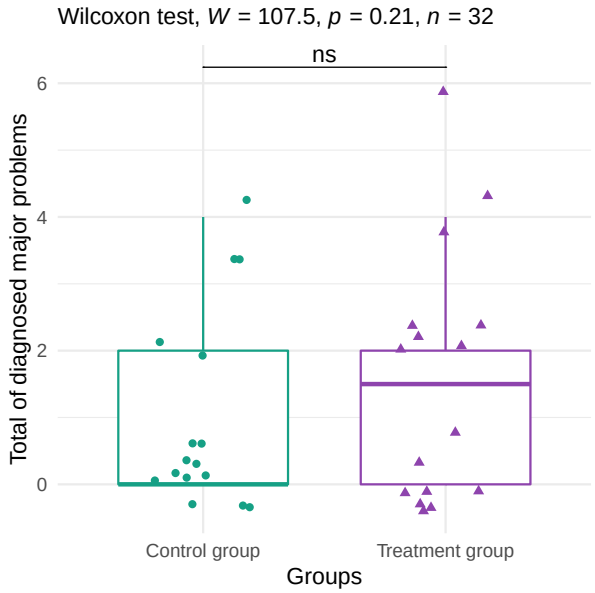
Downstream utility on major problems

Comparing the downstream utility on the number of major problems reported, participants in the control group discovered a minimum of zero and a maximum of four major problems ($\bar{X} = 1.00, M = 0.00, S \approx 1.37$). Meanwhile, in the treatment group, participants discovered a minimum of zero and maximum of six major problems ($\bar{X} = 1.562, M = 1.50, S \approx 1.82$). After a normality test, both results' distribution from the control ($p - value \approx 0.0008$) and treatment ($p - value \approx 0.004$) group significantly differed from a normal distribution. Therefore, we compared the number of diagnosed problems between groups by applying the Wilcoxon signed-rank test. Although the observed results from the treatment group suggest a slight improvement compared with the control group, both groups had similar performance in this measure. Figure 20 illustrates both distributions with jitter and boxplots while indicating the respective statistics.

Overall number of usability problems reported

In addition to comparing the downstream utility of the heuristics, we compared the number of usability problems diagnosed by participants in each group. In the control group, participants discovered a minimum of zero and maximum of five usability problems ($\bar{X} = 1.75, M = 1.00, S \approx 1.95$). Meanwhile, in the treatment group, participants discovered a minimum

Figure 19 – Results of the Wilcoxon signed-rank test.



Source: Elaborated by the author.

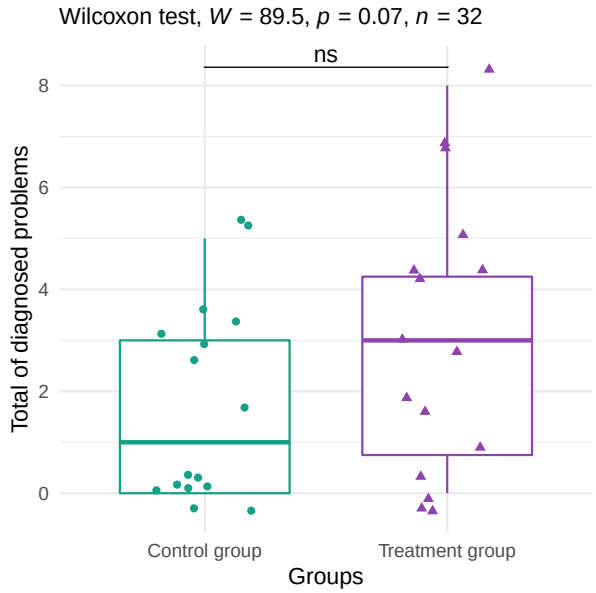
of zero and maximum of eight usability problems ($\bar{X} = 3.125, M = 3.00, S \approx 2.66$). After a normality test, the results' distribution from the control group significantly differed from a normal distribution ($p - value \approx 0.002$), while the results' distribution from the treatment group did not ($p - value \approx 0.14$). Therefore, we compared the number of diagnosed problems between groups by applying the Wilcoxon signed-rank test. Although the results from the treatment group suggest an overall improvement compared with the control group, the difference was not significant. Figure 20 illustrates both distributions with jitter and boxplots while indicating the respective statistics.

Number of minor and cosmetic problems reported

Comparing the number of minor problems reported, participants in the control group discovered a minimum of zero and maximum of one minor problem ($\bar{X} = 0.375, M = 0.00, S = 0.5$). Meanwhile, in the treatment group, participants discovered a minimum of zero and maximum of two minor problems ($\bar{X} = 0.375, M = 0.00, S \approx 0.62$). After a normality test, both results' distribution from the control ($p - value = 2.566e-05$) and treatment ($p - value = 4.803e-05$) group significantly differed from a normal distribution. Therefore, we compared the number of diagnosed problems between groups by applying the Wilcoxon signed-rank test. Although the treatment group's observed results suggest a slight improvement compared with the control group, both groups had similar performance in this measure. Figure 21 illustrates both distributions with jitter and boxplots while indicating the respective statistics.

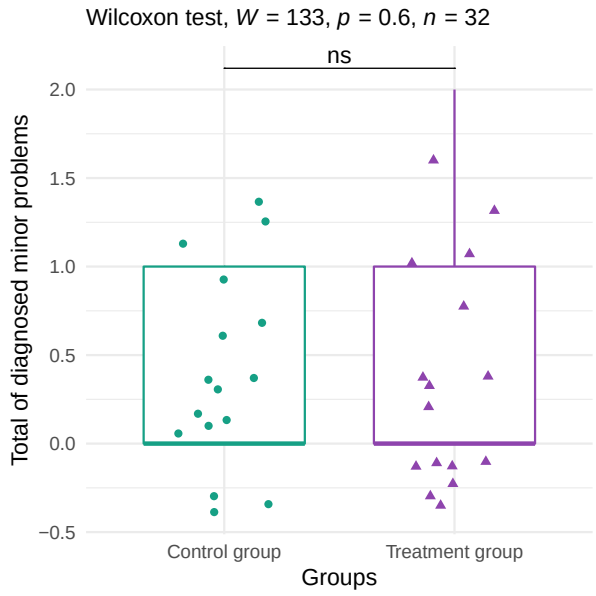
Similar to the number of minor problems, we observed a slight difference in the number

Figure 20 – Results of the Wilcoxon signed-rank test.



Source: Elaborated by the author.

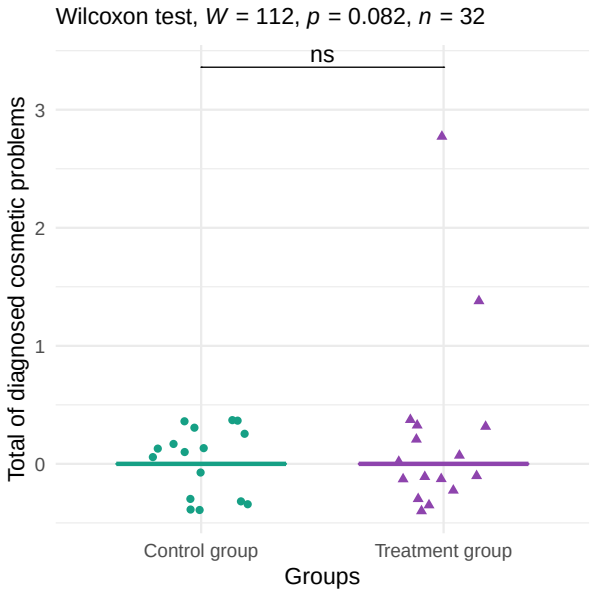
Figure 21 – Results of the Wilcoxon signed-rank test.



Source: Elaborated by the author.

of cosmetic problems reported between groups. Participants in the control group did not find any cosmetic problems. Meanwhile, in the treatment group, participants discovered a minimum of zero and maximum of three cosmetic problems ($\bar{X} = 0.25, M = 0.00, S \approx 0.77$). After a normality test, the results' distribution from the treatment group did not ($p - value = 2.444e - 07$). Figure 22 illustrates both distributions with jitter and boxplots while indicating the results from a Wilcoxon signed-rank test.

Figure 22 – Results of the Wilcoxon signed-rank test.



Source: Elaborated by the author.

The data analysis on the discovery of cosmetic problems suggests that the push# heuristics have higher coverage over the ITSM. However, the observed results are not significant to confirm it.

6.4 Final Remarks

In this chapter, we have specified and evaluated the push# heuristics. We retrieved empirical data from usability studies during a qualitative secondary analysis and, later, clustered such data into themes to compose the new heuristics. After that, we conducted quasi-experiment studies to evaluate the new push# heuristics. We compared the push# heuristics against the state-of-the-art usability heuristics for privacy controls; the ITSM heuristics (JAFERIAN *et al.*, 2014; de Lima Salgado *et al.*, 2020). The analysis of downstream utility showed that the push# heuristics have a more significant performance on the discovery of catastrophic usable privacy problems compared to the ITSM heuristics. These findings led us to accept the hypothesis that *employing new usability heuristics, focused on the domain of privacy policy interfaces for laypeople, enhances the performance of heuristic evaluation in the domain*. Although not significant, the results suggest that the push# heuristics have higher performance on major, minor and cosmetic severity levels as well. The performance of push# heuristics was also higher on the overall number of usability problems diagnosed for the observations of this study.

In addition to creating the push# heuristics, we also created the pug# guidelines. However, it was out of the scope of this thesis to evaluate the pug# guidelines. Therefore, we suggest

future studies to evaluate them.The following chapter concludes this thesis, summarizes its contributions and discusses the limitations and future works.

CHAPTER

7

CONCLUSION

Usability has served the Human-Computer Interaction (HCI) field for more than 20 years, denoting a quality attribute of user interfaces (HORNBÆK, 2018; TRACTINSKY, 2018). Although the theoretical knowledge of the usability definition has challenged the field and motivated fundamental discussions (TRACTINSKY, 2018; BORSCI *et al.*, 2019), a primary practice in HCI pushes the field (BERTELSEN, 2018). In this sense, HCI "must be practical and relevant to people, organizations, or design" (LAZAR; FENG; HOCHHEISER, 2017, p. 5). This thesis sought to create practical usability heuristics to enhance the design and evaluation of privacy policy interfaces. Usable privacy is instead a legislative requirement than only a software quality. Enhancing the transparency of privacy policy interfaces stands as a challenge that computer science must address to enhance users' trust in technology (SLEPCHUK; MILNE, 2020). Instead of providing long and complex privacy policies, we need to design more usable interfaces that empower laypeople to protect their privacy online. In this thesis, we aimed at creating usability criteria (heuristics) for inspecting such interfaces. Consequently, we expect future designs of privacy policy interfaces to provide better interfaces after solving the usability issues diagnosed with our heuristics. To this end, we sought to answer the two consecutive research questions:

– What are the main characteristics of usability issues that laypeople face interacting with privacy policy tools?

– How effective is the performance of usability heuristics composed of these characteristics?

To answer the first question, we conducted a qualitative analysis of the literature. We collected 45 transcripts reporting empirical data from usability tests with users on different

interface models of privacy control. With a thematic analysis, we grouped these transcripts into 19 initial themes representing the main characteristics of usability issues faced by such users when interacting with privacy policy tools. We also provide **19 preliminary privacy and usability guidelines (pug#)** regarding each of the 19 themes. With another round of the thematic analysis, we grouped the 19 initial themes into **six privacy and usability heuristics (push#)**.

To answer the second question, we tested our *hypothesis*: "*employing new usability heuristics, focused on the domain of privacy policy interfaces for laypeople, enhances the performance of heuristic evaluation in the domain*". The results show that the push# heuristics significantly enhances heuristic evaluation performance in the domain. When applied to evaluate privacy policy interfaces for laypeople, the push# heuristics enhances the downstream utility on the number of catastrophic problems discovered. Although we did not find a significant enhancement on the push# regarding major, minor and cosmetic problems, the results also suggest higher performance on these severity levels. We conclude that push# heuristics are valid to evaluate privacy policy interfaces used by laypeople to protect their data privacy.

In addition to answering the research questions, the conduction of this doctoral research also resulted in **a new process for creating new usability criteria**. We adapted the activities proposed by Quiñones and Rusu (2017) to a practical process with data cluster analysis. Bertelsen (2018) stated that developing methods and approaches based on practical usability is one of the lanes we should follow towards the future of the HCI field. Usability as an umbrella concept relates to the need of multiple communities to modify and adapt the concept to their product-specific standards (BORSCI *et al.*, 2019), as for privacy policy interfaces and their related legislative acts. Meanwhile, the technology landscape continuously grows. As new communities arise and technologies advance towards being ubiquitous, the need for adapting practical usability methods for the new needs arises together. We believe it would be beneficial to create domain-specific usability criteria according to their needs. Nevertheless, relying on time-consuming processes to create such criteria might not be the most appropriate for companies such as start-ups (SALGADO *et al.*, 2019b). Our process allows companies to repeat the cluster analysis whenever they retrieve additional data and code it instead of repeating the whole process for all data. Figure 14 summarizes the process. This characteristic makes our process lean and incremental for the needs of companies.

7.1 Additional Contributions

Besides creating six usable privacy heuristics for inspecting the domain of privacy policy interfaces to be used by laypeople, the execution of this thesis' project also raised the following contributions:

A systematic mapping of usability heuristics for privacy policy interfaces designed for laypeople: we present a systematic mapping study of the literature to identify state of the art on usability heuristics for privacy policy interfaces aiming laypeople. See Chapter 5 and de Lima Salgado *et al.* (2020).

Enhancing usability through information architecture of privacy policies: in this thesis, we describe an experiment using cluster analysis mixed with the card sorting technique to evaluate the information architecture of privacy policies and propose a redesign of a privacy control model. See Chapter 3 and Salgado *et al.* (2019a).

Models of heuristic evaluation for novice evaluators: this thesis presents a discussion on different models of heuristic evaluation that aim to improve the overall performance of novice evaluators using collaborative inspection, see Salgado *et al.* (2018).

Usability heuristics for mobile games and elderly players: we employed qualitative analysis to compose usability heuristics for mobile games and elderly players. This study served as a preliminary use of qualitative analysis to compose heuristics from literature content, as shown in Salgado *et al.* (2019b).

A preliminary ontology for usability findings: we proposed a preliminary ontology to describe privacy-related usability findings. Our goal was to help researchers identify the similarity of usability problems when assessing the performance of different usability evaluation methods. See Salgado *et al.* (2019).

Data Glove controlled interface to enhance privacy in health systems: we co-authored an exergame to treat musculoskeletal disorders. The exergame uses gesture interfaces and the assistance of a social robot. The design was helpful to explore the use of gesture interaction as private interactions, which is less public than voice interactions with social robots. See Demoe *et al.* (2020).

Recommendations for the design of usable privacy controls for smart toys: we described a security analysis and usability evaluation of smart toys. As a result, it describes recommendations for further design of privacy policy interfaces. These recommendations are additional to those presented concerning the 19 initial themes and pug# guidelines. See Chapter 4 and Yankson, Salgado and Fortes (2021).

Privacy expectations on user experience regarding connectedautonomous vehicles: we investigated the privacy expectations of potential users of connected-autonomous vehicles. Our goal was to understand the user experience in such a context while getting insights into designing future privacy controls for the domain. See Salgado *et al.* (2020).

7.2 Limitations and Future Work

This thesis created six validated usability and privacy heuristics. We tested its performance on heuristic evaluation of a parental privacy control model for smart toys (RAFFERTY; FANTINATO; HUNG, 2015) to validate them. Although the model aims to protect the data privacy of users' children and not users' privacy, the overall mechanics of creating privacy rules and choosing policies remains the same. Nevertheless, we only tested our heuristics for a mobile interface design. This fact may limit the range of possible human interactions. Future studies can investigate the validity of push# heuristics for other domains, such as human-robot interactions and human-artificial intelligence interactions. These technologies commonly enable human interactions, such as gesture and voice interaction, that are less employed in mobile devices like phones and tablets. Also, this thesis validated the push# heuristics for laypeople as the user profile. Laypeople is an important but still generic, to some extent, user profile. It was necessary to investigate the validity of the push# heuristics for laypeople. Future works remain to investigate whether adaptations can be made to the heuristics to improve its performance for specific user profiles. We suggest that it be primarily for the population with the widest range of user needs, characteristics, capabilities, and eventual assistive technologies.

This thesis provided 19 preliminary design guidelines, the pug# guidelines. We composed them as a preliminary step towards composing usable privacy design guidelines. Nevertheless, future studies must validate them. Also, we suggest that future studies evaluate the order of the guidelines to organize them from most to less strict guidelines. Future studies can also explore grouping the guidelines under the six usable privacy heuristics.

Beyond enhancing the usability of privacy-protection tools, we suggest that future studies explore how usable designs could nudge users to the safest interactions. The unmotivated user property states that managing the security process is not the users' primary goal. As usability depends on users' goals (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018b), understanding the usability of security tools for unmotivated users, such as privacy policy tools, is always problematic. However, the unmotivated user property states that users "*want security in place to protect them while they do those things*" (WHITTEN; TYGAR, 1999). This fact makes security a so-called secondary goal. People "*assume that their security is working, while they focus on their primary goals*" (WHITTEN; TYGAR, 1999).

For this reason, from the unmotivated users' perspective, the security process should take place while they achieve their primary goals. This discussion belongs to nudging research. Nudging research refers to "*interventions (...) aimed at helping users make better online security and privacy decisions - that is, decisions that minimize adverse outcomes or are less likely to be regretted*" (ACQUISTI *et al.*, 2017). Therefore, we suggest future studies to explore what we call **nudging usability**. Nudging usability is a topic of nudging research that concerns summative

or formative usability practices, aiming to nudge unmotivated users toward safer interactions. Nudging usability is about usable security - making "*easy for users to do the right thing*" (THEOFANOS, 2020) - through usability by "*minimizing the risk and the undesirable consequences of use errors*" (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018b). Future studies should create design methods for nudging usability.

BIBLIOGRAPHY

ACQUISTI, A.; ADJERID, I.; BALEBAKO, R.; BRANDIMARTE, L.; CRANOR, L. F.; KOMANDURI, S.; LEON, P. G.; SADEH, N.; SCHAUB, F.; SLEEPER, M.; AL. et. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. **ACM Comput. Surv.**, Association for Computing Machinery, New York, NY, USA, v. 50, n. 3, Aug. 2017. ISSN 0360-0300. Available: <https://doi.org/10.1145/3054926>. Citations on pages 65 and 140.

ACQUISTI, A.; BRANDIMARTE, L.; LOEWENSTEIN, G. Privacy and human behavior in the age of information. **Science**, v. 347, n. 6221, p. 509–514, Jan. 2015. ISSN 0036-8075, 1095-9203. Available: <http://science.sciencemag.org/content/347/6221/509>. Citations on pages 25 and 43.

ACQUISTI, A.; TAYLOR, C. R.; WAGMAN, L. **The Economics of Privacy**. Rochester, NY, 2016. Available: <https://papers.ssrn.com/abstract=2580411>. Citations on pages 44, 55, and 86.

ALBUQUERQUE, A. P. de; KELNER, J. Toy user interfaces: Systematic and industrial mapping. **Journal of Systems Architecture**, Dec. 2018. ISSN 1383-7621. Available: <www.sciencedirect.com/science/article/pii/S138376211830153X>. Citation on page 91.

____. Non-personal Data Collection for Toy User Interfaces. In: **Proceedings of the 52nd Hawaii International Conference on System Sciences**. [s.n.], 2019. Available: <www.scholarspace.manoa.hawaii.edu/handle/10125/59612>. Citation on page 91.

ALBUQUERQUE, d. O. P.; FANTINATO, M.; KELNER, J.; ALBUQUERQUE, A. P. d. Privacy in smart toys: Risks and proposed solutions. **Electronic Commerce Research and Applications**, p. 100922, 2019. ISSN 1567-4223. Available: <www.sciencedirect.com/science/article/pii/S1567422319300997>. Citation on page 91.

ALBUQUERQUE, O. de P.; FANTINATO, M.; KELNER, J.; de Albuquerque, A. P. Privacy in smart toys: Risks and proposed solutions. **Electronic Commerce Research and Applications**, v. 39, p. 100922, 2020. ISSN 1567-4223. Available: <https://www.sciencedirect.com/science/article/pii/S1567422319300997>. Citation on page 69.

ALJOHANI, M.; BLUSTEIN, J.; HAWKEY, K. Participatory Design Research to Understand the Legal and Technological Perspectives in Designing Health Information Technology. In: **Proceedings of the 35th ACM International Conference on the Design of Communication**. New York, NY, USA: ACM, 2017. (SIGDOC '17), p. 39:1–39:3. ISBN 978-1-4503-5160-7. Available: <http://doi.acm.org/10.1145/3121113.3121240>. Citation on page 52.

ALONSO-RÍOS, D.; MOSQUEIRA-REY, E.; MORET-BONILLO, V. A Systematic and Generalizable Approach to the Heuristic Evaluation of User Interfaces. **International Journal of Human–Computer Interaction**, v. 0, n. 0, p. 1–14, Jan. 2018. ISSN 1044-7318. Available: <https://doi.org/10.1080/10447318.2018.1424101>. Citation on page 87.

ANWAR, M.; FONG, P. W. L. A Visualization Tool for Evaluating Access Control Policies in Facebook-style Social Network Systems. In: **Proceedings of the 27th Annual ACM Symposium on Applied Computing**. New York, NY, USA: ACM, 2012. (SAC '12), p. 1443–1450.

ISBN 978-1-4503-0857-1. Available: <http://doi.acm.org/10.1145/2245276.2232007>. Citations on pages 47, 113, and 116.

APTHORPE, N.; VARGHESE, S.; FEAMSTER, N. Evaluating the Contextual Integrity of Privacy Regulation: Parents' IoT Toy Privacy Norms versus COPPA. USENIX Association, USA, p. 123–140, 2019. Available: <dl.acm.org/doi/10.5555/3361338.3361348>. Citation on page 92.

Assistant Secretary for Public Affairs. **Heuristic Evaluations and Expert Reviews**. 2013. Available: </how-to-and-tools/methods/heuristic-evaluation.html>. Citation on page 75.

AïMEUR, E.; LAWANI, O.; DALKIR, K. When changing the look of privacy policies affects user trust: An experimental study. **Computers in Human Behavior**, v. 58, p. 368–379, May 2016. ISSN 07475632. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0747563215302296>. Citations on pages 25 and 45.

BERTELSEN, O. W. Commentary: Usability and the Primacy of Practice. **Human–Computer Interaction**, Taylor & Francis, v. 33, n. 2, p. 182–185, 2018. Available: <https://doi.org/10.1080/07370024.2017.1321991>. Citations on pages 137 and 138.

BERTINO, E. Data Security and Privacy: Concepts, Approaches, and Research Directions. In: **2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)**. [S.l.: s.n.], 2016. v. 1, p. 400–407. Citations on pages 26, 43, 44, 45, 52, 53, 55, 56, 74, 86, 87, 88, and 93.

BERTOLINI, A.; AIELLO, G. Robot companions: A legal and ethical analysis. **The Information Society**, v. 34, n. 3, p. 130–140, May 2018. ISSN 0197-2243. Available: <www.doi.org/10.1080/01972243.2018.1444249>. Citation on page 91.

BORSCI, S.; FEDERICI, S.; MALIZIA, A.; FILIPPIS, M. L. D. Shaking the usability tree: why usability is not a dead end, and a constructive way forward. **Behaviour & Information Technology**, Taylor & Francis, v. 38, n. 5, p. 519–532, 2019. Available: <https://doi.org/10.1080/0144929X.2018.1541255>. Citations on pages 137 and 138.

BORSCI, S.; MACREDIE, R. D.; BARNETT, J.; MARTIN, J.; KULJIS, J.; YOUNG, T. Reviewing and Extending the Five-User Assumption: A Grounded Procedure for Interaction Evaluation. **ACM Trans. Comput.-Hum. Interact.**, v. 20, n. 5, p. 29:1–29:23, Nov. 2013. ISSN 1073-0516. Available: <http://doi.acm.org/10.1145/2506210>. Citation on page 42.

BRAUN, V.; CLARKE, V. Using thematic analysis in psychology. **Qualitative Research in Psychology**, v. 3, n. 2, p. 77–101, Jan. 2006. ISSN 1478-0887. Available: <https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa>. Citation on page 119.

CAINE, K. Local Standards for Sample Size at CHI. In: **Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2016. (CHI '16), p. 981–992. ISBN 978-1-4503-3362-7. Available: <http://doi.acm.org/10.1145/2858036.2858498>. Citations on pages 42, 57, 106, and 125.

CHAMIKARA, M.; BERTOK, P.; KHALIL, I.; LIU, D.; CAMTEPE, S. PPaaS: Privacy Preservation as a Service. **Computer Communications**, v. 173, p. 192–205, May 2021. ISSN 01403664. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0140366421001420>. Citations on pages 25 and 44.

Children's Commissioner. **Who knows what about me? A Children's Commissioner report into the collection and sharing of children's data**. [S.l.], 2018. Available: <www.childrenscommissioner.gov.uk/our-work/digital/who-knows-what-about-me/>. Citations on pages 86 and 91.

CHO, H.; KNIJNENBURG, B.; KOBSA, A.; LI, Y. Collective Privacy Management in Social Media: A Cross-Cultural Validation. **ACM Transactions on Computer-Human Interaction**, v. 25, n. 3, p. 1–33, Jun. 2018. ISSN 10730516. Available: <http://dl.acm.org/citation.cfm?doid=3231919.3193120>. Citation on page 55.

CHOI, S.-S.; CHA, S.-H.; TAPPERT, C. C. A Survey of Binary Similarity and Distance Measures. v. 8, n. 1, p. 6, 2010. Citation on page 122.

Chu, G.; Apthorpe, N.; Feamster, N. Security and Privacy Analyses of Internet of Things Children's Toys. **IEEE Internet of Things Journal**, v. 6, n. 1, p. 978–985, Feb 2019. ISSN 2372-2541. Available: <www.ieeexplore.ieee.org/document/8443103>. Citations on pages 91 and 92.

CLARKE, V.; BRAUN, V. Thematic Analysis. In: TEO, T. (Ed.). **Encyclopedia of Critical Psychology**. New York, NY: Springer New York, 2014. p. 1947–1952. ISBN 978-1-4614-5583-7. Available: <https://doi.org/10.1007/978-1-4614-5583-7_311>. Citation on page 119.

COGNITOYS. **Connected Smart Toys from Cognitoys | Order Yours Today!** 2019. Available: <https://cognitoys.com/>. Citation on page 68.

____. **Privacy | Connected Smart Toys from Cognitoys**. 2019. Available: <https://cognitoys.com/pages/privacy>. Citation on page 71.

CONTI, M.; DEHGHANTANHA, A.; FRANKE, K.; WATSON, S. Internet of Things security and forensics: Challenges and opportunities. **Future Generation Computer Systems**, v. 78, p. 544–546, 2018. ISSN 0167-739X. Available: <https://www.sciencedirect.com/science/article/pii/S0167739X17316667>. Citation on page 68.

CRANOR, L. F.; BUCHLER, N. Better Together: Usability and Security Go Hand in Hand. **IEEE Security Privacy**, v. 12, n. 6, p. 89–93, Nov. 2014. ISSN 1540-7993. Available: <www.ieeexplore.ieee.org/document/7006405>. Citations on pages 45 and 88.

CUNHA, M.; MENDES, R.; VILELA, J. P. A survey of privacy-preserving mechanisms for heterogeneous data types. **Computer Science Review**, v. 41, p. 100403, Aug. 2021. ISSN 15740137. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1574013721000435>. Citations on pages 25 and 44.

CUNHA, M. V. de A.; Unicef; others. **Child Privacy in the Age of Web 2.0 and 3.0: Challenges and opportunities for policy**. [S.l.], 2017. (Innocenti Discussion Paper). Available: <www.unicef-irc.org/publications/926-child-privacy-in-the-age-of-web-20-and-30-challenges-and-opportunities-for-policy.html>. Citations on pages 47 and 92.

DE, L. A.; ZEZSCHWITZ, E. v. Usable privacy and security. **it - Information Technology**, v. 58, n. 5, p. 215–216, 2016. ISSN 1611-2776. Available: <https://www.degruyter.com/view/j/itit.2016.58.issue-5/itit-2016-0034/itit-2016-0034.xml>. Citations on pages 26, 45, 52, 53, 55, 56, 75, 86, 87, 88, and 93.

de Lima Salgado, A.; de Mattos Fortes, R. P.; de Oliveira, R. R.; FREIRE, A. P. Usability heuristics on parental privacy controls for smart toys: From an exploratory map to a confirmatory research. **Electronic Commerce Research and Applications**, v. 42, p. 100984, 2020. ISSN 1567-4223. Available: <http://www.sciencedirect.com/science/article/pii/S1567422320300612>. Citations on pages 26, 30, 77, 108, 124, 125, 126, 135, and 139.

DEMERS, R. A. System Design for Usability. **Commun. ACM**, Association for Computing Machinery, New York, NY, USA, v. 24, n. 8, p. 494–501, Aug. 1981. ISSN 0001-0782. Available: <https://doi.org/10.1145/358722.358730>. Citation on page 33.

DEMOE, M.; URIBE-QUEVEDO, A.; SALGADO, A. L.; MIMURA, H.; KANEV, K.; HUNG, P. C. Exploring Data Glove and Robotics Hand Exergaming: Lessons Learned. In: **2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)**. [S.l.: s.n.], 2020. p. 1–8. Citations on pages 30 and 139.

DÍAZ, G. B.; RÍO, C. M. d. P. Z. D. A Proposal of Usability Heuristics Oriented to E-Banking Websites. In: MARCUS, A.; WANG, W. (Ed.). **Design, User Experience, and Usability: Theory and Practice**. Cham: Springer International Publishing, 2018. v. 10918, p. 327–345. ISBN 978-3-319-91796-2 978-3-319-91797-9. Available: <http://link.springer.com/10.1007/978-3-319-91797-9_23>. Citations on pages 97, 102, and 105.

ELKIN, P. L.; BEUSCART-ZEPHIR, M.-C.; PELAYO, S.; PATEL, V.; NØHR, C. The usability-error ontology. In: **CSHI**. [S.l.: s.n.], 2013. p. 91–96. Citation on page 39.

ERICSSON, K. A.; SIMON, H. A. Verbal reports as data. **Psychological Review**, American Psychological Association, US, v. 87, n. 3, p. 215–251, 1980. Citation on page 39.

Federal Trade Commission ("FTC" or "Commission"). **Children's Online Privacy Protection Rule**. [S.l.]: Federal Register, 2013. The Office of the Federal Register (OFR) of the National Archives and Records Administration (NARA), and the U.S. Government Publishing Office (GPO) at the <FederalRegister.gov> website. Citations on pages 68, 88, and 89.

FIERRO, N.; ZAPATA, C. Usability Heuristics for Web Banking. In: MARCUS, A. (Ed.). **Design, User Experience, and Usability: Design Thinking and Methods**. [S.l.]: Springer International Publishing, 2016. p. 412–423. ISBN 978-3-319-40409-7. Citations on pages 96, 102, and 105.

GARFINKEL, S.; LIPFORD, H. R. **Usable Security: History, Themes, and Challenges**. [S.l.]: Morgan & Claypool Publishers, 2014. (SYNTHESIS LECTURES ON INFORMATION SECURITY, PRIVACY, AND TRUST, v. 5). ISBN 978-1-62705-529-1. Citations on pages 25, 45, 46, 52, 53, 55, 56, 86, 87, 88, 92, 93, 94, 95, 100, and 101.

GHOSH, A. K.; BADILLO-URQUIOLA, K.; GUHA, S.; JR, J. J. L.; WISNIEWSKI, P. J. Safety vs. Surveillance: What Children Have to Say About Mobile Apps for Parental Control. In: **Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2018. (CHI '18), p. 124:1–124:14. ISBN 978-1-4503-5620-6. Available: <http://doi.acm.org/10.1145/3173574.3173698>. Citation on page 68.

GOMES, F. T.; SALGADO, A. d. L.; DUARTE, L. M. C.; SANTOS, F. d. S.; FORTES, R. P. Um simulador visual de leitor de telas para auxílio à interpretação de questões de acessibilidade por avaliadores videntes. **Revista de Sistemas e Computação-RSC**, v. 8, n. 1, 2018. Citation on page 29.

GUMUSSOY, C. A. Usability guideline for banking software design. **Computers in Human Behavior**, v. 62, p. 277–285, Sep. 2016. ISSN 07475632. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0747563216302667>. Citations on pages 96, 101, 103, and 104.

HABIB, H.; PEARMAN, S.; WANG, J.; ZOU, Y.; ACQUISTI, A.; CRANOR, L. F.; SADEH, N.; SCHAUB, F. "It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices. In: **Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems**. Honolulu HI USA: ACM, 2020. p. 1–12. ISBN 978-1-4503-6708-0. Available: <https://dl.acm.org/doi/10.1145/3313831.3376511>. Citations on pages 25, 26, and 45.

HAIR, J. F.; ANDERSON, R. E.; BABIN, B. J.; BLACK, W. C. **Multivariate data analysis: A global perspective**. [S.l.]: Pearson Upper Saddle River, NJ, 2010. Citation on page 122.

HARKOUS, H.; FAWAZ, K.; LEBRET, R.; SCHAUB, F.; SHIN, K. G.; ABERER, K. Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning. In: **Proceedings of the 27th USENIX Conference on Security Symposium**. USA: USENIX Association, 2018. (SEC'18), p. 531–548. ISBN 9781931971461. Available: <dl.acm.org/doi/10.5555/3277203.3277243>. Citation on page 92.

HARTSON, H. R.; ANDRE, T. S.; WILLIGES, R. C. Criteria for evaluating usability evaluation methods. **International journal of human-computer interaction**, v. 13, n. 4, p. 373–410, 2001. Citations on pages 29, 105, 106, 107, and 125.

HARTSON, R.; PYLA, P. S. **The UX Book: Process and guidelines for ensuring a quality user experience**. [S.l.]: Elsevier, 2012. Citation on page 39.

Hello Barbie. **Hello Barbie**. 2019. Available: <http://hellobarbiefaq.mattel.com/>. Citation on page 68.

HERMAWATI, S.; LAWSON, G. Establishing usability heuristics for heuristics evaluation in a specific domain: Is there a consensus? **Applied Ergonomics**, v. 56, p. 34 – 51, 2016. ISSN 0003-6870. Available: <http://www.sciencedirect.com/science/article/pii/S0003687015301162>. Citations on pages 26, 41, 87, 90, 99, and 101.

HERTZUM, M. Commentary: Usability—A Sensitizing Concept. **Human–Computer Interaction**, v. 33, n. 2, p. 178–181, 2018. Available: <https://doi.org/10.1080/07370024.2017.1302800>. Citations on pages 33 and 36.

HOLLOWAY, D.; GREEN, L. The Internet of toys. **Communication Research and Practice**, Routledge, v. 2, n. 4, p. 506–519, 2016. Available: <https://doi.org/10.1080/22041451.2016.1266124>. Citations on pages 70 and 71.

HORNBÆK, K. Dogmas in the assessment of usability evaluation methods. **Behaviour & Information Technology**, Taylor & Francis, v. 29, n. 1, p. 97–111, 2010. Available: <https://doi.org/10.1080/01449290801939400>. Citation on page 75.

____. Commentary: Usability and Theory Building. **Human–Computer Interaction**, Taylor & Francis, v. 33, n. 2, p. 186–189, 2018. Available: <https://doi.org/10.1080/07370024.2017.1321992>. Citation on page 137.

HU, H.; AHN, G.-J.; JORGENSEN, J. Detecting and Resolving Privacy Conflicts for Collaborative Data Sharing in Online Social Networks. In: **Proceedings of the 27th Annual Computer Security Applications Conference**. New York, NY, USA: ACM, 2011. (ACSAC '11), p.

103–112. ISBN 978-1-4503-0672-0. Available: <http://doi.acm.org/10.1145/2076732.2076747>. Citations on pages 47, 113, and 118.

HU, H.; AHN, G. J.; JORGENSEN, J. Enabling Collaborative data sharing in Google+. In: **2012 IEEE Global Communications Conference (GLOBECOM)**. [S.l.: s.n.], 2012. p. 720–725. Citations on pages 47, 113, and 119.

____. Multiparty Access Control for Online Social Networks: Model and Mechanisms. **IEEE Transactions on Knowledge and Data Engineering**, v. 25, n. 7, p. 1614–1627, Jul. 2013. ISSN 1041-4347. Citations on pages 47, 113, and 118.

HUNG, P. C.; FANTINATO, M.; RAFFERTY, L. A study of privacy requirements for smart toys. 2016. Citation on page 69.

HUNG, P. C.; IQBAL, F.; HUANG, S.-C. Children's Privacy Protection Engine for Smart Anthropomorphic Toys. **Engineering in Medical Applications**, p. 15, 2016. Citation on page 91.

HUNG, P. C. K.; FANTINATO, M.; ROA, J. Children Privacy Protection. In: LEE, N. (Ed.). **Encyclopedia of Computer Graphics and Games**. Cham: Springer International Publishing, 2018. p. 1–3. ISBN 978-3-319-08234-9. Available: <www.doi.org/10.1007/978-3-319-08234-9_198-1>. Citations on pages 91 and 92.

HUNG, P. C. K.; TANG, J. K. T.; KANEV, K. Introduction. In: TANG, J. K.; HUNG, P. C. K. (Ed.). **Computing in Smart Toys**. Cham: Springer International Publishing, 2017. p. 1–5. ISBN 978-3-319-62072-5. Available: <www.doi.org/10.1007/978-3-319-62072-5_1>. Citations on pages 47, 52, 68, 70, and 91.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems**. 2010. Available: <www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en>. Citations on pages 35, 88, 89, and 90.

____. **ISO/IEC 25066:2016(en), Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for Usability — Evaluation Report**. [S.l.], 2016. Available: <https://www.iso.org/obp/ui/#iso:std:iso-iec:25066:ed-1:v1:en>. Citations on pages 35, 37, 38, 40, 54, 55, 75, and 90.

____. **ISO/IEC 27000:2016(en), Information technology – Security techniques – Information security management systems - Overview and vocabulary**. 2016. Available: <https://www.iso.org/obp/ui/#iso:std:iso-iec:27000:ed-4:v1:en>. Citations on pages 55, 74, and 75.

____. **ISO 8124-1:2018(en), Safety of toys — Part 1: Safety aspects related to mechanical and physical properties**. 2018. Available: <www.iso.org/obp/ui/#iso:std:iso:8124:-1:ed-5:v1:en>. Citation on page 86.

____. **ISO 9241-11:2018(en), Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts**. 2018. Available: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>. Citations on pages 74, 140, and 141.

____. **ISO/TR 18638:2018, Health informatics — Guidance on health information privacy education in healthcare organization**. 2018. Available: <www.iso.org/obp/ui/#iso:std:iso:tr:18638:ed-1:v1:en>. Citation on page 87.

____. **ISO/IEC TR 20547-1:2020(en) Information technology — Big data reference architecture — Part 1: Framework and application process**. [S.l.], 2020. Available: <https://www.iso.org/obp/ui/#iso:std:iso-iec:tr:20547:-1:ed-1:v1:en:term:3.5>. Citation on page 44.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO); INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC). **ISO/IEC 29100:2011(en), Information technology — Security techniques — Privacy framework**. 2011. Available: <www.iso.org/obp/ui/#iso:std:iso-iec:29100:ed-1:v1:en>. Citations on pages 35, 86, and 92.

JAFERIAN, P.; HAWKEY, K.; SOTIRAKOPOULOS, A.; VELEZ-ROJAS, M.; BEZNOSOV, K. Heuristics for Evaluating IT Security Management Tools. **Human–Computer Interaction**, v. 29, n. 4, p. 311–350, Jul. 2014. ISSN 0737-0024. Available: <http://dx.doi.org/10.1080/07370024.2013.819198>. Citations on pages 40, 77, 78, 79, 80, 94, 95, 101, 102, 105, 108, 109, 124, 125, and 135.

Jakob Nielsen. **10 Heuristics for User Interface Design**. 2018. Available: <https://www.nngroup.com/articles/ten-usability-heuristics/>. Citations on pages 40, 41, and 61.

JANG-JACCARD, J.; NEPAL, S. A survey of emerging threats in cybersecurity. **Journal of Computer and System Sciences**, v. 80, n. 5, p. 973–993, Aug. 2014. ISSN 0022-0000. Available: <http://www.sciencedirect.com/science/article/pii/S0022000014000178>. Citations on pages 45, 52, 55, 75, and 88.

JIANG, X.; LANDAY, J. A. Modeling privacy control in context-aware systems. **IEEE Pervasive Computing**, v. 1, n. 3, p. 59–63, Jul. 2002. ISSN 1536-1268. Available: <www.ieeexplore.ieee.org/document/1037723>. Citation on page 92.

JOHNSTON, J.; ELOFF, J.; LABUSCHAGNE, L. Security and human computer interfaces. **Computers & Security**, v. 22, n. 8, p. 675–684, Dec. 2003. ISSN 01674048. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0167404803000063>. Citation on page 101.

JØSANG, A.; ALFAYYADH, B.; GRANDISON, T.; ALZOMAI, M.; MCNAMARA, J. Security Usability Principles for Vulnerability Analysis and Risk Assessment. In: **Twenty-Third Annual Computer Security Applications Conference (ACSAC 2007)**. [S.l.: s.n.], 2007. p. 269–278. Citations on pages 75, 90, 94, 95, and 101.

JØSANG, A.; ZOMAI, M. A.; SURIADI, S. Usability and Privacy in Identity Management Architectures. In: **Proceedings of the Fifth Australasian Symposium on ACSW Frontiers - Volume 68**. AUS: Australian Computer Society, Inc., 2007. (ACSW '07), p. 143–152. ISBN 192068285X. Available: <www.dl.acm.org/doi/10.5555/1274531.1274548>. Citations on pages 94 and 95.

Juniper Research. **Smart Toy Sales to Grow Threefold to Exceed $15.5 Billion by 2022**. [S.l.], 2017. Available: <www.juniperresearch.com/press/press-releases/smart-toy-sales-to-grow-threefold>. Citation on page 86.

KATSABAS, D.; FURNELL, S.; DOWLAND, P. Using human computer interaction principles to promote usable security. In: **Proceedings of the Fifth International Network Conference (INC 2005), Samos, Greece**. [S.l.: s.n.], 2005. p. 235–242. Citations on pages 95, 101, and 103.

KAUSHIK, K.; JAIN, N. K.; SINGH, A. K. Antecedents and outcomes of information privacy concerns: Role of subjective norm and social presence. **Electronic Commerce Research and**

Applications, v. 32, p. 57–68, Nov. 2018. ISSN 15674223. Available: <www.linkinghub.elsevier.com/retrieve/pii/S1567422318300814>. Citation on page 86.

KELLEY, P. G.; BRESEE, J.; CRANOR, L. F.; REEDER, R. W. A "Nutrition Label" for Privacy. In: **Proceedings of the 5th Symposium on Usable Privacy and Security**. New York, NY, USA: ACM, 2009. (SOUPS '09), p. 4:1–4:12. ISBN 978-1-60558-736-3. Available: <http://doi.acm.org/10.1145/1572532.1572538>. Citations on pages 15, 47, 52, 53, 54, 56, 58, 62, 63, 65, 66, 92, 93, 113, 115, and 116.

KELLEY, P. G.; CRANOR, L. F.; SADEH, N. Privacy as part of the app decision-making process. In: . ACM Press, 2013. p. 3393. ISBN 978-1-4503-1899-0. Available: <http://dl.acm.org/citation.cfm?doid=2470654.2466466>. Citations on pages 53 and 54.

KERCKHOFFS, A. La cryptographie militaire. IX, p. 5–38, 1883. Citation on page 101.

KEYMOLEN, E.; HOF, S. V. der. Can I still trust you, my dear doll? A philosophical and legal exploration of smart toys and trust. **Journal of Cyber Policy**, Routledge, v. 4, n. 2, p. 143–159, 2019. Available: <www.doi.org/10.1080/23738871.2019.1586970>. Citation on page 92.

KOKOLAKIS, S. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. **Computers & Security**, v. 64, n. Supplement C, p. 122–134, Jan. 2017. ISSN 0167-4048. Available: <http://www.sciencedirect.com/science/article/pii/S0167404815001017>. Citations on pages 44 and 55.

KRUMM, J. **Ubiquitous computing fundamentals**. [S.l.]: CRC Press, 2016. Citation on page 39.

____. **Ubiquitous computing fundamentals**. [S.l.]: CRC Press, 2018. Citations on pages 25 and 44.

LAZAR, J.; FENG, J. H.; HOCHHEISER, H. **Research methods in human-computer interaction**. Cambridge, MA, USA: Morgan Kaufmann, 2017. ISBN 978-0-12-805390-4. Citations on pages 40, 41, 42, 75, 87, 90, 104, 130, and 137.

LEWIS, C.; POLSON, P. G.; WHARTON, C.; RIEMAN, J. Testing a Walkthrough Methodology for Theory-based Design of Walk-up-and-use Interfaces. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 1990. (CHI '90), p. 235–242. ISBN 0-201-50932-6. Available: <http://doi.acm.org/10.1145/97243.97279>. Citation on page 38.

LEWIS, J. R. Usability: Lessons Learned ... and Yet to Be Learned. **International Journal of Human–Computer Interaction**, v. 30, n. 9, p. 663–684, 2014. Available: <www.dx.doi.org/10.1080/10447318.2014.930311>. Citations on pages 33, 35, 36, 37, 54, and 90.

LINDGAARD, G. Challenges to Assessing Usability in the Wild: A Case Study. **International Journal of Human–Computer Interaction**, Taylor & Francis, v. 31, n. 9, p. 618–631, 2015. Available: <https://doi.org/10.1080/10447318.2015.1065697>. Citation on page 39.

MAHATODY, T.; SAGAR, M.; KOLSKI, C. State of the Art on the Cognitive Walkthrough Method, Its Variants and Evolutions. **International Journal of Human-Computer Interaction**, v. 26, n. 8, p. 741–785, 2010. Available: <http://dx.doi.org/10.1080/10447311003781409>. Citation on page 38.

MAHMOUD, M.; HOSSEN, M. Z.; BARAKAT, H.; MANNAN, M.; YOUSSEF, A. Towards a Comprehensive Analytical Framework for Smart Toy Privacy Practices. In: **Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust**. New York, NY, USA: Association for Computing Machinery, 2018. (STAST '17), p. 64–75. ISBN 9781450363570. Available: <www.doi.org/10.1145/3167996.3168002>. Citations on pages 86 and 91.

MAYHEW, D. J. Usability + Persuasiveness + Graphic Design = eCommerce User Experience. In: JACKO, J. A. (Ed.). **Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications**. [S.l.]: CRC press, 2012. Citation on page 33.

MAZZIA, A.; LEFEVRE, K.; ADAR, E. The PViz Comprehension Tool for Social Network Privacy Settings. In: **Proceedings of the Eighth Symposium on Usable Privacy and Security**. New York, NY, USA: ACM, 2012. (SOUPS '12), p. 13:1–13:12. ISBN 978-1-4503-1532-6. Available: <http://doi.acm.org/10.1145/2335356.2335374>. Citations on pages 47, 113, and 117.

MCREYNOLDS, E.; HUBBARD, S.; LAU, T.; SARAF, A.; CAKMAK, M.; ROESNER, F. Toys That Listen: A Study of Parents, Children, and Internet-Connected Toys. In: **Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2017. p. 5197–5207. ISBN 978-1-4503-4655-9. Available: <www.doi.acm.org/10.1145/3025453.3025735>. Citations on pages 47 and 86.

____. Toys That Listen: A Study of Parents, Children, and Internet-Connected Toys. In: ____. **Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems**. New York, NY, USA: Association for Computing Machinery, 2017. p. 5197–5207. ISBN 9781450346559. Available: <https://doi.org/10.1145/3025453.3025735>. Citations on pages 68 and 70.

MEHTA, V.; GOOCH, D.; BANDARA, A.; PRICE, B.; NUSEIBEH, B. Privacy Care: A Tangible Interaction Framework for Privacy Management. **ACM Transactions on Internet Technology**, v. 21, n. 1, p. 1–32, Feb. 2021. ISSN 1533-5399, 1557-6051. Available: <https://dl.acm.org/doi/10.1145/3430506>. Citations on pages 25 and 44.

MESZAROS, J.; BUCHALCEVOVA, A. Introducing OSSF: A framework for online service cybersecurity risk management. **Computers & Security**, v. 65, p. 300–313, Mar. 2017. ISSN 01674048. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0167404816301791>. Citation on page 26.

MILLAR, S. A.; MARSHALL, T. P.; CARDON, N. A.; LLP, H.; STREET, G. **The Toy Association White Paper on Privacy & Data Security: New Possibilities and Perils**. 2017. Citation on page 71.

MORVILLE, P.; ROSENFELD, L. **Information Architecture for the World Wide Web**. [s.n.], 2006. ISBN 978-0-596-52734-1. Available: <http://shop.oreilly.com/product/9780596527341.do>. Citation on page 54.

MOTTA, R. C.; OLIVEIRA, K. M. de; TRAVASSOS, G. H. On Challenges in Engineering IoT Software Systems. In: **Proceedings of the XXXII Brazilian Symposium on Software Engineering**. New York, NY, USA: Association for Computing Machinery, 2018. (SBES '18), p. 42–51. ISBN 9781450365031. Available: <https://doi.org/10.1145/3266237.3266263>. Citation on page 68.

MOZILLA. ***Privacy Not Included: A Buyer's Guide for Connected Products**. 2018. Available: <www.foundation.mozilla.org/en/privacynotincluded/categories/Toys&Games/>. Citations on pages 76 and 91.

MULLER, M. J.; MATHESON, L.; PAGE, C.; GALLUP, R. Methods & Tools: Participatory Heuristic Evaluation. **interactions**, ACM, New York, NY, USA, v. 5, n. 5, p. 13–18, Sep. 1998. ISSN 1072-5520. Available: <http://doi.acm.org/10.1145/285213.285219>. Citation on page 101.

MURTAGH, F.; LEGENDRE, P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? **Journal of Classification**, v. 31, n. 3, p. 274–295, Oct. 2014. ISSN 0176-4268, 1432-1343. Available: <http://link.springer.com/10.1007/s00357-014-9161-z>. Citations on pages 59 and 122.

NELSON, B. Children's Connected Toys: Data Security and Privacy Concerns. **Office of Oversight and Investigations Minority Staff Report, US Senate Committee on Commerce, Science, and Transportation**, 2016. Citations on pages 68, 71, 74, and 75.

NIELSEN, J. Enhancing the Explanatory Power of Usability Heuristics. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 1994. (CHI '94), p. 152–158. ISBN 0-89791-650-6. Available: <www.doi.acm.org/10.1145/191666.191729>. Citations on pages 41, 42, 78, 89, and 90.

____. **Severity Ratings for Usability Problems**. 1995. Available: <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>. Citations on pages 78 and 82.

____. **Heuristic Evaluation: How-To**. 2018. Available: <https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>. Citations on pages 42, 88, 89, and 90.

____. **Severity Ratings for Usability Problems**. 2018. Available: <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>. Citation on page 42.

____. **Usability 101: Introduction to usability**. 2021. Available: <https://www.nngroup.com/articles/usability-101-introduction-to-usability>. Citations on pages 33, 34, and 38.

NIELSEN, J.; BUDIU, R. **Mobile usability**. [S.l.]: MITP-Verlags GmbH & Co. KG, 2013. Citations on pages 58, 61, and 62.

NIELSEN, J.; MOLICH, R. Heuristic evaluation of user interfaces. In: ACM. **Proceedings of the SIGCHI conference on Human factors in computing systems**. [S.l.], 1990. p. 249–256. Citation on page 41.

____. Heuristic evaluation of user interfaces. In: **Proceedings of the SIGCHI Conference on Human factors in computing systems Empowering people - CHI '90**. Seattle, Washington, United States: ACM Press, 1990. p. 249–256. ISBN 978-0-201-50932-8. Available: <http://portal.acm.org/citation.cfm?doid=97243.97281>. Citation on page 90.

NORMAN, D. A. **The Design of Everyday Things – Revised and expanded edition**. New York, USA: Basic Books, 2013. 347 p. ISBN 0465072992. Citation on page 39.

NURSE, J. R. C.; CREESE, S.; GOLDSMITH, M.; LAMBERTS, K. Guidelines for usable cybersecurity: Past and present. In: **2011 Third International Workshop on Cyberspace Safety and Security (CSS)**. [S.l.: s.n.], 2011. p. 21–26. Citations on pages 94, 95, and 101.

OATES, M.; AHMADULLAH, Y.; MARSH, A.; SWOOPES, C.; ZHANG, S.; BALEBAKO, R.; CRANOR, L. F. Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration. **Proceedings on Privacy Enhancing Technologies**, v. 2018, n. 4, 2018. Available: <https://content.sciendo.com/view/journals/popets/2018/4/article-p5.xml>. Citations on pages 26, 44, 45, 52, 53, 88, and 92.

Office of the Privacy Commissioner of Canada. **The Personal Information Protection and Electronic Documents Act (PIPEDA)**. 2021. Available: <https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/>. Citation on page 53.

PACI, F.; SQUICCIARINI, A.; ZANNONE, N. Survey on Access Control for Community-Centered Collaborative Systems. **ACM Comput. Surv.**, v. 51, n. 1, p. 6:1–6:38, Jan. 2018. ISSN 0360-0300. Available: <http://doi.acm.org/10.1145/3146025>. Citations on pages 26, 45, 46, 47, 52, 53, 69, 86, 87, 92, 93, 112, 113, and 116.

PARTNERS, P. T. **Jurassic Poke: Hacking a Dino toy | Pen Test Partners**. 2016. Available: <https://www.pentestpartners.com/security-blog/jurassic-poke-hacking-a-dino-toy/>. Citations on pages 72, 73, and 74.

PAZ, F.; PAZ, F. A.; POW-SANG, J. A.; COLLANTES, L. Usability Heuristics for Transactional Web Sites. In: **2014 11th International Conference on Information Technology: New Generations**. [S.l.: s.n.], 2014. p. 627–628. Citations on pages 97 and 102.

POLSON, P. G.; LEWIS, C.; RIEMAN, J.; WHARTON, C. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. **International Journal of Man-Machine Studies**, v. 36, n. 5, p. 741 – 773, 1992. ISSN 0020-7373. Available: <http://www.sciencedirect.com/science/article/pii/002073739290039N>. Citation on page 38.

PREECE, J.; SHARP, H.; ROGERS, Y. **Interaction Design: Beyond Human-Computer Interaction**. 4. ed. USA: John Wiley & Sons, 2015. 584 p. ISBN 978-1-119-08879-0. Citations on pages 34, 37, 38, 39, 40, and 42.

QUIñONES, D.; RUSU, C. How to develop usability heuristics: A systematic literature review. **Computer Standards & Interfaces**, v. 53, p. 89–122, Aug. 2017. ISSN 0920-5489. Available: <http://www.sciencedirect.com/science/article/pii/S0920548917301058>. Citations on pages 31, 111, and 138.

RAFFERTY, L. **A location privacy model and framework for mobile toy computing**. [S.l.]: University of Ontario Institute of Technology (Canada), 2015. Citation on page 70.

RAFFERTY, L.; FANTINATO, M.; HUNG, P. C. K. Privacy Requirements in Toy Computing. In: HUNG, P. C. K. (Ed.). **Mobile Services for Toy Computing**. Springer International Publishing, 2015. p. 141–173. ISBN 978-3-319-21322-4 978-3-319-21323-1. Available: <http://link.springer.com/chapter/10.1007/978-3-319-21323-1_8>. Citations on pages 47, 54, 93, 106, 125, and 140.

RAFFERTY, L.; HUNG, P.; FANTINATO, M.; PERES, S. M.; IQBAL, F.; KUO, S.-Y.; HUANG, S.-C. Towards a Privacy Rule Conceptual Model for Smart Toys. In: **Proceedings of the 50th Hawaii International Conference on System Sciences**. [s.n.], 2017. ISBN 978-0-9981331-0-2. Available: <www.scholarspace.manoa.hawaii.edu/handle/10125/41299>. Citations on pages 15, 52, 53, 54, 57, 58, 60, 61, 62, 64, 65, 66, 68, 70, 71, 86, 91, and 92.

REALPE, P. C.; COLLAZOS, C. A.; HURTADO, J.; GRANOLLERS, A. A Set of Heuristics for Usable Security and User Authentication. In: . ACM Press, 2016. p. 1–8. ISBN 978-1-4503-4119-6. Available: <http://dl.acm.org/citation.cfm?doid=2998626.2998662>. Citations on pages 94, 95, 100, and 103.

REEDER, R. W. **Expandable Grids: A user interface visualization technique and a policy semantics to support fast, accurate security and privacy policy authoring**. Phd Thesis (PhD Thesis) — Carnegie Mellon University, 2008. Citations on pages 47, 56, 113, 114, and 115.

REEDER, R. W.; BAUER, L.; CRANOR, L. F.; REITER, M. K.; BACON, K.; HOW, K.; STRONG, H. Expandable Grids for Visualizing and Authoring Computer Security Policies. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2008. (CHI '08), p. 1473–1482. ISBN 978-1-60558-011-1. Available: <http://doi.acm.org/10.1145/1357054.1357285>. Citations on pages 47, 56, 112, 113, and 114.

REYNAGA, G.; CHIASSON, S.; OORSCHOT, P. C. van. Heuristics for the evaluation of captchas on smartphones. In: . ACM Press, 2015. p. 126–135. ISBN 978-1-4503-3643-7. Available: <http://dl.acm.org/citation.cfm?doid=2783446.2783583>. Citations on pages 96, 102, and 105.

RODE, J.; JOHANSSON, C.; DIGIOIA, P.; FILHO, R. S.; NIES, K.; NGUYEN, D. H.; REN, J.; DOURISH, P.; REDMILES, D. Seeing Further: Extending Visualization As a Basis for Usable Security. In: **Proceedings of the Second Symposium on Usable Privacy and Security**. New York, NY, USA: ACM, 2006. (SOUPS '06), p. 145–155. ISBN 978-1-59593-448-2. Available: <http://doi.acm.org/10.1145/1143120.1143138>. Citations on pages 47, 112, 113, and 114.

RUSU, C.; RONCAGLIOLO, S.; RUSU, V.; COLLAZOS, C. A Methodology to Establish Usability Heuristics. p. 4, 2011. Citation on page 100.

SALGADO, A. d. L.; FORTES, R. P. d. M.; OLIVEIRA, R. R. d.; FREIRE, A. P. Usability heuristics on parental privacy controls for smart toys: From an exploratory map to a confirmatory research. **Electronic Commerce Research and Applications**, v. 42, p. 100984, 2020. ISSN 1567-4223. Available: <https://www.sciencedirect.com/science/article/pii/S1567422320300612>. Citation on page 85.

SALGADO, A. d. L.; FORTES, R. Pontin de M.; HUNG, P. C.; MOREIRA, D. d. A. A Method for Classifying Usability Findings to Enhance Validation of New Heuristics. **Revista de Sistemas e Computação-RSC**, v. 9, n. 1, 2019. Citations on pages 29 and 139.

SALGADO, A. d. L.; SINGH, B.; HUNG, P. C. K.; JIANG, A.; LIU, Y.-H.; WHELER, A. P. d. A.; GABER, H. A. Preliminary Tendencies of Users' Expectations about Privacy on Connected-Autonomous Vehicles. In: **2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)**. [S.l.: s.n.], 2020. p. 296–301. Citations on pages 30 and 139.

SALGADO, A. de L.; AMARAL, L. A. do; CASTRO, P. C.; FORTES, R. P. de M. Designing for Parental Control: Enriching Usability and Accessibility in the Context of Smart Toys. In: TANG, J. K.; HUNG, P. C. K. (Ed.). **Computing in Smart Toys**. Cham: Springer International Publishing, 2017. p. 103–125. ISBN 978-3-319-62072-5. DOI: 10.1007/978-3-319-62072-5_7. Available: <https://doi.org/10.1007/978-3-319-62072-5_7>. Citations on pages 48, 92, and 93.

SALGADO, A. de L.; DIAS, F. S.; MATTOS, J. P. R.; FORTES, R. P. de M.; HUNG, P. C. K. Smart toys and children's privacy: usable privacy policy insights from a card sorting

experiment. In: **Proceedings of the 37th ACM International Conference on the Design of Communication**. Portland Oregon: ACM, 2019. p. 1–8. ISBN 978-1-4503-6790-5. Available: <https://dl.acm.org/doi/10.1145/3328020.3353951>. Citations on pages 30, 51, 93, and 139.

SALGADO, A. de L.; FEDERICI, F. M.; FORTES, R. P. de M.; MOTTI, V. G. Startup Workplace, Mobile Games, and Older Adults: A Practical Guide on UX, Usability, and Accessibility Evaluation. In: **Proceedings of the 37th ACM International Conference on the Design of Communication**. New York, NY, USA: Association for Computing Machinery, 2019. (SIGDOC '19). ISBN 9781450367905. Available: <https://doi.org/10.1145/3328020.3353948>. Citations on pages 30, 138, and 139.

SALGADO, A. de L.; FREIRE, A. P. Heuristic Evaluation of Mobile Usability: A Mapping Study. In: KUROSU, M. (Ed.). **Human-Computer Interaction. Applications and Services**. Cham: Springer International Publishing, 2014. p. 178–188. ISBN 978-3-319-07227-2. Available: <www.link.springer.com/chapter/10.1007/978-3-319-07227-2_18>. Citations on pages 41, 87, and 90.

SALGADO, A. de L.; PEREIRA, F. H. S.; FREIRE, A. P. User-Centred Design and Evaluation of Information Architecture for Information Systems. In: **Handbook of Research on Information Architecture and Management in Modern Organizations**. [S.l.]: IGI Global, 2016. p. 219–236. Citation on page 55.

SALGADO, A. de L.; RODRIGUES, S. S.; FORTES, R. P. M. Evolving Heuristic Evaluation for Multiple Contexts and Audiences: Perspectives from a Mapping Study. In: **Proceedings of the 34th ACM International Conference on the Design of Communication**. New York, NY, USA: ACM, 2016. (SIGDOC '16), p. 19:1–19:8. ISBN 978-1-4503-4495-1. Available: <http://doi.acm.org/10.1145/2987592.2987617>. Citations on pages 26, 41, 87, and 90.

SALGADO, A. de L.; SANTOS, F. de S.; FORTES, R. P. de M.; HUNG, P. C. K. Guiding Usability Newcomers to Understand the Context of Use: Towards Models of Collaborative Heuristic Evaluation. In: WONG, R.; CHI, C.-H.; HUNG, P. C. K. (Ed.). **Behavior Engineering and Applications**. Cham: Springer International Publishing, 2018. p. 149–168. ISBN 978-3-319-76430-6. Available: <https://doi.org/10.1007/978-3-319-76430-6_7>. Citations on pages 29, 36, 43, and 139.

SANTOS, F. de S.; SALGADO, A. de L.; FORTES, R. P. de M. Um mapeamento sistemático sobre acessibilidade e usabilidade no desenvolvimento de jogos digitais para idosos. **iSys-Brazilian Journal of Information Systems**, v. 11, n. 2, p. 63–90, 2018. Citation on page 29.

SASSE, M. A.; SMITH, M. The Security-Usability Tradeoff Myth [Guest editors' introduction]. **IEEE Security Privacy**, v. 14, n. 5, p. 11–13, Sep. 2016. ISSN 1540-7993. Citations on pages 45, 87, 88, and 90.

SASSE, M. A.; SMITH, M.; HERLEY, C.; LIPFORD, H.; VANIEA, K. Debunking Security-Usability Tradeoff Myths. **IEEE Security Privacy**, v. 14, n. 5, p. 33–39, 2016. Citations on pages 71, 72, 74, and 75.

SAURO, J. The Relationship Between Problem Frequency and Problem Severity in Usability Evaluations. **J. Usability Studies**, v. 10, n. 1, p. 17–25, Nov. 2014. ISSN 1931-3357. Available: <http://dl.acm.org/citation.cfm?id=2817310.2817312>. Citation on page 42.

____. **MeasuringU: Rating the Severity of Usability Problems**. 2018. Available: <https://measuringu.com/rating-severity/>. Citation on page 42.

SCHAUB, F.; BALEBAKO, R.; CRANOR, L. F. Designing Effective Privacy Notices and Controls. **IEEE Internet Computing**, v. 21, n. 3, p. 70–77, May 2017. ISSN 1089-7801. Citations on pages 25, 44, 45, 52, 53, 54, 57, and 92.

SCHLEGEL, R.; KAPADIA, A.; LEE, A. J. Eyeing Your Exposure: Quantifying and Controlling Information Sharing for Improved Privacy. In: **Proceedings of the Seventh Symposium on Usable Privacy and Security**. New York, NY, USA: ACM, 2011. (SOUPS '11), p. 14:1–14:14. ISBN 978-1-4503-0911-0. Available: <http://doi.acm.org/10.1145/2078827.2078846>. Citations on pages 47, 112, 113, and 116.

SILVERMAN, D. **Qualitative Research**. [S.l.]: SAGE, 2016. Google-Books-ID: 9FALDAAAQBAJ. ISBN 978-1-4739-8484-4. Citation on page 111.

SLEPCHUK, A. N.; MILNE, G. R. Informing the design of better privacy policies. **Current Opinion in Psychology**, v. 31, p. 89–93, Feb. 2020. ISSN 2352250X. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2352250X19301319>. Citations on pages 25, 26, 45, and 137.

Somerset Recon Inc. **Hello Barbie Initial Security Analysis**. [S.l.], 2016. 30 p. Available: <https://static1.squarespace.com/static/543effd8e4b095fba39dfe59/t/56a66d424bf1187ad34383b2/1453747529070/HelloBarbieSecurityAnalysis.pdf>. Citations on pages 72 and 73.

SQUICCIARINI, A. C.; LIN, D.; SUNDARESWARAN, S.; WEDE, J. Privacy Policy Inference of User-Uploaded Images on Content Sharing Sites. **IEEE Transactions on Knowledge and Data Engineering**, v. 27, n. 1, p. 193–206, Jan. 2015. ISSN 1041-4347. Citations on pages 52 and 53.

STILL, J. D. Cybersecurity Needs You! **interactions**, v. 23, n. 3, p. 54–58, Apr. 2016. ISSN 1072-5520. Available: <http://doi.acm.org/10.1145/2899383>. Citations on pages 45, 52, 55, and 88.

STREIFF, J.; DAS, S.; CANNON, J. Overpowered and Underprotected Toys Empowering Parents with Tools to Protect Their Children. In: **IEEE HUMANS AND CYBER SECURITY WORKSHOP (HACS 2019)**. [S.l.: s.n.], 2019. Citation on page 92.

SUN, Y.; HUANG, Z.; KE, C. Obtaining P3P privacy policies for composite services. **The Scientific World Journal**, Hindawi, v. 2014, 2014. Citation on page 115.

TESSIAN; HANCOCK, J. **Psycology of Human Error: Understand the mistakes that compromise your company's cybersecurity**. [S.l.]: Tessian, 2020. Citation on page 45.

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION. **REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)**. [S.l.]: Official Journal of the European Union, 2016. Citations on pages 44, 46, 68, and 89.

The European Parliament and The Council of The European Union. **Guidelines on Transparency under Regulation 2016/679 (WP260 rev.01)**. [S.l.]: European Commission website, 2017. Citation on page 45.

THEOFANOS, M. Is Usable Security an Oxymoron? **Computer**, v. 53, n. 2, p. 71–74, 2020. Citation on page 141.

TOGNAZZINI, B. **First Principles of Interaction Design (Revised & Expanded)**. 2014. Available: <http://asktog.com/atc/principles-of-interaction-design/>. Citation on page 101.

TRACTINSKY, N. The Usability Construct: A Dead End? **Human–Computer Interaction**, Taylor & Francis, v. 33, n. 2, p. 131–177, 2018. Available: <https://doi.org/10.1080/07370024.2017.1298038>. Citation on page 137.

ULANOFF, L. **Smart Dino Toy is powered by a super computer**. 2015. Available: <https://mashable.com/2015/02/16/smart-dino-toy-powered-by-ibm-watson/>. Citation on page 68.

UNICEF Innovation. **Memorandum on Artificial Intelligence and Child Rights**. 2019. Available: <www.unicef.org/innovation/reports/memoAIchildrights>. Citation on page 92.

UNION, E.; EUROPARAT (Ed.). **Handbook on European data protection law**. 2018 edition. ed. Luxembourg: Publications Office of the European Union, 2018. (Handbook / FRA, European Union Agency for Fundamental Rights). ISBN 978-92-871-9849-5 978-92-9491-903-8 978-92-9491-901-4. Citation on page 89.

VALENTE, J.; CARDENAS, A. A. Security & Privacy in Smart Toys. In: **Proceedings of the 2017 Workshop on Internet of Things Security and Privacy**. New York, NY, USA: ACM, 2017. (IoTS&P '17), p. 19–24. ISBN 978-1-4503-5396-0. Available: <www.doi.acm.org/10.1145/3139937.3139947>. Citations on pages 86 and 91.

____. Security & Privacy in Smart Toys. In: **Proceedings of the 2017 Workshop on Internet of Things Security and Privacy**. New York, NY, USA: Association for Computing Machinery, 2017. (IoTS&P '17), p. 19–24. ISBN 9781450353960. Available: <www.doi.org/10.1145/3139937.3139947>. Citation on page 91.

VALLEE, H. Quay-de la; SELBY, P.; KRISHNAMURTHI, S. On a (Per)Mission: Building Privacy Into the App Marketplace. In: . ACM Press, 2016. p. 63–72. ISBN 978-1-4503-4564-4. Available: <http://dl.acm.org/citation.cfm?doid=2994459.2994466>. Citations on pages 53, 54, 56, and 58.

VILBERGSDOTTIR, S. G.; HVANNBERG, E. T.; LAW, E. L.-C. Assessing the reliability, validity and acceptance of a classification scheme of usability problems (CUP). **Journal of Systems and Software**, v. 87, p. 18–37, 2014. ISSN 0164-1212. Available: <https://www.sciencedirect.com/science/article/pii/S0164121213002136>. Citation on page 39.

von Solms, R.; van Niekerk, J. From information security to cyber security. **Computers & Security**, v. 38, p. 97–102, 2013. ISSN 0167-4048. Cybercrime in the Digital Economy. Available: <https://www.sciencedirect.com/science/article/pii/S0167404813000801>. Citation on page 46.

VéLIZ, C. Privacy and digital ethics after the pandemic. **Nature Electronics**, v. 4, n. 1, p. 10–11, Jan. 2021. ISSN 2520-1131. Available: <http://www.nature.com/articles/s41928-020-00536-y>. Citation on page 25.

WANG, Y.; GOU, L.; XU, A.; ZHOU, M. X.; YANG, H.; BADENES, H. VeilMe: An Interactive Visualization Tool for Privacy Configuration of Using Personality Traits. In: **Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2015. (CHI '15), p. 817–826. ISBN 978-1-4503-3145-6. Available: <http://doi.acm.org/10.1145/2702123.2702293>. Citations on pages 47, 113, and 117.

WARREN, S. D.; BRANDEIS, L. D. The Right to Privacy. **Harvard Law Review**, v. 4, n. 5, p. 193–220, 1890. ISSN 0017811X. Available: <http://www.jstor.org/stable/1321160>. Citations on pages 43 and 44.

WASH, R.; ZURKO, M. E. Usable Security. **IEEE Internet Computing**, v. 21, n. 3, p. 19–21, May 2017. ISSN 1089-7801. Citations on pages 45 and 88.

WHARTON, C.; RIEMAN, J.; LEWIS, C.; POLSON, P. Usability Inspection Methods. In: NIELSEN, J.; MACK, R. L. (Ed.). New York, NY, USA: John Wiley & Sons, Inc., 1994. chap. The Cognitive Walkthrough Method: A Practitioner's Guide, p. 105–140. ISBN 0-471-01877-5. Available: <http://dl.acm.org/citation.cfm?id=189200.189214>. Citation on page 38.

WHITTEN, A.; TYGAR, J. D. Why Johnny Can'T Encrypt: A Usability Evaluation of PGP 5.0. In: **Proceedings of the 8th Conference on USENIX Security Symposium - Volume 8**. Berkeley, CA, USA: USENIX Association, 1999. (SSYM'99), p. 14–14. Available: <http://dl.acm.org/citation.cfm?id=1251421.1251435>. Citations on pages 34, 46, 55, and 140.

WIRESHARK Developer's Guide. 1998. Available: <https://www.wireshark.org/docs/wsdg_html_chunked/index.html>. Citations on pages 76 and 78.

WOHLIN, C. Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. In: **Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering**. New York, NY, USA: ACM, 2014. (EASE '14), p. 38:1–38:10. ISBN 978-1-4503-2476-2. Available: <http://doi.acm.org/10.1145/2601248.2601268>. Citations on pages 87, 93, 94, 95, and 113.

XIA, Z.; WANG, X.; SUN, X.; WANG, Q. A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data. **IEEE Transactions on Parallel and Distributed Systems**, v. 27, n. 2, p. 340–352, Feb. 2016. ISSN 1045-9219. Citation on page 92.

YANG, J.; LU, Z.; WU, J. Smart-toy-edge-computing-oriented data exchange based on blockchain. **Journal of Systems Architecture**, v. 87, p. 36 – 48, 2018. ISSN 1383-7621. Available: <www.sciencedirect.com/science/article/pii/S1383762118300638>. Citation on page 91.

YANKSON, B.; IQBAL, F.; HUNG, P. C. K. Privacy Preservation Framework for Smart Connected Toys. In: ____. **Computing in Smart Toys**. Cham: Springer International Publishing, 2017. p. 149–164. ISBN 978-3-319-62072-5. Available: <doi.org/10.1007/978-3-319-62072-5_9>. Citation on page 92.

____. 4P Based Forensics Investigation Framework for Smart Connected Toys. In: **Proceedings of the 15th International Conference on Availability, Reliability and Security**. New York, NY, USA: Association for Computing Machinery, 2020. (ARES '20). ISBN 9781450388337. Available: <https://doi-org.ez67.periodicos.capes.gov.br/10.1145/3407023.3409213>. Citations on pages 70 and 71.

YANKSON, B.; IQBAL, F.; LU, Z.; WANG, X.; HUNG, P. Modeling Privacy Preservation in Smart Connected Toys by Petri-Nets. In: **Proceedings of the 52nd Hawaii International Conference on System Sciences**. [s.n.], 2019. Available: <www.scholarspace.manoa.hawaii.edu/handle/10125/59610>. Citation on page 92.

YANKSON, B.; SALGADO, A. L.; FORTES, R. P. Recommendations to enhance privacy and usability of smart toys. In: **Proceedings of the 54th Hawaii International Conference on System Sciences**. [S.l.: s.n.], 2021. p. 1868. Citations on pages 31, 67, 125, and 139.

YERATZIOTIS, A.; GREUNEN, D. V.; POTTAS, D. Recommendations for usable security in online health social networks. In: **2011 6th International Conference on Pervasive Computing and Applications**. [S.l.: s.n.], 2011. p. 220–226. Citations on pages 96 and 103.

YERATZIOTIS, A.; POTTAS, D.; GREUNEN, D. V. A Usable Security Heuristic Evaluation for the Online Health Social Networking Paradigm. **International Journal of Human–Computer Interaction**, v. 28, n. 10, p. 678–694, Oct. 2012. ISSN 1044-7318. Available: <https://doi.org/10.1080/10447318.2011.654202>. Citations on pages 94, 95, 103, and 104.

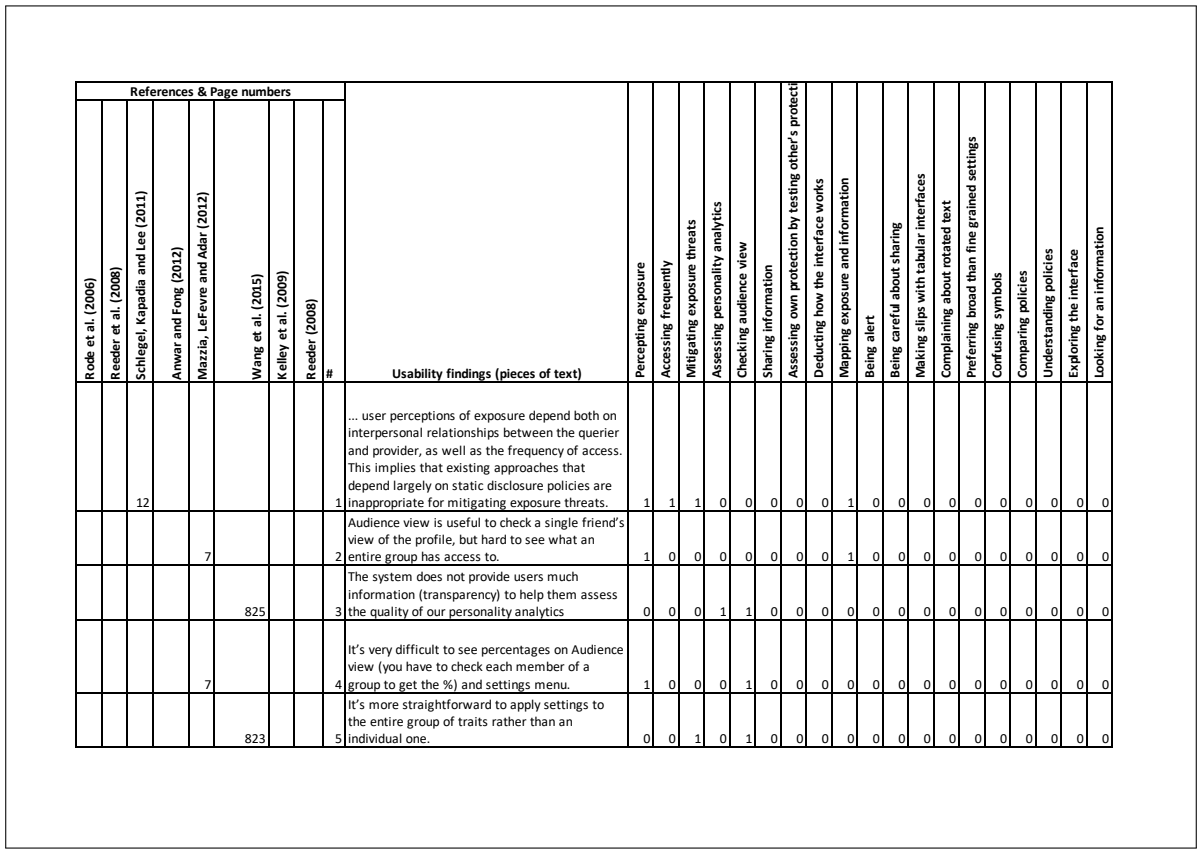
YERATZIOTIS, A.; POTTAS, D.; GREUNEN, D. van. A three-phase process to develop heuristics. In: **Proceedings of the 13th ZA-WWW Conference**. [S.l.: s.n.], 2011. Citation on page 100.

YUSOP, N. S. M.; GRUNDY, J.; VASA, R. Reporting Usability Defects: A Systematic Literature Review. **IEEE Transactions on Software Engineering**, v. 43, n. 9, p. 848–867, 2017. Citation on page 39.

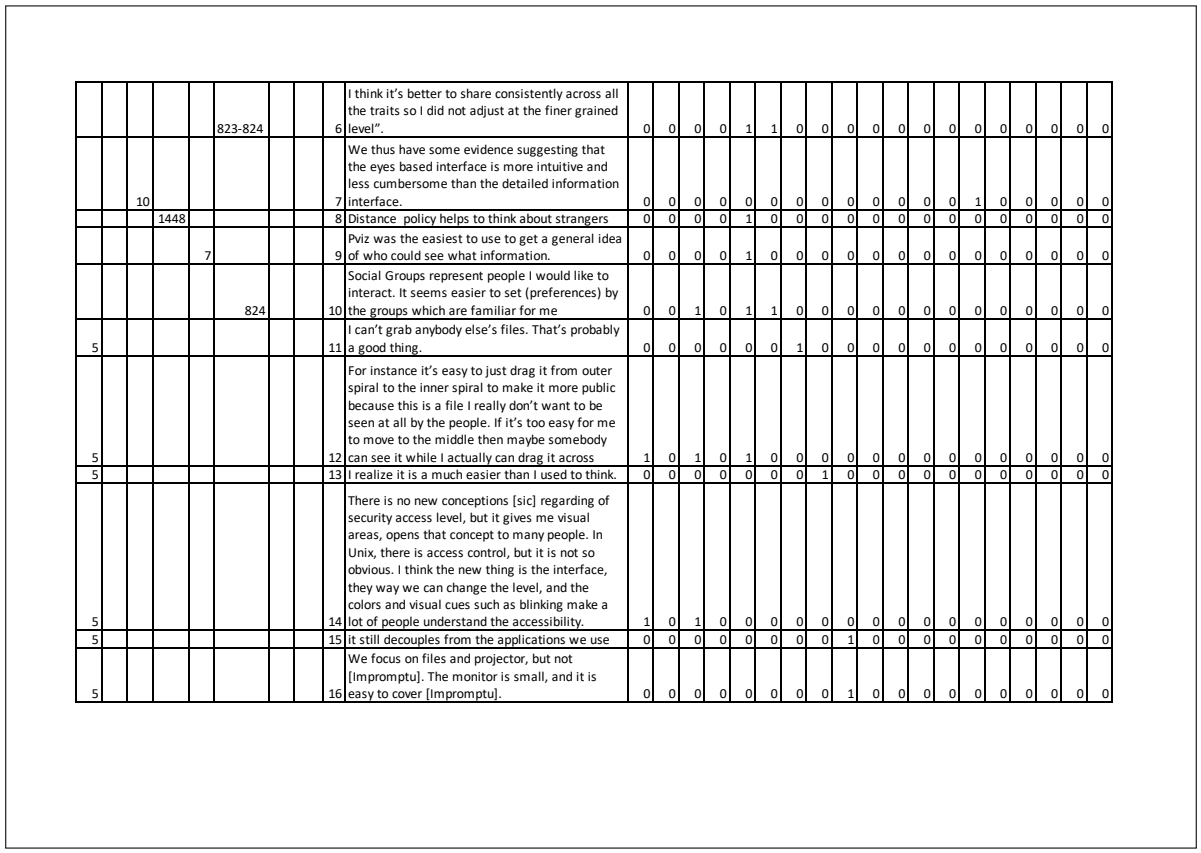
APPENDIX

A

DATASET OF TRANSCRIPTS FROM THE THEMATIC ANALYSIS.



References & Page numbers									Usability findings (pieces of text)	Percepting exposure	Accessing frequently	Mitigating exposure threats	Assessing personality analytics	Checking audience view	Sharing information	Assessing own protection by testing other's protecti	Deducting how the interface works	Mapping exposure and information	Being alert	Being careful about sharing	Making slips with tabular interfaces	Complaining about rotated text	Preferring broad than fine grained settings	Confusing symbols	Comparing policies	Understanding policies	Exploring the interface	Looking for an information
Rode et al. (2006)	Reeder et al. (2008)	Schlegel, Kapadia and Lee (2011)	Anwar and Fong (2012)	Mazzia, LeFevre and Adar (2012)	Wang et al. (2015)	Kelley et al. (2009)	Reeder (2008)	#																				
		12						1	… user perceptions of exposure depend both on interpersonal relationships between the querier and provider, as well as the frequency of access. This implies that existing approaches that depend largely on static disclosure policies are inappropriate for mitigating exposure threats.	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
				7				2	Audience view is useful to check a single friend's view of the profile, but hard to see what an entire group has access to.	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
					825			3	The system does not provide users much information (transparency) to help them assess the quality of our personality analytics	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
				7				4	It's very difficult to see percentages on Audience view (you have to check each member of a group to get the %) and settings menu.	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
					823			5	It's more straightforward to apply settings to the entire group of traits rather than an individual one.	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0



					823-824			6	I think it's better to share consistently across all the traits so I did not adjust at the finer grained level".	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		10						7	We thus have some evidence suggesting that the eyes based interface is more intuitive and less cumbersome than the detailed information interface.	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
			1448					8	Distance policy helps to think about strangers	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
				7				9	Pviz was the easiest to use to get a general idea of who could see what information.	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
					824			10	Social Groups represent people I would like to interact. It seems easier to set (preferences) by the groups which are familiar for me	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5								11	I can't grab anybody else's files. That's probably a good thing.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
5								12	For instance it's easy to just drag it from outer spiral to the inner spiral to make it more public because this is a file I really don't want to be seen at all by the people. If it's too easy for me to move to the middle then maybe somebody can see it while I actually can drag it across	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5								13	I realize it is a much easier than I used to think.	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
5								14	There is no new conceptions [sic] regarding of security access level, but it gives me visual areas, opens that concept to many people. In Unix, there is access control, but it is not so obvious. I think the new thing is the interface, they way we can change the level, and the colors and visual cues such as blinking make a lot of people understand the accessibility.	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5								15	it still decouples from the applications we use	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
5								16	We focus on files and projector, but not [Impromptu]. The monitor is small, and it is easy to cover [Impromptu].	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

5								17	whose files are whose...you know what's important to share." Further, Impromptu "emphasizes what to explore, what's important.	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6								18	Yes, because it's saying read-only but, you know, initially it was – change the whole document on this – in this area	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6								19	It made it clear if someone can see or view files, but just a little bit. The visualization was a little helpful.	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6								20	but maybe if there could be a mechanism to see whose reading your file right now. Does that exist?	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6								21	Someone asks who opened my file? It looks like someone edited my file!	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
6								22	OOOH, the ring is who has it open, or who has ownership of it. Cool!	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6								23	Oh. Oh, cool. The little – so the little outer ring on the dots is. like, who' s got it open, or who' s got ownership of it right now. Yes, yes. Oh, that's cool. {Sound of computer chime}. So, I guess I'm the only person that actually went over (inaudible) so I can just trim some of my stuff off (inaudible), if that' s cool	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6								24	I was very careful. I didn't give other participants 'write' access to my files. Others were not so careful. They left some files writable. I changed one participant's justification to make him greedy. That is one of the scary things, which partly makes this an interesting scenario	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
6								25	Unsure who can access - there is no control access	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
6								26	It would be good if you could grant very limited access to just one person—a finer granularity that is not just for all people but for a specific user.	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	11							27	First were off-by-one errors, in which the participant had lined up the mouse in a desired column, then moved it to find the desired row, only to have the mouse slip, unnoticed, from the desired column into an adjacent column. Participants then made the right changes to the wrong principal.	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
	11							28	The second common type of error, which was responsible for the lower-than-expected Grid accuracy scores in the conflict-simple tasks, occurred in searching for resources. Some resources in our teaching assistant scenario had similar names, for example, a folder called "Assignment 1" and a file called "assignment1.pdf." Some participants, when searching for the keyword "assignment," came across the "assignment1.pdf" file first and assumed they had found the correct resource, although the task required changes to the folder.	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	11							29	The third common type of error resulted from the 90-degree angle of the text in the tree at the top of the Grid interface. Participants reported that this rotated text was difficult to read.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
			1448					30	after using the tool, I understand that lot of people can access my profile than what I thought	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
			1448					31	The tool is very helpful, provides more control, and without (the tool) it is impossible to understand who has access to what profile item	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			1448					32	The tool would have been more useful if the nodes were color coded based on access scenarios.	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
				9				33	I'd like to merge and reconfigure some of the groups after they were initially generated.	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

				10				34	It reminded me that I don't really manage my privacy settings as well as I should since everyone could access all of my information.	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
					823			35	I found it easiest to do a quick broad brush setting, rather than having to tinker with each trait, which was fiddly and time consuming.	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
						10		36	Several participants were confused by the symbols we used to indicate opt-in and opt-out.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
						10		37	...symbol...one participant did not understand what "out"meant, saying, "I've been messing things up because I thought 'out' meant 'out of the question.'"	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
						11		38	several participants in the lab study were completely unfamiliar with the terms opt-out and opt-in, and they assumed that the terms meant exactly the same thing.	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
						11		39	Participants also realized the benefits of the label for comparison: "This may actually be the biggest advantage of this system because you can put down two polices that are formatted the same and see the exact differences between them. It's really easy."	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
						11		40	"I guess I'll look to see which policy has more blue," exactly capturing one of our intended design goals.	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

							148	41	"The policy itself is not that – very clear at all."	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
							148	42	"Yeah, I'm just confused right now what I'm looking at. OK, so there's these color-coded things…"	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
							148	43	"I'm not sure how to find this, I'm just kind of looking around."	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
							148	44	"This is a really convoluted presentation."	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
							148	45	"There's no mention of credit cards at all, which is kind of strange, so I'm kind of playing the 'is it under this heading [referring to statement header boxes]' game again."	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1

APPENDIX

B

DATASET OF USABILITY PROBLEMS FOUND BY PARTICIPANTS IN THE VALIDATION PROCESS.

participant_id	expertise_years_linkedin	group	n	p1	p2	p3	p4
1	0	0	2	4	2	0	0
2	6	0	5	0	0	2	3
3	5	0	5	0	2	0	0
4	2	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	0	0	3	0	0	0	0
7	0	0	4	0	0	0	0
8	3	0	3	0	0	2	3
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	1	0	0	0	0	0	0
12	5	0	3	0	2	0	0
13	1	0	0	0	0	0	0
14	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0
16	1	0	3	0	0	2	0
17	4	1	2	0	0	2	0
18	1	1	0	0	0	0	0
19	0	1	3	0	0	0	0
20	3	1	3	0	2	0	0
21	0	1	1	0	0	0	0
22	0	1	3	0	0	0	0
23	3	1	3	0	0	0	0
24	0	1	2	0	0	0	0
25	0	1	0	0	0	0	0
26	0	1	8	0	0	0	0
27	2	1	3	0	0	0	0
28	1	1	7	0	0	0	0
29	0	1	0	0	0	0	0
30	0	1	0	0	0	0	0
31	0	1	4	0	0	0	0
32	4	1	7	0	0	0	3

participant_id	p5	p6	p7	p8	p9	p10	p11	p12
1	0	0	0	0	0	0	0	0
2	3	4	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	4	4	3	0	0	0
7	0	0	0	0	3	4	2	4
8	3	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	2	0
20	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0
23	0	0	0	4	3	0	0	0
24	0	0	0	0	0	0	0	4
25	0	0	0	0	0	0	0	0
26	0	0	0	0	3	4	2	4
27	0	0	0	0	0	0	0	4
28	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0
31	0	0	0	0	3	4	0	0
32	0	4	0	0	3	0	0	0

participant_id	p13	p14	p15	p16	p17	p18	p19	p20
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	3
3	0	0	3	3	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	0	0	0	0	3	0	0	0
13	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0
16	0	0	0	3	0	0	0	0
17	2	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0
19	0	0	0	0	3	0	0	0
20	0	0	3	0	0	0	0	0
21	0	0	0	0	0	0	0	0
22	2	0	3	0	0	0	0	0
23	0	0	0	3	0	0	0	0
24	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0
26	0	4	3	3	3	0	0	0
27	0	0	0	0	3	0	0	0
28	0	0	0	3	0	3	1	3
29	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0
31	0	0	0	3	0	0	0	0
32	0	0	0	0	0	0	0	0

participant_id	p21	p22	p23	p24	p25	p26	p27	p28
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	3
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	3	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0
19	0	0	0	4	0	0	0	0
20	0	0	0	4	0	0	0	0
21	0	0	0	0	4	0	0	0
22	3	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0
24	0	0	0	0	4	0	0	0
25	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0
27	0	0	0	4	0	0	0	0
28	3	1	1	0	0	0	0	0
29	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0
31	0	0	0	4	0	0	0	0
32	0	0	0	0	0	3	3	3

participant_id	p29	p30	p31
1	0	0	0
2	0	0	0
3	3	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	0
14	0	0	0
15	0	0	0
16	3	0	0
17	0	0	0
18	0	0	0
19	0	0	0
20	0	0	3
21	0	0	0
22	0	0	0
23	0	1	0
24	0	0	0
25	0	0	0
26	0	0	0
27	0	0	0
28	0	0	0
29	0	0	0
30	0	0	0
31	0	0	0
32	3	0	0


ICMC
USP
SÃO CARLOS