Faculdade de Ciências Exatas e da Engenharia

Mestrado em Engenharia Informática

Privacy in the Internet of Things: Fostering User Empowerment through Digital Literacy

Nelson Ernesto Freitas Vieira

Orientado por:

Mary Alejandra Luiz Barreto

Constituição do júri de provas públicas:

Nome completo (categoria), Presidente

Nome completo (categoria), Vogal

Nome completo (categoria), Vogal

Saturday 9th September, 2023

Resumo

Os dispositivos da Internet das coisas estão por todo o lado, desde o nascimento da computação ubíqua que se prevê que a vida quotidiana do ser humano contenha milhões de dispositivos que controlam todos os aspectos da nossa vida. Hoje em dia, temos carros inteligentes, casas inteligentes, cidades inteligentes, dispositivos portáteis, entre outros, que utilizam vários tipos de dispositivos e vários tipos de redes para comunicar. Estes dispositivos criam novas formas de recolha e tratamento de dados pessoais de utilizadores e não utilizadores. A maioria dos utilizadores finais nem sequer tem conhecimento ou tem pouco controlo sobre as informações que estão a ser recolhidas por estes sistemas. Este trabalho adopta uma abordagem holística a este problema, começando por fazer uma revisão da literatura, depois conduzindo um inquérito para saber mais sobre o conhecimento geral do público e, finalmente, com base na informação recolhida, é proposto um sistema que dá aos utilizadores informações sobre os dispositivos que estão nas proximidades e como proteger os dados que não querem partilhar com esses dispositivos. Este sistema é capaz de detetar que tipo de dispositivos estão nas proximidades, que tipo de dados são recolhidos por esses dispositivos, mostrar opções de privacidade ao utilizador quando é possível fazê-lo e o que pode ser feito para proteger dados indesejados de serem recolhidos.

Keywords: privacidade · Internet das Coisas · computação ubíqua · assistente de privacidade

Abstract

Internet of things devices are everywhere, since the birth of ubiquitous computing that human every day life is envisioned containing millions of devices that control every aspect of our lives. Today we have smart cars, smart houses, smart cities, wearables among other things that use various types of devices and various types of networks to communicate. These devices create new ways of collecting and process personal data from users and non-users. Most end users are not even aware or have little control over the information that is being collected by these systems. This work takes an holistic approach to this problem by first doing a literature review, then conducting a survey to learn more about the general knowledge of the public, and finally, based on the information gathered, a system is proposed that gives users information about the devices that are nearby and how to protect the data that they do not want to share with these devices, this system is capable of detecting what type of devices are nearby, what kind of data is collected by these devices, show privacy choices to the user when it is possible to do so and what can be done to protect unwanted data from being collected.

Keywords: privacy · Internet of Things · ubiquitous computing · privacy assistant

Acknowledgment

Agradeço ...

Table of Contents

List of Figures ... vi

List of Tables ... vii

1 Introduction ... 1
2 State of the Art ... 5
2.1 Privacy Paradox ... 5
2.2 Differential Privacy ... 5
2.3 Privacy in IoT: Approaches ... 6
2.4 Proposed Solutions ... 9
2.4.1 User Awareness ... 9
2.4.2 Legislation ... 10
2.4.3 Privacy through Security ... 10
2.4.4 Architecture / Frameworks ... 11
2.4.5 Blockchain ... 12
2.4.6 Privacy Assistants ... 13
2.4.7 Other Proposals ... 15
2.5 Privacy Challenges ... 15
3 Methodology ... 18
3.1 Stage 1: User perceptions ... 19
3.2 Stage 2: An application of theory into practice ... 22
4 Results ... 28
4.1 Stage 1: Survey ... 28
4.2 Stage 2: Application ... 38
5 Discussion ... 40
5.1 Systematic literature review Findings ... 40
5.2 Study analysis and an application of privacy literacy ... 46
6 Challenges ... 48
7 Future work ... 49
8 Conclusion ... 50

References ... 51

List of Figures

1 Genre distribution of participants. 21
2 Age ranges of participants. 22
3 Education qualifications distribution of participants. 22
4 Low level prototype of (a) homepage, (b) about and (c) FAQ pages. 25
5 Medium level prototype of (a) homepage, (b) about and (c) FAQ pages. 26
6 High level prototype of (a) homepage, (b) about and (c) FAQ pages. 26
7 Live version of the application with pages: (a) homepage, (b) about and (c) FAQ pages. 27
8 Perspectives of participants on the importance of personal data privacy. 28
9 Participant responses indicating whether they know techniques to guarantee privacy and the protection of their data when using the internet. 29
10 Participant responses indicating whether they agree or disagree with these statements. . . 29
11 Responses to security and privacy being synonymous. 30
12 Familiarity with general IT terms. 30
13 Responses related to phone usage. 31
14 Participant responses indicating whether they agree or disagree with these statements. . . 31
15 Responses related to privacy regulations. 32
16 Awareness of the duties of a data protection officer. 33
17 Interest in knowing where and how personal information is used by organizations. 33
18 Participant responses indicating whether they agree or disagree with these statements. . . 34
19 Responses related to concerns of organizations and individuals handling of private data. 35
20 Responses regarding the input of false personal data when creating an account on an online platform. 36
21 Awareness of data brokers. 37
22 Participant responses indicating whether they agree or disagree with these statements. . . 37
23 Familiarity with IoT terms. 38
24 Distribution of papers, per year since 2010, on privacy in the Internet of Things, based on Google Scholar database results. 40
25 Distribution of papers on privacy in IoT, since 2020, by keywords searches on Google Scholar, BASE and CORE databases. 41
26 Distribution of literature per IoT domain. 45
27 IoT privacy contextual diagram. 86
28 Data flow diagram. 87
29 Swimlane diagram. 88
30 Use cases diagram. 91

List of Tables

1 Structure of questionnaire . . . 20
2 General information of papers . . . 44
4 Requirements table . . . 90
5 Use case 1 - device information query . . . 92
6 Use case 2 - statistics query . . . 93
7 Use case 3 - add a device . . . 94
8 Use case 4 - edit a device's data . . . 94
9 Use case 5 - delete a device . . . 95
10 Use case 6 - create an account . . . 96
11 Prioritisation table using the MoSCoW technique . . . 97
12 Features prioritisation table . . . 97
13 Highest priority requirements ordered . . . 97

List of Acronyms

AI Artificial Intelligence

BASE Bielefeld Academic Search Engine

CCPA California consumer privacy act

CCTV Closed-circuit television

DCM Deep computation model

DP Differential privacy

DPDCM Double-projection deep computation model

EU European Union

GDPR General data protection regulation

ICT Information and communications technology

IP Internet Protocol

ISO International Organization for Standardization

IT Information technology

IoT Internet of Things

LDP Local differential privacy

M2M Machine-to-machine

MIT Massachusetts Institute of Technology

NIST National Institute of Standards and Technology

PA Privacy assistant

PPA Personalized privacy assistant

PPDPDCM Privacy-preserving double-projection deep computation model

PbD Privacy-by-Design

RFID Radio-frequency identification

RIA Regulatory impact assessment

RQ Research question

SLR Systematic literature review

SaaS Software as a service

UK United Kingdom

USA United States of America

VM Virtual Machine

VPN Virtual private network

1 Introduction

Privacy as we know it is a somewhat recent concept [1, 2], before the digital age there was barely any notion of privacy for most people. For many centuries most people used to reside in small communities where they were continuously involved in one another's lives. Even more recent is the concept that privacy is a crucial component of personal security, in contrast to the necessity of public security. Privacy has traditionally been considered a luxury and is still frequently recognized as nicety as opposed to an essential need, even though it is acknowledged as a human right, as present in article 12 of the Universal Declaration of Human Rights [3]: "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks". The right to privacy is recognized in more than 120 national constitutions [4]. Privacy can be defined [5, 6] as the right to govern how personal information and data is collected, stored, and used, it frequently involves handling sensitive information with care, and as such, organizations must be open and honest about the kind of data they plan to gather, why they need it, and where and with whom they plan to share it. Users should have the right to control their shared information.

This definition can cause some confusion with the idea of security [7] and although privacy and security are interconnected, security involves measures taken to safeguard data from risk, threat or danger, it frequently alludes to safety. It is the practice of keeping users' personal information and data safe and preventing unauthorized access to it. The primary contrast between privacy and security is that the former deals with personal information to individuals and how they want their data used and maintained, whilst the latter deals with its protection from possible threats. Security can exist without privacy, but the opposite is not true. For managing sensitive and personal data, privacy and computer security are equally crucial. Users should be aware of the internal procedures regarding the collection, processing, retention, and sharing of personal information.

Concerns about digital privacy have been growing [8–10] in the last few years, especially after the Anonymous decentralized hacker group cyber attacks, WikiLeaks and Snowden's leaked top secret documents from United State's National Security Agency. These concerns can be noted with the increase of written literature on the subject, when searching for terms like "privacy", "online privacy", "digital privacy" in Google Scholar, ACM Digital Library or Science Direct it can be seen that, in the last 5 years, it returns about 5 000 000, 650 000 and 80 000 documents respectively, including articles, books, conference papers etc.

Privacy has become such an important concern in the digital age in which we live because most technology organizations heavily rely on the advertising sector [11], which in turn depends on customer data. Following the 2000s, many organizations offered their services to the general public without charging the user, however, because these organizations depend on revenue to stay afloat, the user is ultimately treated as a commodity for their data because they are not required to pay to access the service.

Internet of Things is a term that first appeared in the 1990s, and it may be linked to Mark Weiser's paper on ubiquitous computing [12] and the growth of devices of all sizes that communicate with one another to do various tasks, making Weiser's dream a reality. The first use of the term *Internet of Things* was in 1999 by British technology pioneer Kevin Ashton [13], executive director of the Auto-ID Center at Massachusetts Institute of Technology (MIT), to describe a system in which items may be connected to the internet by sensors. He came up with the phrase while giving a presentation for Procter & Gamble to highlight the value of linking Radio-Frequency Identification (RFID) tags used in corporate supply chains to the internet in order to count and track goods without the need for human assistance. These devices are used in various applications, starting at home [14] with thermostats, fridges, microwaves, etc, moving on to smart cars [15], the educational system [16], our clothes and our watches [17] and even into outer space [18]. IoT resources may include IoT equipment (like smart home assistants and autonomous vehicles), IoT services (like video analytics services linked to smart cameras and indoor position tracking systems), or IoT apps (like smart TV remote apps) that track and use information about us. Internet of Things is now widely used to describe situations in which a range of objects, gadgets, sensors, and ordinary items are connected to the internet and have computational capabilities.

The Internet of Things can be defined as: “An open and comprehensive network of intelligent objects that have the capacity to auto-organize, share information, data and resources, reacting and acting in face of situations and changes in the environment” [19].

The idea of using computers and networks in order to monitor and manage devices is nothing new, despite the term *Internet of Things* being relatively recent. Wireless technology improvements in the 1990s permitted the widespread adoption of corporate and industrial machine-to-machine (M2M) solutions for equipment monitoring and operation. Many early M2M solutions, on the other hand, relied on proprietary purpose-built networks or industry-specific standards rather than internet standards. To connect devices other than computers to the internet is not a new concept. A Coke machine at Carnegie Mellon University’s Computer Science Department [20] was the first ubiquitous device to be linked to the internet. The system, which was created in 1982, remotely observed the out of stock lights on the pressing buttons of the vending machine and broadcast the state of each row of the vending machine on the network so that it could be accessed using the Name/Finger protocol through a terminal. In 1990, a toaster that could be turned on and off over the internet that was created by John Romkey [21], was demonstrated at the Interop Internet Networking show.

IoT is one of the fastest growing technologies [22], it is predicted that it will grow into the trillions of devices by 2030 [23], and with this expansion new security vulnerabilities and data gathering threats appear, making these devices an ideal target for privacy violations and inadequate customer disclosure of device capabilities and data practices aggravates privacy and security issues.

When the earliest computers were created, privacy was not even considered a concern because they were utilized for basic calculations, it was only in the decades that followed [24], as computers became connected to one another that privacy gradually came to the forefront. In 1973, the US Department of Health, Education, and Welfare published *Records, Computers and the Rights of Citizens, Report of the Secretary’s Advisory Committee on Automated Personal Data Systems* [25], one of the first documents on digital privacy and an important first step that would form the basis for modern privacy legislation. In 1977 a revision of privacy policies would be published by the Privacy Protection Study Commission [26]. Meanwhile, in the 1980s, computers were becoming more ubiquitous in workplaces and increasing popularity in people’s homes, sparking debate regarding digital privacy. With the introduction of the World Wide Web at the start of the following decade, privacy concerns began to increase.

Privacy in IoT systems in not seen as a crucial factor in the design and development stages [27], instead emphasis is placed on enhancing system quality, providing management controls, and maximising productivity. Specific standards for privacy options have been imposed by data privacy regulations including the General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA), but even these regulations have been criticized [28–32].

Privacy is considered to be one of the most important concerns in IoT by individuals, according to a study conducted by Mozilla [33] with nearly 190,000 participants, some papers [34, 35] also consider privacy a primary issue alongside security while others also highlight energy efficiency [36].

The primary motivation for writing this dissertation was to investigate the dissonance between individuals high regard for their privacy and their behaviours and to determine whether this happens because individuals are lacking the knowledge needed to make better and informed decisions. The focus on IoT was chosen since there isn't as much research being done to explore privacy issues as there is for the web. As IoT devices are becoming more prevalent, new methods of communicating, gathering, and analysing data emerge. It is much more fruitful to investigate the issue of privacy in the setting of the IoT because there is already a sizeable body of research focusing on web or mobile privacy as opposed to IoT privacy. This dissertation will mainly focus on privacy concerns in IoT systems in a holistic manner in order to present fresh perspectives and serve as a repository for the collected understanding of these issues, challenges and addressed gaps.

This work contributes to the overall body of research by compiling and reviewing other works with the perspective of privacy as a distinct subject matter rather than an extension of security, as many publications imply. By conducting a survey on the perception of individuals on privacy in IoT systems from the majority viewpoint of Portugal, as 60% of participants were portuguese. Additionally, a mobile application was developed that is capable of doing its purpose on its own without having to rely on additional platforms.

The approach for this dissertation is formulated in two phases, the first of which is a systematic literature review (SLR), and the second is a questionnaire, and a mobile application. These stages are discussed in the several chapters that make up the dissertation's structure: Chapter 2 is composed of a systematic literature review that focuses on gathering the state of the art in terms of the most relevant approaches to privacy issues in IoT, as well challenges inherit in IoT and other relevant topics that help give an overview of IoT privacy. Finally, following a review of each work, the main lessons learned are outlined, the most significant gaps in the literature are highlighted, and suggestions for future research that could be undertaken to tackle IoT privacy are provided. This chapter aims to address these two research questions:

RQ1: What approaches are currently being considered to address privacy issues in IoT?

RQ2: What issues are prevalent in IoT that make it challenging to protect individuals privacy?

Chapter 3 describes the methodology of this work, it is composed of a questionnaire and a mobile application. The questionnaire aimed to gather the general privacy concerns of individuals, their online habits, their understanding of privacy concepts and their relation with IoT devices. The mobile application was designed and developed based on the findings of the SLR and the questionnaire, it seeks to help users broaden their knowledge of IoT and privacy with the goal to enable them to make more informed decisions. Two research questions will be considered in this chapter:

RQ3: What are the perceptions of individuals on online privacy?

RQ4: How can users be empowered to protect their privacy in IoT systems?

The results of the questionnaire and the usability tests of the application are covered in Chapter 4, while the discussion of these results is reserved for Chapter 5, which will also address the research questions posed for phase 2 of this work. Chapter 6 discusses the challenges confronted throughout the making of this work, from the SLR to the development of the application. Future work is discussed on Chapter 7. Finally, Chapter 8 presents the conclusion.

2 State of the Art

This section provides an overview of the recent literature with the themes that were found to be more relevant for this work.

2.1 Privacy Paradox

The use of a variety of digital devices have numerous advantages, but they also bring with them the ubiquity of data capturing equipment, therefore, it is understandable why the majority of online users have serious concerns about the privacy of their personal data. However, the opinions expressed are starkly at odds with the reality, according to Thomson et al. [37] report on the state of privacy, that just one in four European users read the terms and conditions in their entirety prior to making an online purchase or subscribing to a service, 59% admitted to only quickly scanning the terms and conditions before completing a purchase, while 14% admitted to never reading them at all, 30% of the respondents would even swap their email address to win a reward, or entry into a raffle, while 17% would do so to get an app and 30% would do it for money.

This is what is called a privacy paradox, there have been multiple papers written on this subject [38–42], some papers attempt a theoretical explanation while others attempt an empirical one. There has been very different interpretations or explanations of this paradox, a few papers [43–45] apply the theoretical concept of the *homo economicus* [46], which is the representation of people as beings who constantly act in a way that is logical and self-interested, not worrying about morality or ethics, and who do so to the best of their ability, to the context of privacy. Individuals' decision-making processes can be greatly influenced by a variety of cognitive biases and heuristics, and these decisions can vary greatly from individual to individual, according to several studies on individual choice behaviour [47–49]. According to several articles [50, 51], this paradox might be explained by the fact that some people have genuinely experienced online privacy assaults and that most privacy views are therefore based on heuristics or second-hand accounts. Taddicken's study [52] argues that peer pressure is the reason people have this contradictory behaviour, Norberg et al. [53] explains this paradox by suggesting that while perceived risk affects reported attitudes and behavioural intentions, trust has a direct impact on privacy behaviour, while others [49, 54] rely on quantum theory, meaning there is indeterminacy in the decision making, in other words, it means that an individual's decision's result can be determined at the time of the decision and not before. Brandimarte et al. [55] have explored the idea that when it comes to their data privacy, users have an *illusion of control*.

This paradox has been proven to be vitiated by a number of empirical studies [50, 56–58], online privacy practices are founded on separate privacy mindsets and so they are not inherently paradoxical.

Another concept worth analysing is differential privacy (DP) which relates more closely to the survey that will be conducted but also to the general collection and analysis of user data by applications and systems.

2.2 Differential Privacy

The notion of differential privacy, according to Michael Kearns [59], is based on three important principles. The first being that "differential privacy requires that adding or removing the data record of a single individual not change the probability of any outcome by much". The second principle being that "no outside observer can learn very much about any individual because of that person's specific data". And the third important principle being that "for every individual in the dataset, and for any observer no matter what their initial beliefs about the world were, after observing the output of a differentially private computation, their posterior belief about anything is close to what it would have been had they observed the output of the same computation run without the individual's data".

DP has the potential to significantly increase individual privacy protection, by purposefully adding noise into a dataset, it gives plausible deniability to any individual who may have had their data exploited while still being able to calculate statistics with relatively high precision. Although algorithms that deal with notions of fairness, ethics, and privacy are hard to implement because of the subjectivity of these concepts, and DP algorithms are no exception, they can still help in regards to tackling technology's intrinsic ethical and moral issues.

Zhao and Chen [60] conduct a SLR on DP for unstructured data. The authors present DP methods for sensitive content in image, audio, video, and text data. They compare the various methods and perform utility analyses for each method, highlighting the benefits and drawbacks of each, the utility loss is measured in experimental evaluations between the actual data and its obfuscated variant. They come to the conclusion that DP as well as its variations give stringent privacy protections for unstructured data against attackers with unpredictable background knowledge. They also suggest potential future study subjects that have yet to be investigated.

Integrating federated learning and local differential privacy (LDP) is proposed by Zhao et al. [61] to facilitate crowdsourcing applications to achieve the machine learning model while avoiding privacy threats and reducing communication costs. The authors provide four mechanisms depending on the privacy budget that is wanted. Based on experimental results of real-world datasets, the suggested mechanisms provide stronger privacy protection when compared to other comparable mechanisms; nevertheless, further testing is required to establish the validity of these mechanisms on production systems. No future research subjects are provided, only the intention of applying these mechanisms to deep neural networks.

Similar to differential privacy, there exists different algorithms that aim to preserve privacy such as Google's box blurring algorithm [62] that is used in the Google Map's street view, Microsoft's Visor [63] which is a video-analytics-as-a-service tool and Shokri and Shmatikov's [64] system for collaborative deep learning, however, in general, these algorithms struggle with high computational cost, internal attacks, or non-provable privacy.

2.3 Privacy in IoT: Approaches

There have been a number of SLRs [65–70] and systematic mapping reviews [71, 72] done to study privacy and security issues in IoT.

In Gupta and Ghanavati's [65] SLR, the authors review papers with methodologies and techniques that identify privacy risks or notify users about these risks. They divide the literature into the semantic-based, data-driven, source code analysis, survey, blockchain and architectural and framework based approaches. The authors' findings show that: most works concentrate on single IoT devices when addressing privacy threats; When analysing privacy issues, key privacy factors such as data reduction and data aggregation are overlooked; existing studies ignored the sensitivity of the obtained information; most useful studies did not include a diverse range of users when assessing privacy problems; no work has been done to discover compliance difficulties between an IoT application and different privacy rules; and current research does not place a premium on providing individuals with real-time privacy notices. Overall the authors do an extensive literature review and provide threats to the study's validity, however, this SLR has the following limitations: the authors only chose papers and not dissertations, thesis or books and from the selected papers, only the ones written in english were considered.

Kühtreiber et al. [66] evaluate the frameworks and tools established for developers, specifically in the case of IoT, and find that current solutions are difficult to use, only successful in limited scenarios, and insufficient to handle the privacy problems inherent in IoT development. This study lacks a comprehensive gap review of the chosen literature, along with research questions (RQs) establishing the significance of the articles chosen.

Sicari et al. [67] examine current research and ongoing activities that focus on IoT privacy and security solutions. The authors start by describing the requirements for IoT privacy and security, such as access control, confidentiality, and authentication. The authors then conduct a literature study in connection to these three needs. The authors came to the conclusion that IoT privacy issues have only been partially examined and that further attention is required. The study does, however, have some shortcomings: the prior research analysis focuses primarily on security requirements while ignoring privacy considerations; the authors do not conduct a thorough gap analysis on the publications examined; and they do not provide a comprehensive summary of future research topics in the area of IoT privacy that require more attention.

Lin et al. [68] undertake a literature review to identify security and privacy vulnerabilities in the three IoT architecture layers: network, perception, and application. The authors describe the first six fundamental security properties for these tiers as confidentiality, integrity, availability, identification and authentication, privacy, and trust. Then, the authors look at a variety of security threats for each of the three stages. The authors wrap up by giving a succinct summary of many privacy-preserving data techniques, including the stages of data collection, data aggregation, and data analysis. The authors do, however, largely focus on the IoT's security components and consider privacy to be one of the most crucial security aspects, rather than viewing privacy as a distinct concern. Despite this, the authors devote a portion of the paper to privacy and the majority of what they address concerns security.

Yang's study [69] examines the literature from the perspective of IoT phases, such as data collection, transmission, and storage, or in other words, perception, network and application layers. In these phases the author explores security protocols at the physical layer, network solutions, and data storage and sharing approaches. There is a balance of security approaches or solutions and non-security related approaches discussed, besides the mentioned security protocols, the author proposes a software-defined network, for the network layer, that enables a global view of the network which helps to identify malicious traffic patterns. The author also suggest the use of differential privacy, privacy-preserving data mining, blockchain or holochain for the application layer. The author presents a high-level overview of privacy issues and as a result does not offer a comprehensive analysis of the various approaches or make many suggestions for future study directions.

In Zubaydi et al. [70], different works that address the integration of blockchain technology that aim to tackle various aspects of privacy are analysed. This SLR has an overview of the IoT architecture, the various blockchain types and algorithms and IoT challenges. The authors provide an overview of the chosen papers and present a detailed comparison between them before analysing the papers using several categories which include generic approaches, healthcare, smart environments, IoT device gateway, IoT information systems, management systems and other approaches. The authors also present a considerable number of open issues that future works should focus on. Because this SLR is about blockchain technology, the focus is on privacy through security. Overall this is a very in-depth SLR on the use of blockchain technology to address privacy issues.

The Khanna and Kaur [73] research paper provides a comprehensive literature review on the IoT. 194 publications have been examined and categorized by a particular domain, including independent living, smart grid, smart house, environment monitoring, healthcare, industrial processing, and smart agriculture. Following the assessment of the literature, the authors propose the following unresolved issues: availability, reliability, mobility, performance, data confidentiality, network management, scalability, interoperability, security, and privacy. They argue that because IoT is evolving quickly, current research will be unable to keep up, the data produced by various sensors is extremely important and needs to be managed and evaluated with extreme care, and extensive research in this field will make a significant advancement with the goal of requiring no human intervention.

The research conducted by Tzafestas [74] focuses on an overview of broad ethical questions, ideas, and theories that can be applied in the IoT, IoT security, privacy, and trust elements, and the role of governments. Before discussing the role of governments, a distinction between ethics and law is made. Governments play an important role in mediating the future of IoT, especially since many plan to develop smart cities. Governments should also ensure that IoT products and solutions are used exclusively for their specified goal. Privacy protection, security, usability, user experience, trust, and safety have been identified as the primary ethical problems in IoT. The author concludes that in order to address sophisticated and tricky unethical and illegal behaviour across the IoT, more specific IoT legislation and ethics norms for IoT need to be devised and constantly updated.

Based on Ziegeldorf et al. [75] analysis of the literature, the following are the most prominent privacy concerns in IoT:

1. The most prominent concern is *identification*, which binds an identifier, such as a name and location, with an individual's identity, this also enables and aggravates other threats;
2. *Localization and tracking* is the threat of detecting an individual's locations through numerous techniques, such as GPS, internet traffic, or smartphone location. This threat requires *identification* of some kind;
3. In e-commerce, *profiling* is often used for personalization. Organizations collect information about individuals in order to deduce their interests via association with other profiles and data sources.
4. *Interaction and presentation* allude to the sharing of private information with an unintended audience while doing so through a public medium. IoT applications often need extensive user interaction, it is expected that users of these systems will obtain information via smart devices in their immediate surroundings and that users will interface with systems in creative, natural

ways. However, many of those modes of communication and presentation are already available to the broader public, making them apparent to anybody around. When personal information is transferred between a system and its user, privacy is breached.

5. *Lifecycle transitions* occur when an IoT device is sold, utilized by its owner and eventually disposed of. There may be an expectation that the object deletes all information, yet smart devices frequently keep massive volumes of data about their own past throughout their entire existence. This might contain personal images and videos, which are not always erased following ownership transfer.

6. *Inventory attacks* involve unauthorized entry and the acquisition of information about the presence and characteristics of personal things. Malicious users might use inventory data to profile the property and break in.

7. *Linkage* is the process of connecting disparate systems, when systems are connecting different data sources, there is a higher danger of unauthorized access and data leakage.

2.4 Proposed Solutions

This section presents a variety of solutions separated by themes that arose from the structured literature research in order to address the disconnect between privacy ideas across systems and users.

2.4.1 User Awareness

Although individuals have a certain expectation of privacy when using IoT devices, privacy preferences are complex, as noted by [76], some individuals value some aspects more than others. There are multiple studies on privacy in IoT but not many focus on user privacy awareness or privacy literacy.

Koohang et al. [77] propose a research model with the goal to examine the relationship between IoT awareness and IoT privacy, security, and trust, and how this influences IoT continuing usage. To validate the model, they conducted a study with 299 participants and discovered that as users increase their awareness of IoT security and privacy threats, their privacy and privacy knowledge increases, and as users become savvier about IoT privacy and security, they place more trust in IoT devices, which affects their continued intention to use these IoT devices. Previous publications provide support for the proposed model like [78, 79]. The authors acknowledged the method's shortcomings due to use of a traditional statement-based method survey in the study and proposed a scenario-based method approach and a random sample to validate the results, but they did not elaborate any further on future research directions.

An interactive theatre experience was developed by Skirpan et al. [80] as a case study of an innovative approach to gather user awareness about their online behaviour with regard to privacy. This was created to try to prove that a simulated experience with a credible privacy problem may encourage people to take action before actually encountering a catastrophe. The plot of the play consist in a fledgling tech company that unveiled its revolutionary artificial intelligence (AI) technology while dealing with a company whistleblower and an untimely zero-day hack on their system. The public is able to interact with the actors and influence how the story plays out. Audiences and actors were given the chance to try on roles, behaviors, and opinions that they would not normally have access to in ordinary life. The authors had interviews and surveys done after the plays with audience members however they only did interviews halfway through production and only a small fraction of the audience actually participated in this data collection, they also noted that after contacting people months after the interviews that they did not really changed their behaviour regarding their privacy rights.

2.4.2 Legislation

Some papers seek to improve legislation [81, 82] because otherwise, in their view, privacy rights won't be respected if they are not enforceable legally, they defend that without the express agreement of the individual concerned, private information obtained by IoT devices must not be retained or processed in any form, and necessary procedures must be taken to guarantee that the data collected is not that of an unrelated individual. Before regulations like the GDPR, CCPA or the Artificial Intelligence Act, some papers such as Weber [83] and Ziegeldorf et al. [75] urged for better regulations to protect personal privacy in IoT systems. But better protection legislation for individuals would also create opposition from most companies that want to extract as much private data from their users without (m)any restrictions in order to increase their profit margins.

The research conducted by Hadzovic et al. [84] focuses on present initiatives aimed at establishing an IoT and AI regulatory and legislative framework in the European Union (EU), as well as its pertinence in developing nations, the authors choice to focus on EU is due to EU's claim for becoming a global leader [85]. The authors identify three steps toward the development of an IoT and AI regulatory and legislative framework. The first step is to develop a national AI strategy, which is an important blueprint for increasing AI adoption and should be developed in accordance with a country's strategic priorities. To that end, a proactive information and communications technology (ICT) regulator could initiate and encourage national AI strategy development. The second step in developing a regulatory framework for IoT and AI involves the consideration of various aspects of IoT and AI and navigating through many overlapping policy areas to determine the rules. The legislation must be designed to be future-proof and not restrict technological development. The ICT regulator, which plays a central role in improving innovation and developing the electronic communications market, could also play a central role in the context of IoT and AI. It could serve as a coordinating authority within an advisory committee before the establishment of a national supervisory authority for AI. The new regulatory framework should be prepared and assessed using the regulatory impact assessment (RIA) method to select the best option for the country. And the third step in the development of a regulatory framework for IoT and AI involves redefining the role of regulatory authorities. The state must establish a national supervisory authority, but it is also important to involve civil society, the private sector, and academia in the process to ensure success. This approach is known as multistakeholder governance development.

2.4.3 Privacy through Security

Sun et al. [86] design a lightweight communication strategy for a remote-control system, employing two types of Virtual-Spaces to achieve the aim of identity announcement and data exchange. They constructed a prototype system of the scheme and tested it on the Freenet, demonstrating that the method can effectively resist the influence of flow analysis on communication anonymity while preserving communication data security.

Motivated by privacy concerns in IoT, Xiong et al. [87] propose a locally differentially private packet obfuscation mechanism as a defence against packet-size side-channel assaults in IoT networks. Because a quantifiable measure of privacy risk was required for preserving privacy, the authors chose LDP in order to protect individual IoT devices given that it could mask packet sizes prior to transmission. The effectiveness of hiding the packet size from various smart home IoT devices was empirically demonstrated, and it was also clear that this mechanism works really well when using high bandwidth, but many IoT devices purposefully use low bandwidth for a variety of reasons, so bandwidth-constrained users must optimally tune their privacy preferences and trade off privacy with bandwidth. For further research the authors suggested addressing timing side-channel attacks.

2.4.4 Architecture / Frameworks

Antunes et al. [88] do a SLR on federated learning in the area of healthcare and make an architecture proposal. The procedure known as federated learning allows for the distributed training of machine learning models using remotely hosted datasets without the requirement to gather and hence jeopardize data. The fundamental goal of the proposed architecture is to allow healthcare institutions that have access to sensitive medical information to use it in distributed data analysis and machine learning research while ensuring patient confidentiality. Because information transmitted among institutions need confidentiality guarantees for learning model parameters and analysis results, the architecture can adopt a number of ways based on a zero-trust security paradigm [89]. Furthermore, the institutions develop a learning algorithm verification system that can store and disseminate manifestos, as well as engage in distributed analytic procedures that need unanimous agreement from all participants. This study also demonstrates what previous literature implies, that homomorphic encryption and differential privacy are effective approaches for preventing data breaches without incurring prohibitively high computing costs.

Opara et al. [90] present a system for spotting possible problems with privacy or security regulations in the early stages of development, this approach is intended at developers. The paper proposes a domain-specific ontology for modelling IoT security and privacy policies, a notation for representing and validating IoT security and privacy policies, a set of guidelines and rules for detecting IoT policy errors, and a tool for visually modelling and capturing IoT security policies and discovering policy problems. Although the framework that is presented is theoretically promising it has not been tested in a real environment so the effectiveness can't yet be measured. The authors also do not compare their proposal with others already available.

A Privacy by Design (PbD) framework is proposed by Perera et al. [91] to assist software developers in incorporating concerns about data privacy into the design of IoT applications. The proposed framework consists in a set of guidelines. The authors conducted two studies, one interview based in person and the other online, the first study was done with software engineers with various levels of experience and the second was done with master's students. The authors note that developers do not have a privacy mindset, meaning they do not regard privacy as an primary aspect of the application design, giving more importance to other aspects like functionality or security. They also mentioned that using this framework must be context-based, since it may be difficult and prone to over-analysis by engineers. When citing related works, they state that various PbD frameworks exist but none are dedicated to IoT, which is incorrect as evidenced by other works [92, 93], with [94, 95] being more recent examples. It is true, however, that there are relatively few IoT PbD frameworks.

Q. Zhang et al. [96] propose a double-projection deep computation model (DPDCM) and a privacy-preserving (PPDPDCM) variant for big data feature learning. The DPDCM extends the deep computation model (DCM) by replacing all hidden layers with double-projection layers. Because the learning algorithm is very time-consuming the authors choose to leverage cloud computing to increase efficiency by crowdsourcing data on the cloud, but in big data sets there are a large number of private data so to solve this problem the authors propose a PPDPDCM by using the BGV scheme to encrypt the private data. After testing the proposed models they conclude that the DPDCM is marginally better than the conventional DCM with a 3% to 4% improvement, but the PPDPDCM performs similar to DPDCM without a noticeable classification accuracy drop while still preserving individual's private data.

There exists frameworks like the NIST Cybersecurity Framework [97], published by the National Institute of Standards and Technology (NIST), which includes a number of recommendations for reducing organizational cybersecurity risks. Although this framework is primarily focused on security, it can also be used to reduce privacy threats. ISO/IEC 27400:2022 [98], developed by the International Organization for Standardization (ISO), is a standard that provides guidelines on risks, principles and safeguards for the security and privacy of IoT systems. In addition, ISO has produced about 208 privacy-related standards [99], some of which have been published and others of which are currently in development. This is a small number when compared to the roughly 1305 security-related standards [100], but there may be some overlap between these standards. Even those ISO privacy based standards have a good amount of intersection with security related guidelines.

2.4.5 Blockchain

Blockchain is an option to guarantee privacy in IoT because of zero-knowledge proofs [101], ring signatures [102] and mixing [103] among other techniques [104].

A zero-knowledge proof is a cryptographic technique that enables one party to demonstrate to another that a certain statement is true without disclosing any information other than the validity of that statement. Completeness, soundness, and zero-knowledge are the three requirements that must be satisfied by a zero-knowledge proof method. Completeness states that if a statement is genuine, an honest party will be convinced of it by another honest party. Soundness indicates that a nefarious party should only have a small chance of convincing an honest party that a statement is true. Zero-knowledge states that the method must only tell one party whether or not the other party is disclosing the truth.

Ring signatures create a single, recognizable signature that is used to sign a transaction by combining a number of partial digital signatures from diverse users. This group, known as the ring, can be chosen at random from the outputs that other users have made to the blockchain. A ring signature has the security property that it should be computationally expensive to determine which of the group's members' keys was used to produce the signature, this is because it obfuscates the input side of a transaction. A user's anonymity cannot be taken away from their signature, and any group of users can act as a signing group automatically.

Mixing is the process of blending possibly traceable digital assets with others to obscure the original assets' sources. This is frequently done by pooling source assets from different inputs for a long period and at random intervals, then spitting them back out to destination addresses. Since they are all packed together and then delivered at random intervals, it is very difficult to pinpoint particular assets. Due to the fact that cryptocurrencies provide a public record of every transaction, mixers have been developed to improve cryptocurrency privacy. Because of their emphasis on secrecy, mixers have been used to launder money using cryptocurrency.

R. Zhang et al. [104] outline some disadvantages concerning these techniques. For example, the authors assert that zero-knowledge proofs are less efficient than other techniques, that ring signatures cannot reveal the identity of the signer in the event of a dispute, and that centralized services run an increased risk of leaking users' personal information when it comes to mixing.

Yu et al. [105] discuss various implementations of blockchain that provide privacy through security, based on different categories like data integrity, data sharing and authentication and access control, most implementations proposed use a peer-to-peer network based on the Ethereum platform, with the exception being using any consortium or private blockchain, consortium meaning a blockchain where consensus is managed by a pre-selected set of nodes and private meaning access permission and read/write authority are tightly controlled. The authors use privacy as a proxy for security, they also do not discuss the weak and strong points of each implementation or make any comparison, they also do not provide further research questions.

Ali et al. [106] suggest a software stack that combines peer-to-peer file sharing with blockchain smart contracts to offer IoT users control over their data and do away with the necessity for centralized IoT data management. Blockchain smart contracts are used in the proposed 'modular consortium' architecture to regulate access while establishing responsibility for both data owners and other parties that users grant access to.

2.4.6 Privacy Assistants

There exists a number of privacy assistants in the market. Privacy assistants have the objective of giving the user flexibility in choosing the preferred privacy options in available applications, most are used in smartphones, very few are made for devices in the IoT.

The Carnegie Mellon University CyLab, which is the university's security and privacy research institute, started developing in 2019 an IoT Infrastructure that intended to be free of privacy leaks and software covered by their Secure and Private IoT Initiative 2019, this project would fall under their main research theme of Trust. In this project they started the design of a Personalized Privacy Assistant (PPA) [107], this would involve the use of semi-structured interviews with 17 participants to examine user perceptions of three hypothetical PPA implementations, each of which is potentially more autonomous, while outlining the advantages and disadvantages of each implementation. The interviews were divided into three sections: exploratory, anchoring and the PPA; While the exploratory phase's purpose was to learn about participants' attitudes and understanding of IoT, the anchoring phase aimed to normalize participants' basic understanding of how IoT functions. In order to get people to think about potential privacy concerns towards the end of the anchoring section, the authors asked participants about their opinions on data privacy. In the PPA section, it was proposed the idea of a PPA for IoT as a potential future project. The authors clarified that the PPA could distinguish between active data requests such as a gadget asking biometric information from the user's health tracker and passive data collection such as a smart device with a microphone that could record people's utterances while they were nearby. The Notification, Recommendation, and Auto implementations of an IoT PPA were the three that the authors and attendees discussed. Notification PPAs can determine which adjacent devices are requesting data and alert users to those devices' presence and requests so that users can approve or reject each request. Building on notification PPAs, recommendation PPAs offer advice to individuals on how to share their data based on their preferences. The user's data sharing decisions would be made by auto PPAs. This would lessen the cognitive load on individuals but also take away their ability to influence the process. They found that the participants' attitudes regarding the various implementations were generally favourable, although they also voiced worries, which varied depending on the degree of automation. Given the divergent motivations of participants some desired increased control, while others wished to avoid being overtaken by notifications and the lack of agreement regarding the optimal PPA implementation.

After the design phase, the institute implemented a privacy assistant (PA) [108], the authors called it IoT Assistant. Because the predominant approach of "notice and choice" for data privacy protection, they decided the PA would also fall into this approach, but because many systems implement notice as a form of consent, without sometimes offering choices to the end user, they also wanted this work to provide a conceptual framework that views user-centred privacy choice as well as a taxonomy for practitioners to use when designing meaningful privacy choices for their systems. The authors define meaningful privacy choices as "the capabilities provided by digital systems for users to control different data practices over their personal data", They extend the notion of privacy choices with five facets: effectiveness (the opportunity to establish privacy preferences that precisely and completely match the data collection and use methods that a user is okay with), efficiency (the capacity to specify these options with the least amount of effort and time), user awareness (where significant privacy options should be prominently and clearly communicated to users), comprehensiveness (users should understand their options, how they affect the gathering and potential use of their data, as well as what conclusions might be drawn from this data and the potential repercussions of these conclusions) and neutrality (meaningful privacy decisions should not be subject to manipulation or bias). The IoT Assistant offers four privacy settings, giving end users a variety of alternatives to better suit their varied privacy preferences and as a result, privacy options are more effective in the IoT environment. The IoT Assistant acts as a centralized privacy choice platform by implementing various privacy options, allowing individuals to more effectively govern their data privacy in IoT. The three IoT system discovery modes that the IoT Assistant supports are QR codes, push notifications, and location-based map interfaces. These discovery tools are probably going to make users more aware of the installed IoT devices and the privacy options they have. Additionally, the united viewpoint of the integrated notification and option in the the IoT Assistant gives succinct yet thorough information regarding IoT data practices to help users better understand the implications of their privacy choices. Additionally, the authors work to implement the integrated notice and option in the IoT Assistant without bias or framing, attempting to offer individuals a neutral space to execute their privacy choices. Although the authors view the IoT Assistant as a significant step towards "meaningful privacy options" in IoT, this assistant still has many problems, such as the fact that it is still in its early stages of development and that there hasn't been much growth given that it was created in 2020 and we are in 2023. Maybe the main reason this application was not able to be developed further is that the application itself serves to show the user the data that is already in the IoT infrastructure that was created before, and as such it is not capable of identifying new IoT devices without the end users themselves create on the infrastructure's main webpage [109] a new entry for the device in question that the user wants to interact with. Another reason that cripples this application as well as others that seek to provide better privacy in IoT systems is that many systems do not offer any type of privacy choices to the end user or to other users that are not the intended end users but the devices are still collecting data about.

The IoT infrastructure that was developed [109] is built on an open, distributed design that allows for the deployment and management of IoT resources to be carried out by any number of actors. Part of this infrastructure is the Internet of Things Resource Registry, it is a web platform that enables resource owners to declare not only the place where a resource is deployed but also data practices like the reason(s) for a particular data collecting process, the level of detail in the data being gathered, retention, the recipients of the data, and more. Additionally, it discloses any user-configurable privacy settings that might be connected to a particular resource.

2.4.7 Other Proposals

Zhu et al. [110] present a hybrid sensor system that safeguards privacy while also monitoring parking availability. The authors merged IoT sensing with crowdsensing and enhanced it with privacy-preserving methods. The authors employed physical hazy filters to mask IoT sensors and a cryptographic technique based on cryptographic commitments, zero-knowledge proofs, and anonymous credentials in crowdsensing. In addition, they used crowdsourcing to create a machine learning model for parking recognition in the presence of foggy filters. Their paper included proof-of-concept prototypes such as a Raspberry Pi system and a mobile app, as well as an evaluation study of the machine learning model and the effects of crowdsourcing.

Lola et al. [111] propose a system that manages IoT device network communication by having manufacturers declare their device's data collection intentions while simultaneously allowing IoT users control over their data privacy and security, and thus providing transparency to IoT systems. The system's design includes tools for analysing packets sent by IoT devices and executing network traffic control rules. The goal is to enable the declaration and verification of IoT device communication intentions, as well as to govern such communication in order to detect potential security and privacy violations. This system's limitations include the fact that it can only handle non-encrypted network traffic, only working in TCP/IP networks, and end-users' inability to adequately establish their user policies due to lack of technical knowledge. The authors suggest using machine learning to improve this system, homomorphic encryption and/or federated learning could also be used.

IoT sniffers are usually used to detect problems in the networks, they rarely are used to provide privacy for the users.

The LTEye project [112] is an open platform that provides granular temporal and spatial analytics on the performance of LTE radios without access to private user data or provider assistance. Despite the presence of multipath, LTEye uses a revolutionary extension of synthetic aperture radar to communication signals in order to precisely pinpoint mobile users.

2.5 Privacy Challenges

IoT is a composed of a complex web of architectures, applications and technologies. In terms of architectures, it can be decomposed in three layers: the perception layer, the network layer and the application layer.

The perception layer, also known as the sensor layer, interacts with physical objects and components via smart devices (RFID, sensors, actuators, and so on). Its key objectives are to connect objects to the IoT network and to monitor, collect, and analyze status information about these things using deployed smart devices. This layer can often be unreliable, for instance with autonomous vehicles where they find it hard to read road signs or to predict if certain objects are inanimate or not, but this unreliability also brings privacy even though some of the data might be unusable. Noise can also be added in this layer to provide extra privacy.

In the network layer there are many competing networks like ZigBee, Z-Wave, Bluetooth Low Energy, LoRa, Wi-fi, etc., this layer is fragmented specially in regards to wireless networks and that makes it very difficult to create an IoT architecture that can use various networks and have the various devices communicate with each other, even though interoperability is seen as a very important factor in IoT. Some of these networks are open standard protocols while others are proprietary and use different protocols of communication, use different frequencies, different ranges and different data rates. When creating an IoT architecture the designers often think of how to solve specific problems and use what is best for the current needs, and the way that IoT is fragmented doesn't help in providing progress.

The application layer receives data from the network layer and uses it to execute essential services or operations. This layer, for example, can provide the storage service to backup incoming data into a database or the analysis service to analyze received data in order to predict the future state of physical devices. This layer encompasses a wide range of applications, each with its own set of requirements. A few examples are smart grids, smart transportation, and smart cities.

Several major challenges that need to be addressed have been identified by Qu et al. [113], including the absence of a theoretical foundation, the need to balance privacy and data utility, and the over-complexity of system isomerism. Isomerism is a concept borrowed from chemistry, which refers to molecules that have the same molecular formula, but distinct arrangements of atoms in space [114]. The design of IoT structures lacks mathematical foundations and is based on empirical methods, which hinders the development of IoT. Optimizing IoT performance based solely on human experience is difficult, as is implementing privacy protection mechanisms without theoretical guidance. Adversaries can take advantage of this to increase the success rates of their attacks. Trade-off optimization must be based on scientific theory and quantitative analysis, but this is complicated by the presence of multiple parties with dynamic characteristics and diverse requirements. The lack of a theoretical foundation leads to non-uniform quantitative measurements and introduces uncertainty into trade-off optimization. The large number of standards and protocols contributes to the over-complication of system isomerism, which hinders communication and system integration. Ensuring effective IoT applications without wasting resources leaves less for privacy protection, but lightweight privacy protection cannot meet all requirements. Adversaries can also exploit structural information to launch various attacks.

In the context of big data privacy, Ranjan et al. [115] point out several challenges, such as obtaining user consent, giving users absolute freedom to control their data, full transparency of the data life cycle, anonymity technology, security throughout the dataflow and stakeholder responsibility. It's challenging to create technologies that efficiently and effectively solicit consent from users, given that each user has limited time and technical expertise to participate in the process, to resolve this the authors suggest integrating the principles and methods from both human-computer interaction and cognitive sciences. Data owners should have complete control over their data, yet current solutions provide restricted user access. Users should be able to select hardware and software, choose the sort of data they share and grant access privileges to, as well as be allowed to revoke or modify previous consents. Service providers should not disable functionality or change membership fees, to encourage consent. Regarding transparency of the data life cycle, without clear user approval, service providers shall not use previously gathered data for any other purpose. Given that technology should offer users anonymity, a comprehensive architecture for anonymization is necessary to facilitate complete anonymity in the IoT. This anonymity needs to be guaranteed at various stages, such as data modelling, storage, routing, communication, analytics, and aggregation. It is the duty of all stakeholders to secure the infrastructure, the data collecting and transfer process, and the individuals who use the devices. Security upgrades should require as little user intervention as possible. This paper is one of the few that outlines stakeholder responsibility. The authors identify five major stakeholders: device manufacturers, IoT cloud services and platform providers, third-party application developers, government and regulatory bodies, and individual users and non-users. Device manufacturers must be responsible for including privacy-preserving techniques into their devices, informing users on the types of data collected by their devices, explain the data processing processes used, and specify when and how data will be extracted, they must also provide users with the option to disable hardware components and third-party developers with programming interfaces. IoT cloud services and platform providers must be responsible to use common standards so it allows users to have a choice of providers and to easily move data between them. Local software and hardware gateways, such as mobile phones, can encrypt, process, and filter data locally, minimizing the quantity of data transferred to the cloud and the risk of user privacy violations. Application developers are responsibility for making sure their programs are free of malware and for giving users accurate information. Users must expressly consent before using any features of the app, and they must have the choice of which features to enable as well as the flexibility to withdraw, grant, or change their consent at any time. Users should also have full access to the data collected by IoT devices. Government and regulatory bodies should be responsible for the enforcement of standardization and legal efforts, while also allowing interoperability among different IoT solutions, a fair marketplace, and competition. The authors suggest a governing body similar to the World Wide Web Consortium for the IoT to oversee standardization and certification processes and that the IoT certification model should be much broader because it might need to certify both hardware products and software services. The majority of IoT systems currently in use are primarily geared toward users, although some IoT system types can also have an impact on non-users. The authors suggest using notifications that are similar to closed-circuit television (CCTV) surveillance in public spaces, but because monitoring and actuation tasks are difficult, it may be required to use interactive and digital tools to educate non-users.

Except for the papers focusing on the privacy paradox, very few papers discussed stakeholder responsibility. The ethical side of privacy is also not often examined. Privacy literacy is, however, somewhat more widely studied, though primarily outside the context of IoT, similar to the studies on the privacy paradox.

3 Methodology

The overall work is comprised of two phases which will be described in the following paragraphs. The first phase, which is primarily described on Chapter 2, focuses on gathering the state of the art in terms of the most relevant topics, from which the main privacy concepts were selected to be explored in the first stage of phase two with the creation of a questionnaire with the aim to collect user perceptions of privacy. The second stage of phase 2 consisted in developing an application, partially based on the information generated by the survey, that can identify what sort of devices are around, what kind of data is gathered by these devices, present privacy options to the user when available, and what can be done to prevent undesirable data from being collected.

The first phase involved conducting a systematic literature review to gather the most relevant papers discussing methodologies and techniques for the protection of user privacy data with a focus on IoT systems. This SLR focused on papers from the last 13 years, from 2010 until 2023, since papers published prior to 2010 become out of date with the evolution of technology. Although certain aspects of privacy have remained constant throughout the years, this historical perspective on privacy is primarily discussed in the introduction to this work.

The SLR followed Keshav's three-pass approach [116] and PRISMA 2020 [117] guidelines when selecting the papers for review, the principles suggested by Kitchenham and Charters [118] were also taken into account in order to ensure the transparency and reliability of the review. On Keshav's approach, first the title would be read, then the abstract, the introduction and conclusion and briefly skim the rest of the paper and then decide if it was worth reading any further. If the document passed the inclusion criteria, which will be discussed further ahead, then the document would be read in its entirety while ignoring any tables, figures, images or graphs. If the paper failed to present any interesting idea, approach, or technique it would be discarded, but if not, it would be read carefully from the beginning again in order to fully understand what it presents.

It is challenging to curate the literature and gain an in-depth understanding of its different aspects due to the vast array of approaches that address the topic IoT privacy. As a result, specific databases were used because of the sheer quantity of papers contained within them makes research easier. The following were the primary databases of the SLR:

- Google Scholar;
- ScienceDirect;
- IEEE Explore Digital Library;
- ResearchGate;
- Elicit;
- BASE.

Other supplementary databases used, during the course of the research, include: CORE, AIS, ACM Digital Library, Semantic Scholar, Baidu Scholar, RefSeek and Science.gov. These databases were used in to search certain works that would have not been found otherwise.

The papers were collected by searching the databases with keywords, a broad spectrum of results were obtained using generic terms like. The primary search terms were used between quotation marks, e.g. "Internet of Things", so that results include all words in sequence, operators AND

and OR were also used as was a minus sign before a keyword to remove said keyword from the search results. Many search terms were utilized, however the most often used ones, based on their frequency, include: “Digital literacy”, “Differential privacy”, “Privacy paradox”, “Machine learning”, “Blockchain”, “User awareness”, “User knowledge”, “Blockchain”, “Privacy concerns”, “Privacy perceptions”, “Regulation”, “Framework”, “Security”, “Deep learning”, “Approach”. Most of these search terms also included the terms “Privacy” and/or “Internet of Things” or any variants like “IoT” or “IoT privacy”.

The SLR attempts to summarize and evaluate IoT privacy concerns, as well as ideas, techniques, or methodologies to overcome those challenges. The focal point in this phase was answering the following question: Does the paper present a new methodology or interesting angle to tackle users’ privacy concerns?

As referenced before, only papers published from 2010 until 2023 are considered, these works must also be published in journals, conference proceedings, dissertations, thesis or technical reports. Exclusion criteria include: presentations, editorials, abstracts or commentaries. Works can cover any area as long as they deal with privacy in the Internet of Things, if the paper does not cover IoT then at least it must cover privacy aspects that can be applied to IoT, as is the case of the privacy paradox.

From database searches, a total of 229 papers were found. Applying the inclusion and exclusion criteria to these papers bring the total number of papers down to 95, excluding 134 papers. After reading the full texts of the remaining 95 papers, 47 were excluded, making the number of total papers in the SRL be 48.

Having collected the major findings of the SLR, this work then aimed to conduct a throughout study split into two stages, which will compose phase 2 of this work.

The second phase was evaluated on two stages, the first one consisted on doing a questionnaire on people’s general privacy concerns, while using and interacting with IoT devices. The SLR helped on the creation of the questionnaire to assess general user’s knowledge on privacy concepts, their habits and concerns, their understanding of privacy rights, and what they do to safeguard those rights. The goal of this study was to both understand the privacy paradox and collect insights on how to address privacy issues in IoT devices.

3.1 Stage 1: User perceptions

This questionnaire aims to understand people’s perception of IoT and their privacy practices online. It also serves to better understand and demystify the privacy paradox and to help provide a solution to the privacy issue in IoT, which will be discussed on Section 3.2.

The questionnaire consisted of 86 questions divided into 7 sections to gauge users’ digital literacy, the first section being about general privacy questions, then about the predisposition to data sharing, to concerns with privacy then about daily digital routines, then about profile identification, subsequently about IoT general knowledge before a final part about non-identifiable demographic data. The questionnaire’s structure is shown in more detail in Table 1, the full questionnaire is presented on Appendix 8.

#	Section	Details
1	General knowledge and attitudes towards privacy	This section's goal is to ask generic questions regarding the participants awareness of information privacy.
2	Disposition for sharing personal information	This section is designed to elicit generic inquiries about the participants willingness to provide personal information.
3	Privacy concerns	This section strives to elicit questions about potential concerns about disclosing personal information.
4	Current online habits and practices	This section includes general questions with regard to working with the internet in everyday activities.
5	Profile identification	This section gathers more particular questions concerning employing profiles to make it more straightforward to generate tailor-made interactions.
6	Knowledge and habits regarding the Internet of Things	This section contains questions about participants' usage patterns for IoT devices as well as questions that aim to understand their level of literacy.
7	Demographic data	This section is for gathering broad demographic information that allows to characterize the participants in statistical terms.

Table 1: Structure of questionnaire

Great care was taken when it comes to this questionnaire's data collection, in order to not identify any individual or group of individuals, for instance, when it comes to differential privacy, any data that might identify someone will not be disclosed, even though the data might suffer from some inaccuracy because of this.

The scale that was used in the questionnaire, which is from one to seven, is based on the work of Philip K. Masur [119], it was chosen because it provides a more nuanced understanding of the knowledge of participants. This scale was developed as an online privacy concerns scale so it fits perfectly on this questionnaire, another scale, also developed by Masur with Teutsch and Trepte, that could be used but was ultimately not chosen is the Online Privacy Literacy Scale [120], although the questionnaire does contain some of the main aspects of this scale like knowledge of data collection and analysis practices by institutions and online service providers knowledge of data protection law, knowledge of technical aspects of data protection and knowledge of data protection strategies. This survey was partially based in a study done in the Philippines by the government in the context of their privacy act of 2012 [121], this was the second survey done on the country's population. It was also inspired by Alves's master's thesis [122], which was about citizen's perception about privacy in the wake of GDPR.

This survey was constructed in Google Forms and distributed through the internet, the intent would be that this would reach as many individuals as possible, besides Google Forms itself, it was used other online venues for distribution like social media, forum websites, through in person discussions and some participants responded when also doing the application usability tests.

The questionnaire was available for completion until August 30, 2023, and during the time that it was open 46 participants responded. Several online survey dissemination services were used to acquire participants, all the services used were based on the goodwill of the participants, there was no financial incentive for completing the survey. Most of the services used were software as a service (SaaS), these platforms are based on credits for filling in other questionnaires available, this makes the process of acquiring participants very tedious as many questionnaires need to be filled in to get a reasonable number of participants (at least 150 to 200 participants). Disseminating the questionnaire in this way does not entail any additional cost, but it may mean that the results obtained in this way may not be as honest as possible, as some participants may be filling in this questionnaire quickly just to get the number of participants for their own questionnaires, but there is also no way to guarantee that if this questionnaire was carried out with some financial incentive that participants would fill it in as honestly as possible. In addition to dissemination by the various services, social networks were also as well as it was personally disseminated to family and friends. One possible method of dissemination would be in person, house to house, but this would be a very slow way to get responses, not to mention that people might feel obligated to respond, which could be considered to be unethical, and the answers might have been answered in a less than honest manner.

From the respondents, 47% are male and 51% are female, while 2% don't identify as neither, see Figure 1. 40% of the participants are younger than 25 years old, 31.11% are aged between 26 and 35, 9% are between 36 and 45 years old, 18% are older than 46 but younger than 65 and 2% are have more than 65 years, as shown on Figure 2.

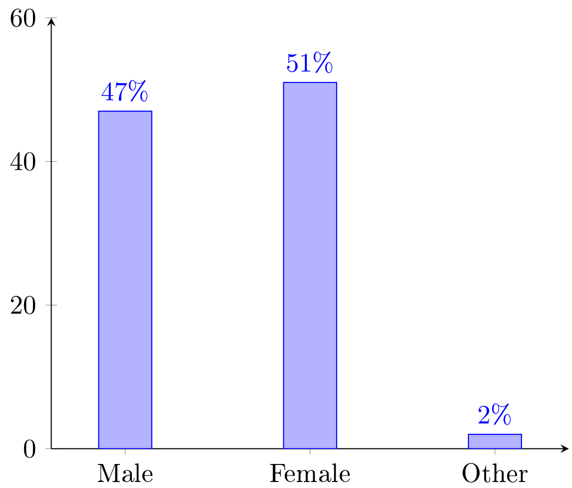


Fig. 1: Genre distribution of participants.

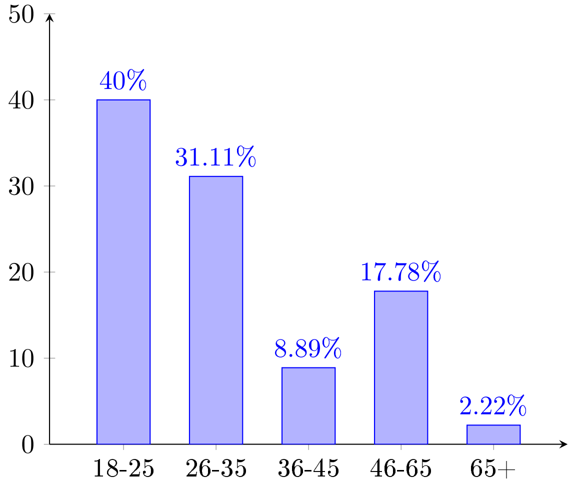


Fig. 2: Age ranges of participants.

Most of the respondents have a bachelor's degree, 66.67% to be exact, 17.78% have only finished high school, 2.22% only have a basic education, 11.11% have a master's degree and 2.22% have a doctorate, as pictured on Figure 3.

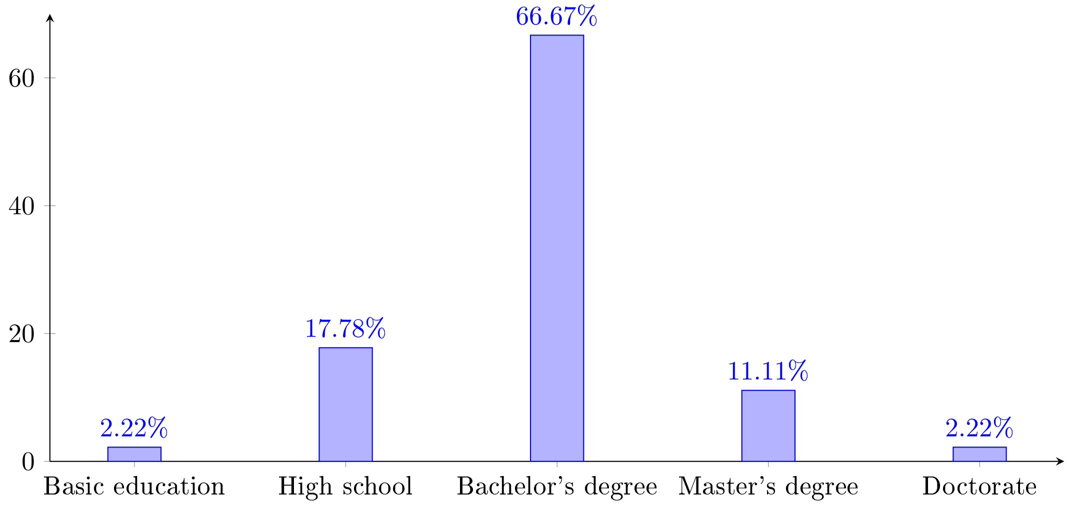


Fig. 3: Education qualifications distribution of participants.

3.2 Stage 2: An application of theory into practice

This work proposes an application that gives users information about IoT devices that inhabit their surroundings, like the type of information these devices collect and what privacy options are available. This application is developed for mobile phones due to the fact that these are the most used devices and people take them everywhere they go, this is important because the application uses georeferencing to show the location of the IoT devices. This application has two main objectives, the first is to inform and educate users in order improve their digital literacy on this particular field (privacy on IoT systems, and IoT in general) and the other being to give users a way to make informed decisions to protect their private data, in a concise and convenient place. Generally the application will show the geolocation of the IoT devices, what type of device it is, what type of data is being collect by the device. The application will not detect the devices by itself, this will be done by the users themselves, in the first iterations of the application it was discussed that the application itself would automatically detect the devices by using some kind of sniffer and would categorize what type of device it was and what type of data it was collecting but it was discovered that this approach was too complex and so it was not feasible to do with the constraints of this thesis. The application is developed with Flutter, other options considered were React Native or a progressive web application, but Flutter uses ahead of time and just in time compilation, with Dart as it is programming language, while React Native uses the Javascript programming language that was never created for mobile programming, so it uses a bridge to convert Javascript to native components for Android or iOS. Flutter has better performance and as such it was the chosen framework for this application.

The first step before creating any prototypes or starting the development was creating a software requirements specification, as can be read on Appendix 8, delineating the scope and vision for the application; the involved stakeholders; containing contextual, data-flow and swimlane diagrams; software requirements, including business, technology, functional and non-functional requirements; use cases and requirements prioritisation.

Users can, when they start the application for the first time, freely use it to see which devices are in their vicinity, information about the devices, information about the application itself, and information about privacy in general and more specifically privacy in IoT systems which they can use to improve their digital literacy. What they cannot do is add a new device to the application or edit a device's information. The user has to create an account first to do these operations. The decision to add an account creation before the user can add or edit a device is to prevent bad actors to add bogus data to the application making unusable for the majority of people, this solution doesn't completely solve this issue (because bad actors can still create an account and add bogus data anyway) but it helps to slow down the insertion of bad data.

Upon account creation the only data entry that can be considered sensitive that the user has to input is an email address. After the user has created an account and logs in, the user can add devices to the application with the following information:

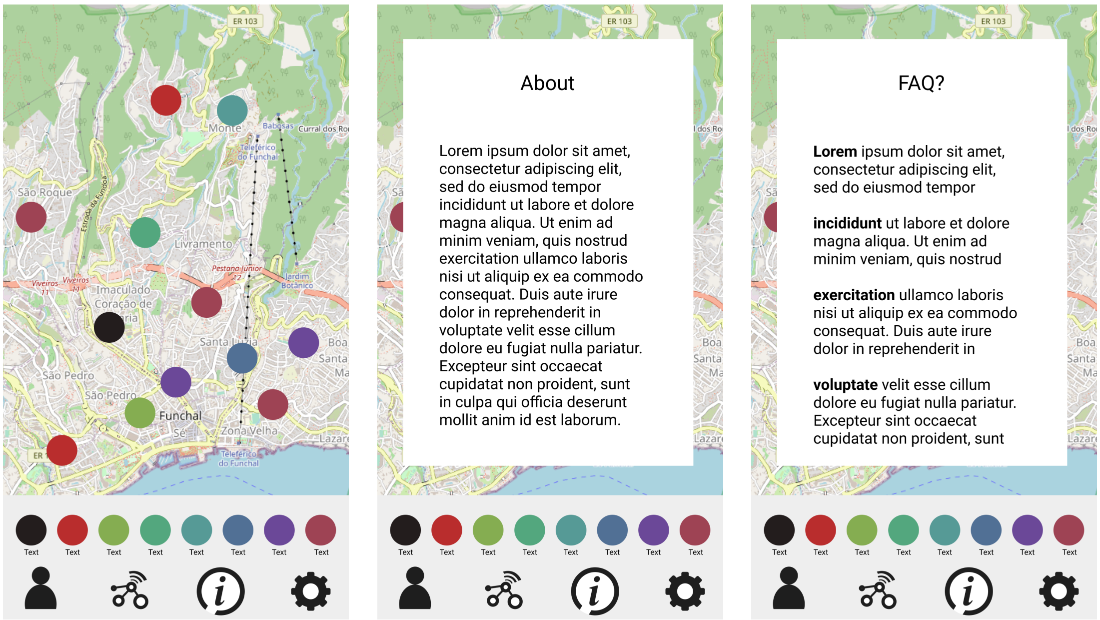
- The **name of the device**: This serves to differentiate between the various devices on the application and as such should be unique to each device, it is used on various routes and is one of the first fields that users see about a device. A single device does not have an *official* name, what is more probable is that the device has a model name or is part of a system with it's own name. The user creates the name, this could be abused by bad actors but it is extremely discouraged. It is used for aesthetic reasons.
- The **category of the device**: This is used to categorize each devices main type of information that the device is collecting. These can be of the following:
  - **Visual**: The device mainly collects visual information with maybe a video camera.
  - **Audio**: The device mainly collects audio information with a sound recorder.
  - **Presence**: The device can detect the presence of nearby objects or persons. This is not the same as the location category because the device does not know the location of an

individual, it merely knows that the individual is nearby. These type of devices can be used, for example, to collect information about how many people frequent a specific store.

- **Location**: The device can detect the exact or approximate location of an individual, it can use GPS to get this kind of information.
- **Biometrics**: The devices collects biometric data, this can be the number of steps an individual (or animal) takes, or health related data like the heart beat.
- **Environment**: These type of devices collect environmental data, they can be used for agriculture or weather forecasting by collecting, for example, temperature, humidity or wind speed/direction data.
- **Unique identification**: This category is for a device that can uniquely identify an individual, the device itself most likely is not capable of doing it but with other information that the device has access to, it can be used to cross reference of information and as such uniquely identify a person. An example of this would be a device a device that can collect visual data and with facial recognition used against other data in a database it can uniquely identify an individual.

- The **purpose for the data collection**: Defines what is the purpose for the collection of the data, if a device collects temperature and humidity data and is used by a weather based company or government agency then the purpose for the data collection is for weather forecasting.
- **Who has access to the collected data**: Disclose an individual or group of individuals that have access to the data of the device, if the device is part of a closed system it can be that only an individual as access to the data but most likely various groups of people have access to the data, some with more data than others depending on the permissions they have. If the device publicizes its data then everyone has access to it.
- **For how long is the data stored**: Pinpoint the duration of the stored data in the device or system, due to legislation passed in various countries this duration has a limit, in some cases the data cannot be stored for more than one year.
- **Can the data identify anyone**: Used to quickly identify if a particular device can identify an individual or not. If the device belongs to the "Unique identification" category then this should be active.
- **What is being done with the data**: This could be assumed to be similar to the **purpose** field mentioned above but it should be used to diagnose what is being done now with the data collected, in certain situations it might coincide with the purpose for the data collection.
- **Privacy options**: The user can insert an url for the device's privacy options, in some cases the device, or company, has a website with information on privacy options, or privacy policy. If the device uses a mobile application then a link to the this application can be inserted here.
- **Coordinates of the device**: Used to express the latitude and longitude of the device so that it can be shown on the map, on the homepage of the application.
- **Who owns the device**: Who is the device owner, if it belongs to an organization then the name of the organization should appear here otherwise if the device belongs to a person then the person's name should **not** appear, it should say private or something similar.

The user is not required to provide information to satisfy all items on this list, the only information that is required in order to add a device to the application is the name, category and coordinates of the device, all other information is optional but should be provided for the sake of guaranteeing a good experience to other users of the application. The information provided should be verified by the user beforehand so that bogus data does not clutter the application, in this case there are no absolute ways of guaranteeing this but the maintainer of the application edit wrong information or in some cases remove it, other users can also edit any device data. This is an open platform so it is expected that users act in good faith.

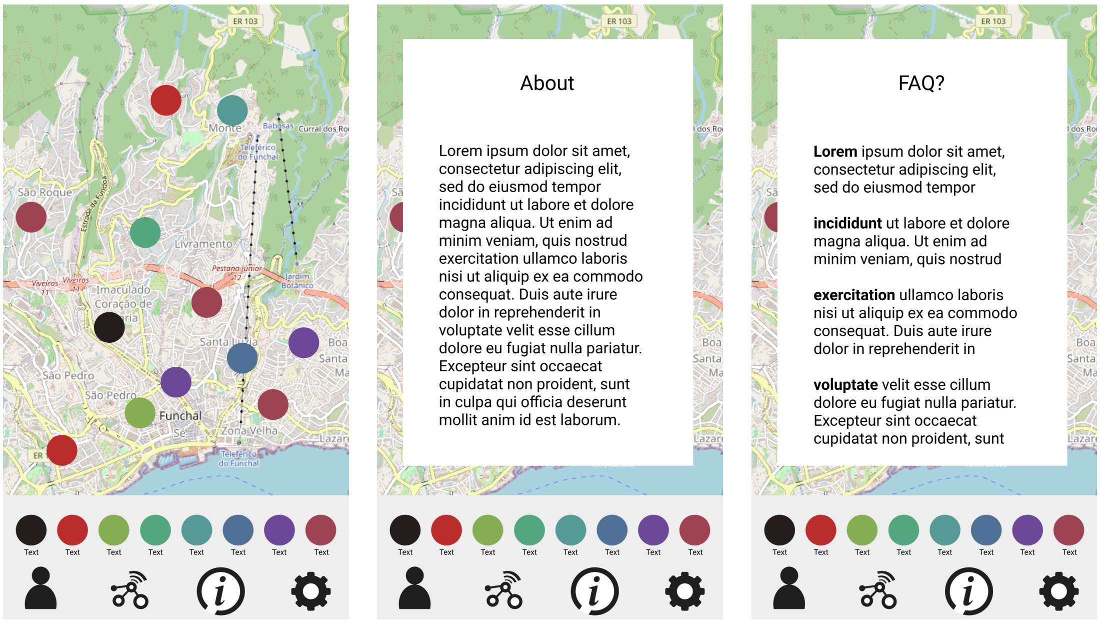


(a) (b) (c)

Fig. 4: Low level prototype of (a) homepage, (b) about and (c) FAQ pages.

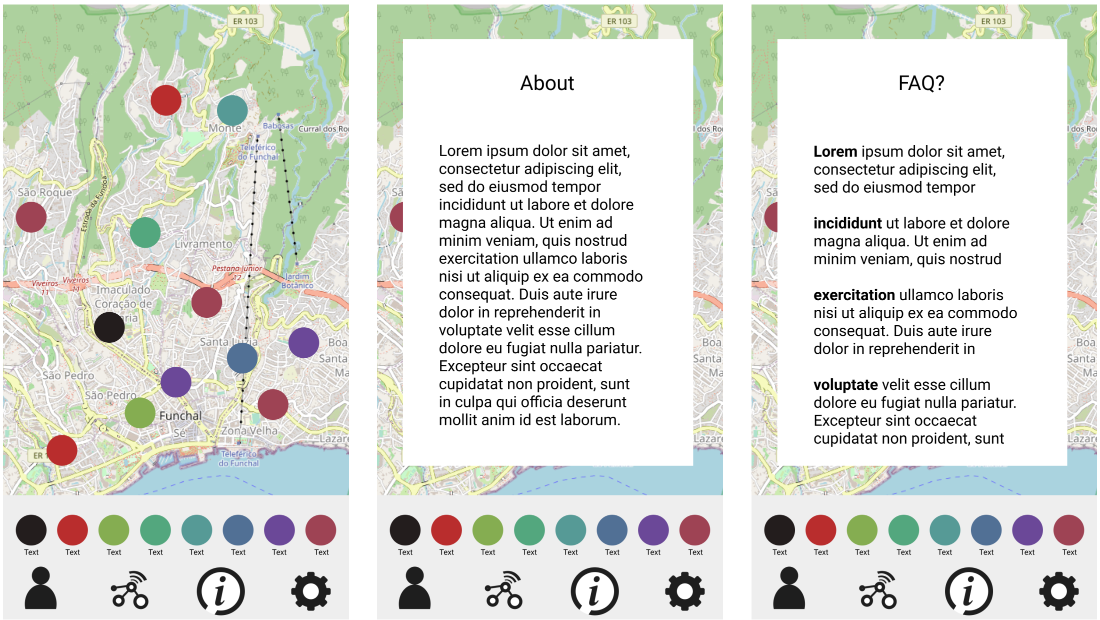
After creating the software requirements specification, the prototypes were created. For the creation of the prototypes the following tools were used: Figma and GIMP.

At first a low level prototype was made in order to understand the general design and user interaction of the application. Figure 4 shows tree pages of the low level prototype, these are the homepage, about and faq pages, this prototype has a navigation menu on the bottom where the other pages of the application can be selected along with some information above the page icons, this information is supposed to be the categories of the devices, the logic would be that the user could tap one of these categories and only devices of that category should be displayed on the map. It can be seen that between the three pages the map stays in the background and the various pages work like an overlay on the homepage, this would be changed in subsequent prototype versions.



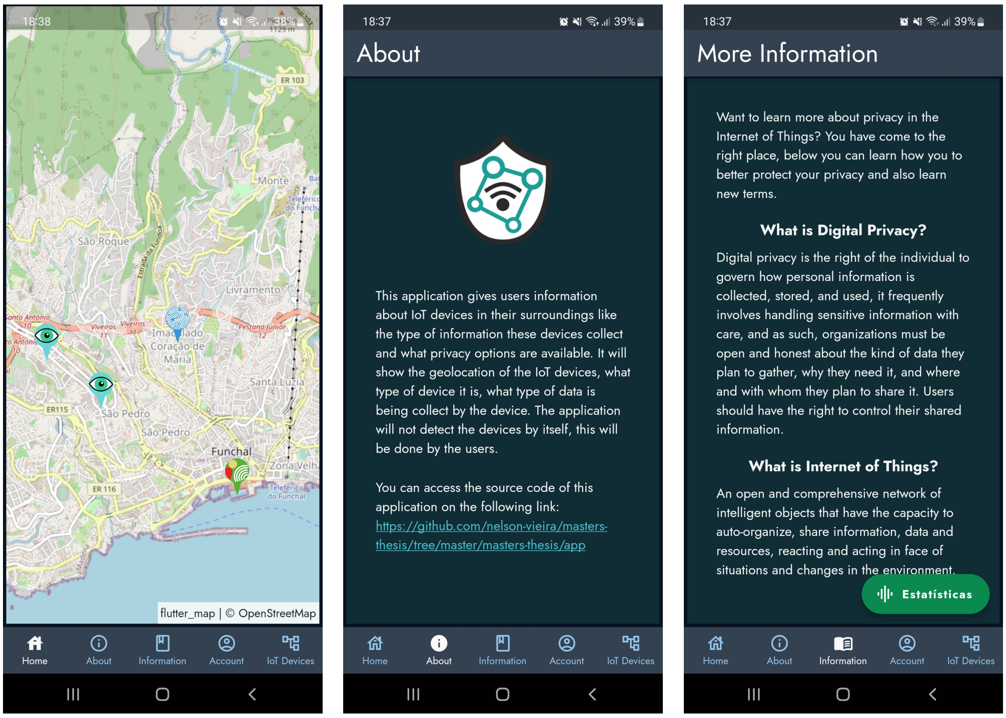
(a) (b) (c)

Fig. 5: Medium level prototype of (a) homepage, (b) about and (c) FAQ pages.



(a) (b) (c)

Fig. 6: High level prototype of (a) homepage, (b) about and (c) FAQ pages.



(a) (b) (c)

Fig. 7: Live version of the application with pages: (a) homepage, (b) about and (c) FAQ pages.

Usability tests were conducted in person with 7 participants of different ages, professional fields and qualifications. Before doing the tests, some questions were made to gather the general level of digital literacy related to IoT and privacy, then the participants were asked to fill in the survey, if they had not done it yet, as this gives some insight into what the application is about. The usability tests consists of single ease questions and system usability scale, as can be seen on appendix 8. The single ease question was used after the participant performed each task, the participant would answer how difficult they though the task was in a scale of 1 to 7. The system usability scale was used after the participants performed all tasks.

4 Results

This chapter presents the results of the questionnaire about participants' perceptions of privacy, IoT and theirs online behaviors, as well as the usability tests that were conducted for the mobile application.

4.1 Stage 1: Survey

As would be expected, the majority of participants agreed that data privacy is important to them and as such should be protected, as can be inferred on Figure 8. Most participants also have some knowledge of behaviours or techniques to do in an online environment, as shown on Figure 9, be it connected to the internet or on some local network, like not sharing too much private information (or none at all) with strangers, revealing only the bare minimum necessary information to use a particular system, use of VPNs, different and strong passwords, 2 step authentication methods between others. Because a good portion of participants are from engineering areas this question might be skewed.

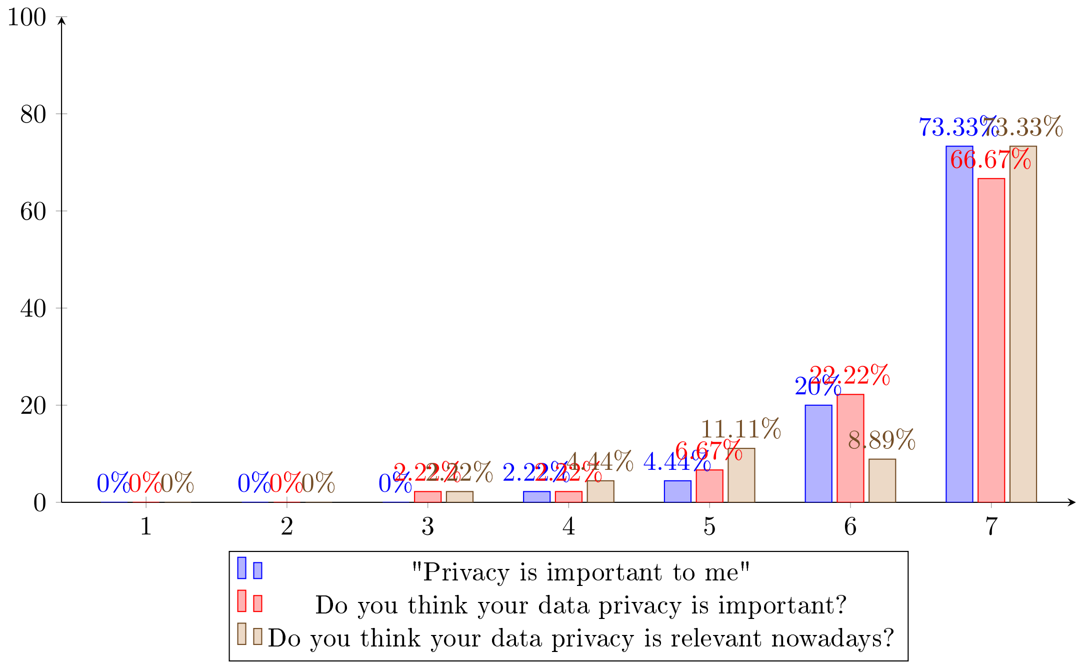


Fig. 8: Perspectives of participants on the importance of personal data privacy.

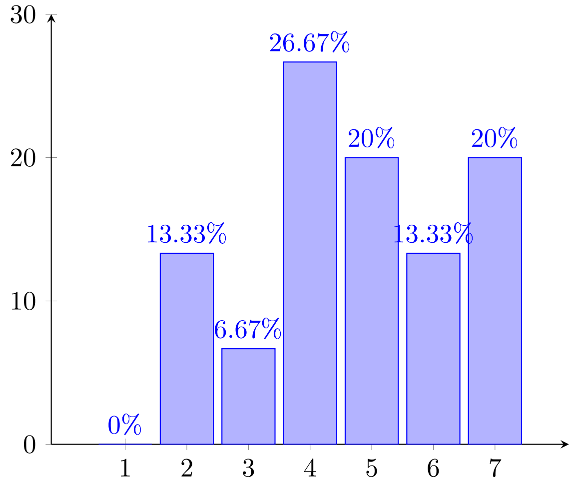


Fig. 9: Participant responses indicating whether they know techniques to guarantee privacy and the protection of their data when using the internet.

Most participants consider data privacy as a human and consumer right, as shown on Figure 10, even if they have no knowledge of article 12 of the Universal Declaration of Human Rights.

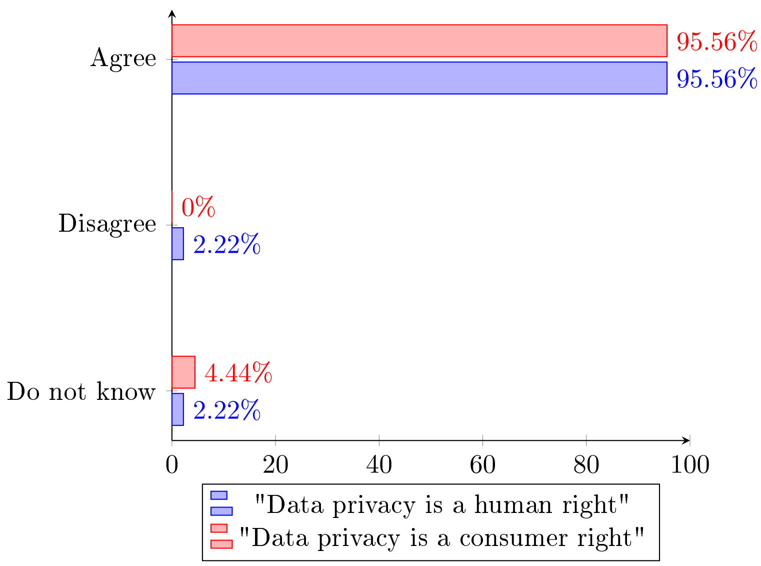


Fig. 10: Participant responses indicating whether they agree or disagree with these statements.

When asked to define digital privacy, most participants did not know how to properly define it, giving generic answers while some even gave a one word answer, some participants gave incomplete or adjacent related answers. Only approximately 16% of participants supplied a concrete answer that was close to the definition presented in Section 1. Curiously, no one mentioned security, which contradicts with the responses represented in Figure 11, where 53.33% of participants believe that privacy and security are synonymous.

Participants have some digital literacy of IT terms, as shown on Figure 12, most know the more popular terms like *wi-fi*, *cookie* and *data protection*, but as the terms become more esoteric the

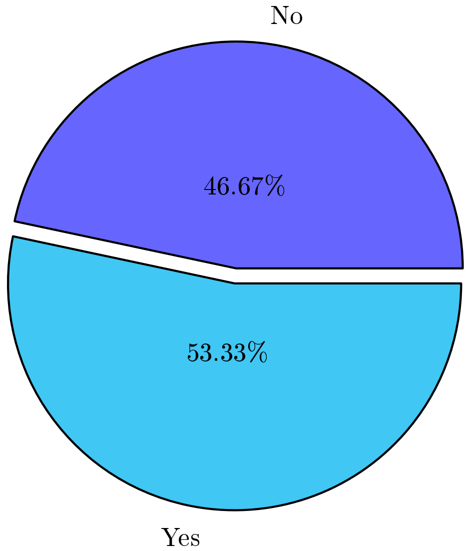


Fig. 11: Responses to security and privacy being synonymous.

general knowledge starts to drop. Some terms, even after getting some popularity, are still mostly unknown to the majority of people like *blockchain* or *Internet of Things*.

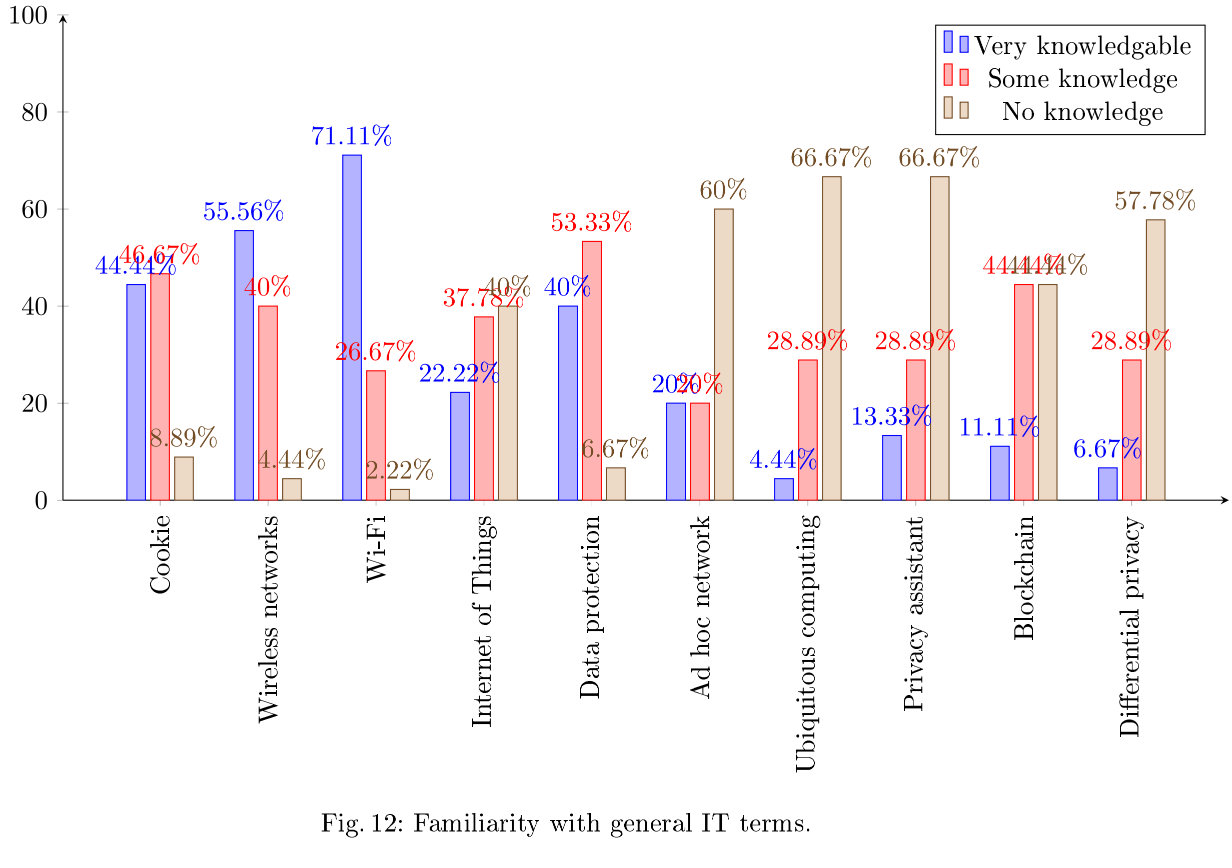


Fig. 12: Familiarity with general IT terms.

Regarding users' online habits, all participants have, or have access to, a smartphone and they use it in their daily lives, most participants concede that they spend a lot of their daily time using it, like shown on Figure 13, and are somewhat worried but do not actively try to protect their data privacy, which goes against their earlier responses regarding their worry of personal data privacy. When asked if they accept cookies, respondents occasionally do but are unsure what they do or their importance on the user experience.

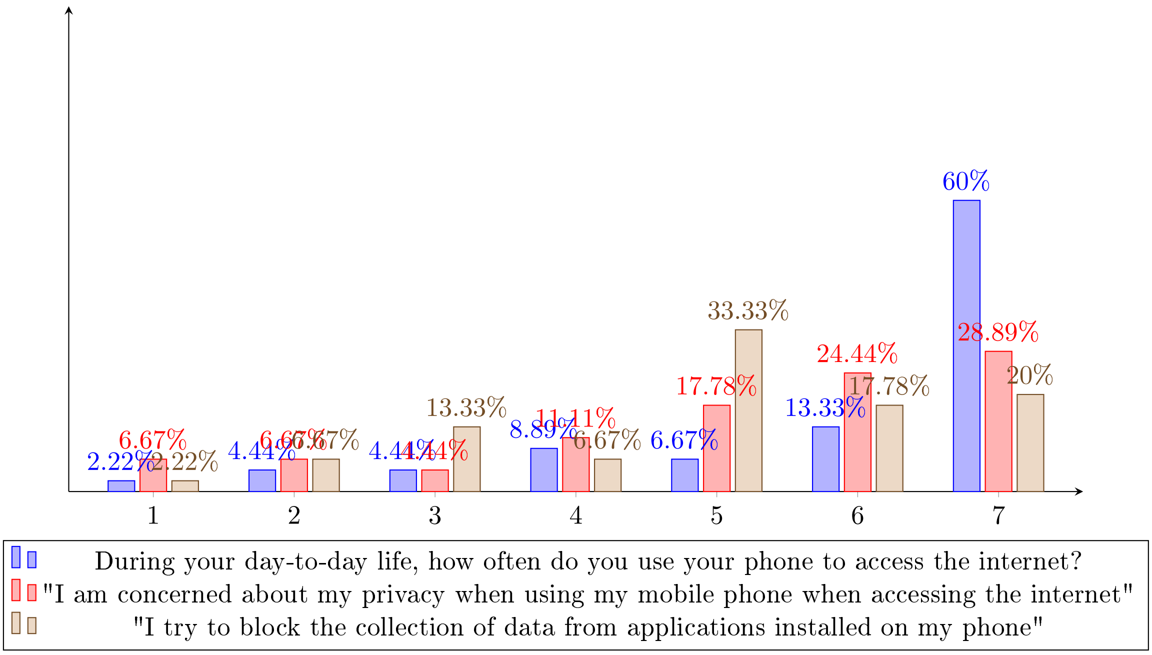


Fig. 13: Responses related to phone usage.

When asked about the concept of profiling, only half of the participants are aware of the term but more than half consider that their online activity contributes to its development, see Figure 14.

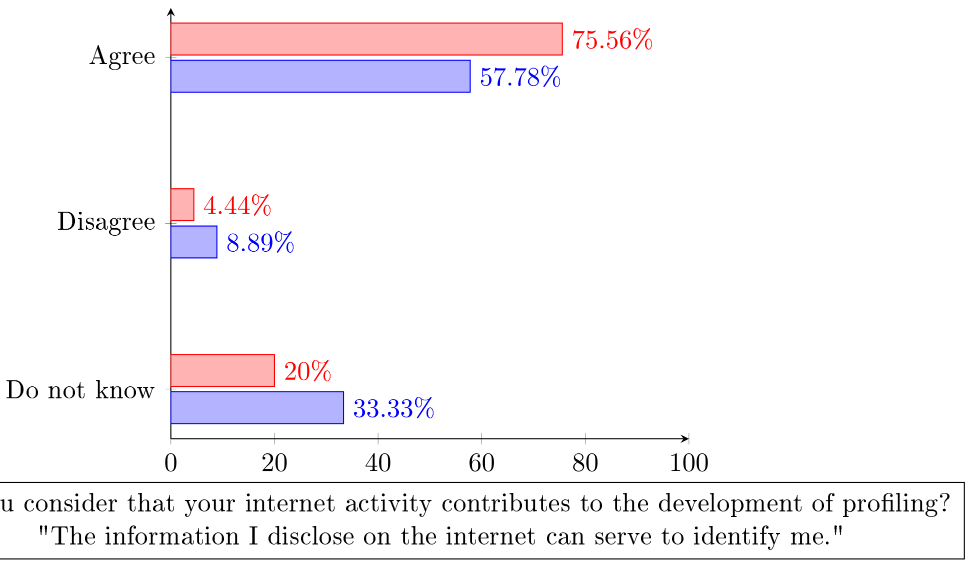


Fig. 14: Participant responses indicating whether they agree or disagree with these statements.

Regarding participants relation with regulations, most participants are aware of regulations related to digital privacy, like GDPR or CCPA, but it's not absolutely clear to them what they represent and they do not show a clear interest in know more about these regulations, as shown on Figure 15. It is not clear why there is a certain disinterest by most participants in regulations, it might be that most find it frustrating to read, or there is some faith that they work and as such there is no need to know more about them. Notably 0,47% of the participants responded correctly to the question about GDPR's definition of personal data.

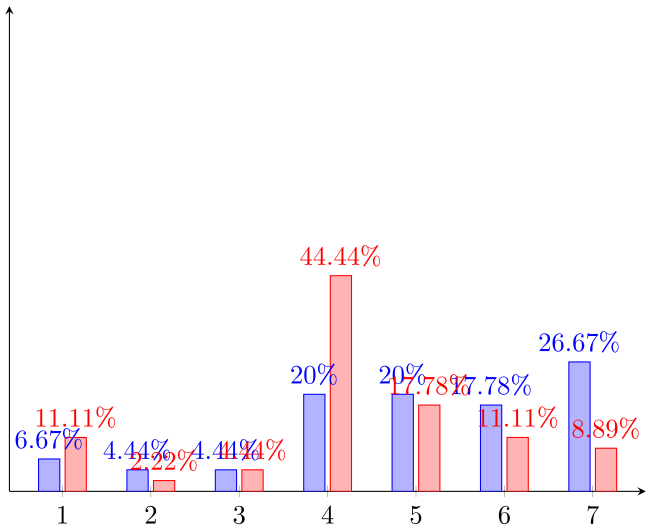


Fig. 15: Responses related to privacy regulations.

Most participants, 73.33%, do not know what a data protection officer is or what are their duties, as can be seen on Figure 16. An internal role known as a data protection officer acts as an independent spokesman for a company's policies regarding the processing and application of customer data. The GDPR, which requires the appointment of a data protection officer by all businesses that offer goods or services to consumers in the European Union and collect data as a result, led to the creation of this job. The data protection officer keeps up with data protection regulations and policies, conducts internal privacy audits, and ensures that all other compliance-related data-related concerns are current.

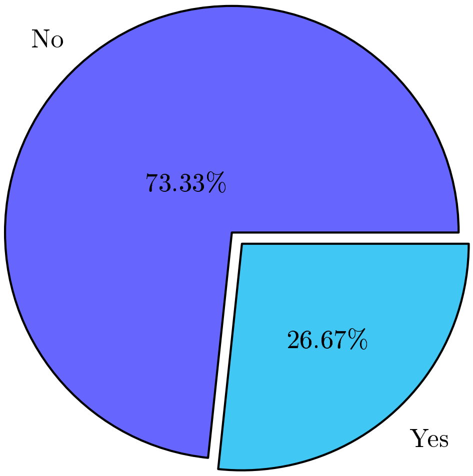


Fig. 16: Awareness of the duties of a data protection officer.

As expressed before, participants consider the protection of private data important and as such 82.22% are interested in knowing where and how their data is used by organizations, as shown on Figure 17.

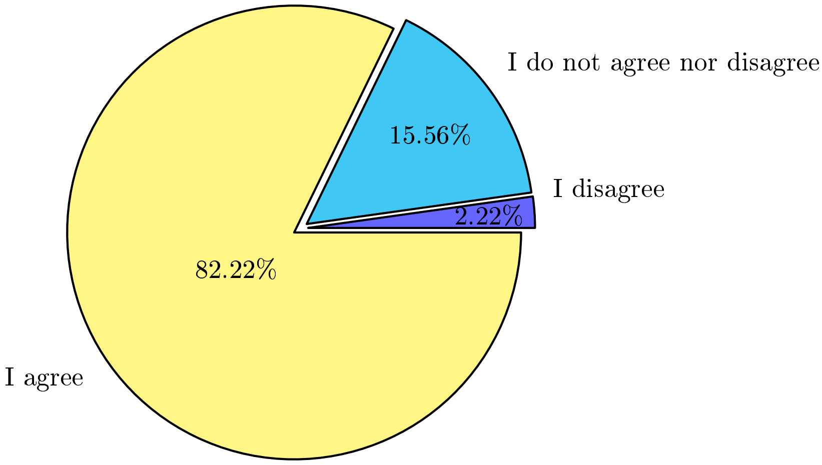


Fig. 17: Interest in knowing where and how personal information is used by organizations.

Figure 18 depicts questions related to privacy notices. Participants are interested in knowing how their data is used by organizations, and privacy notices are one way to convey such information, but not all privacy notices are created equal, specially in the IoT, yet 53.33% are not familiar with the reason data is collected but are still curious. Most participants agree that lengthy texts and font size, 86.67% and 77.78% respectively, affect their willingness to read them, 71.11% confess to agreeing to the notices just to use the service due to a fear of missing out. In regards to brand trust, participants are divided with 33.33% agreeing to not reading the privacy notice if they trust the organization and 33.33% disagreeing.

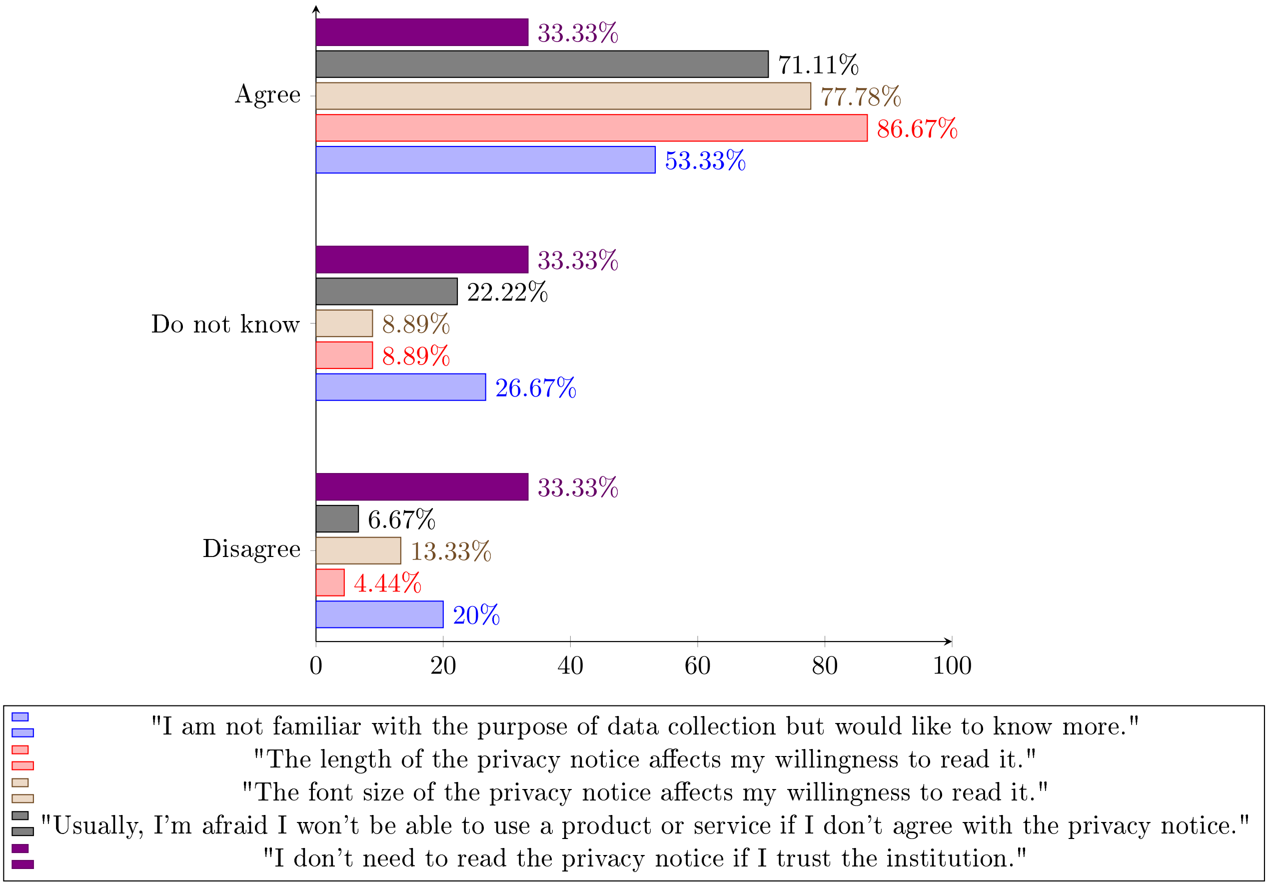


Fig. 18: Participant responses indicating whether they agree or disagree with these statements.

Figure 19 depicts questions related to participants privacy related concerns. Participants answered almost unanimously to all questions with most answers being 7, the questions that deviate from the rule concern public institutions or intelligence services analysing participants' online movements and organizations collecting and using participants' online activity respectively where most responded between a 5 and 6 meaning that some participants are not as concerned when public institutions or organizations collect their online activities. Participants show more concern when organizations collect data without their consent or when others engage in identity theft.

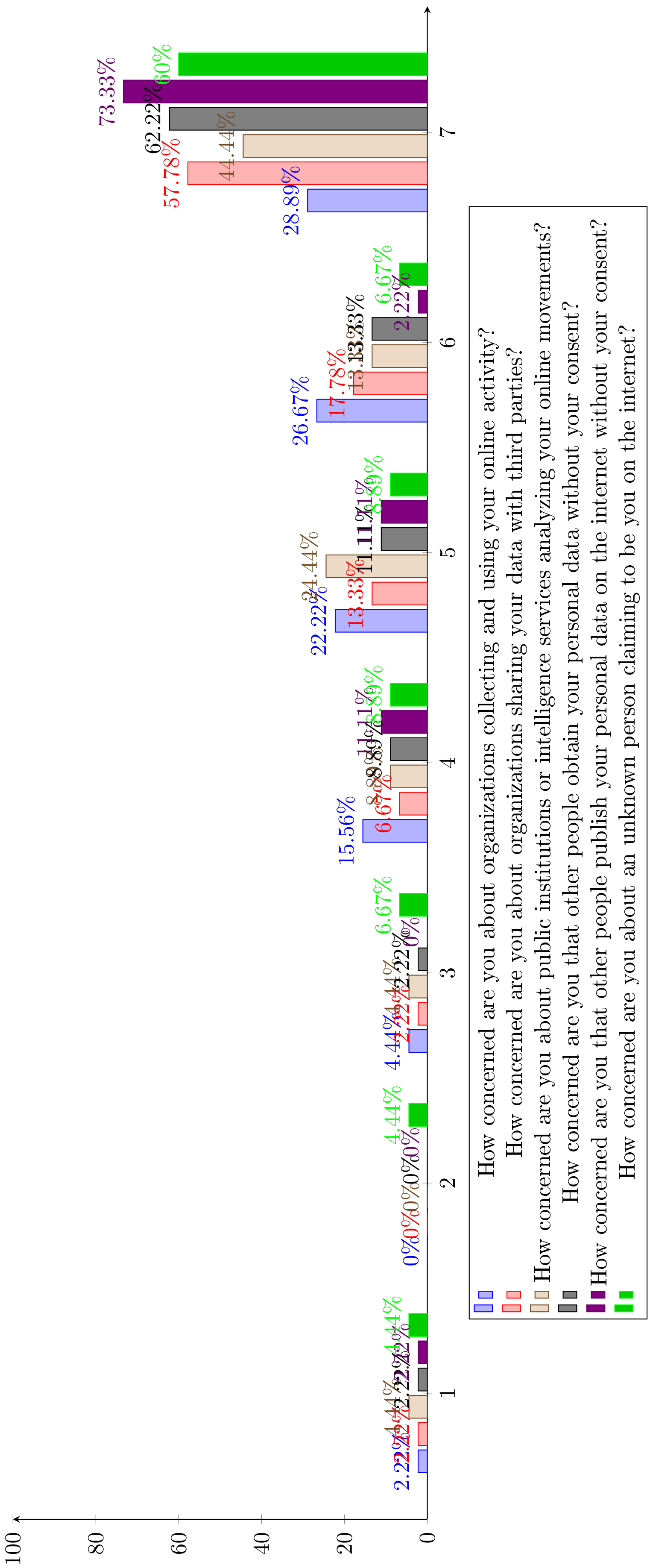


Fig. 19: Responses related to concerns of organizations and individuals handling of private data.

The majority of participants admitted to inputting varying amounts of incorrect information, with only 15.56% of participants said they have never entered fake information when setting up an account on a platform, and 6.67% said they always do, as demonstrated on Figure 20. The most typical inaccurate information entered by participants concerns their age, phone number, complete name, birthday, electronic mail, username, home address and country of citizenship. The most common reasons participants give for entering this false information are when creating a temporary account or they find that the information is not relevant, some stated that they do not want to provide any information that may be used to identify them.

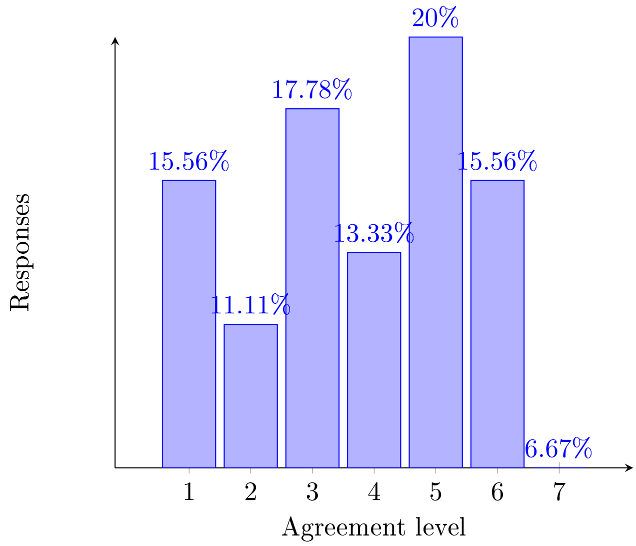


Fig. 20: Responses regarding the input of false personal data when creating an account on an online platform.

Most participants are unaware of how digital data flows, as 82.22% state being unfamiliar with data brokers, according to Figure 21. But they are aware that organisations collect their private data and use it in various ways, even without knowing exactly what they do with the data according to previous responses. If more individuals were aware of the function of a data broker, which plays an important role on data flow and can be an individual or a company which specializes in collecting and selling private data, their decisions undoubtedly be affected by this knowledge.

In terms of IoT usage and literacy, about 2/3 of the participants have interacted in some way with a device, and from those who have used an IoT device, only close to a third (35.56%) actually have a device in their homes, as depicted on Figure 22. Even though most participants have interacted with an IoT device they had trouble actually explaining what Internet of Things actually is. When asked to describe it, 48% of the participants wrote some kind of “do not know” variation, from the other 52% that gave a more elaborate answer 70% gave a generic answer that could be applied to a number of things, some described it like objects that are connected to the internet. Only two participants managed to describe IoT accurately. Most participants know devices that can be characterized as belonging to the Internet of Things like smart homes devices, including smart appliances like fridges, smart locks or virtual assistants that function with natural language commands, so even though the majority of respondents do not know how to define they can recognize IoT devices.

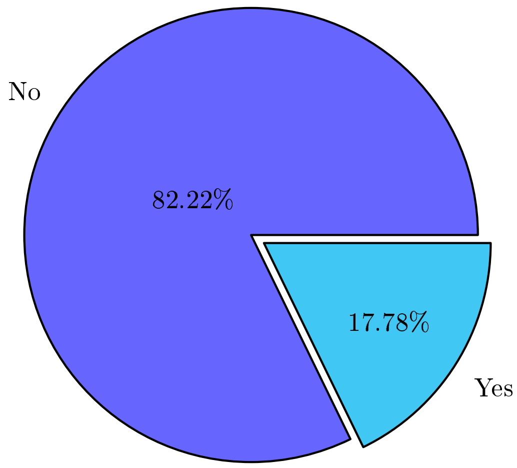


Fig. 21: Awareness of data brokers.

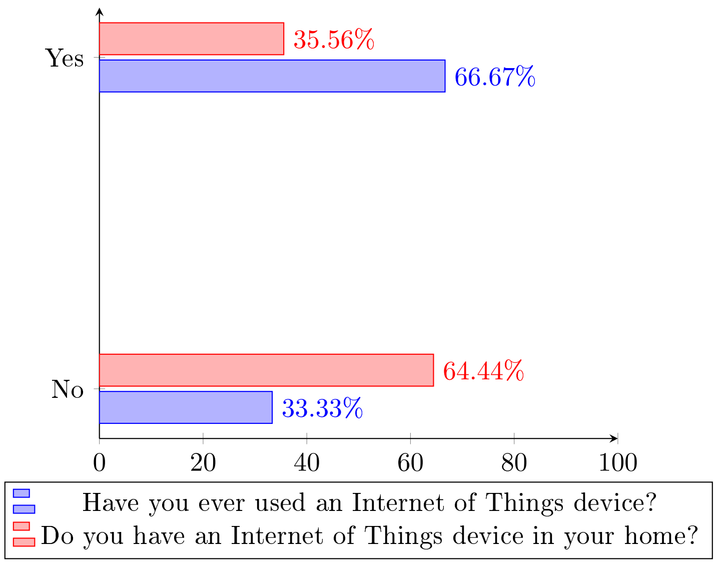


Fig. 22: Participant responses indicating whether they agree or disagree with these statements.

Familiarity with devices does not extend to most IoT related terms, as illustrated in Figure 23, the most recognized terms are the ones who have more visibility in everyday day like smart home, smart vehicle or fully autonomous car. Esoteric terms like the ones describing networks or protocols are mostly unknown to most respondents, it is known only to those that have certain qualifications or who work in IT or similar areas.

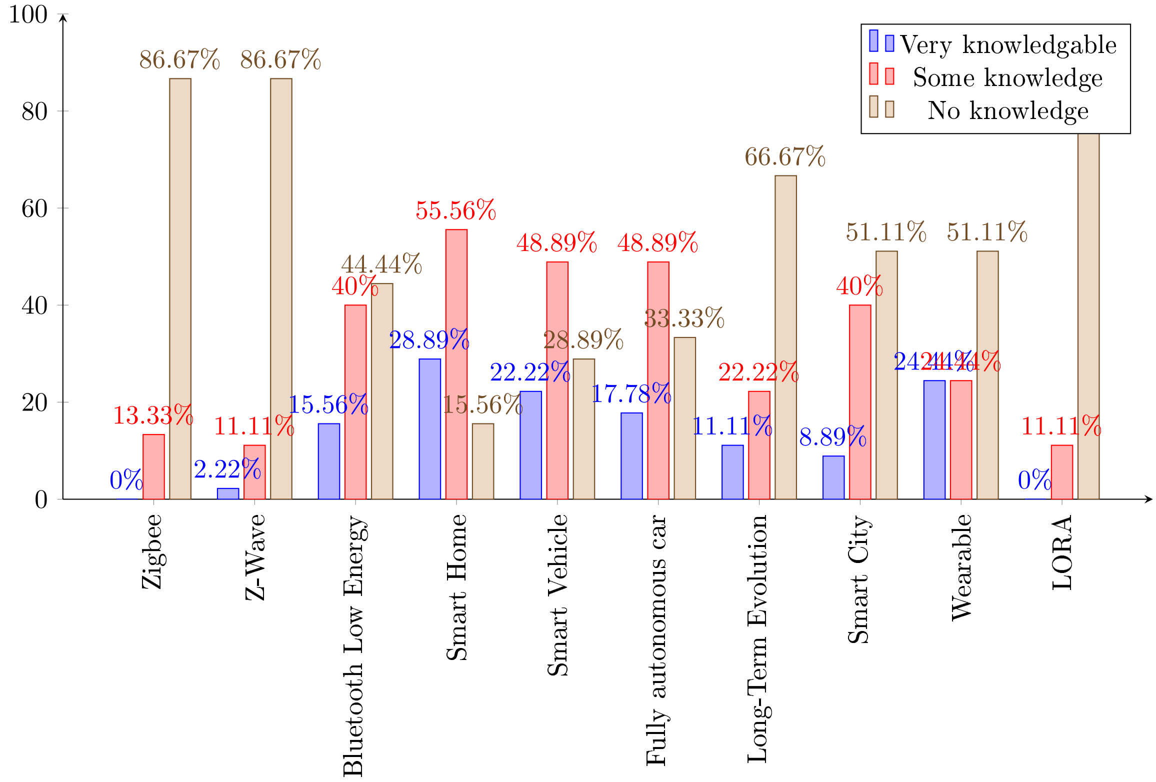


Fig. 23: Familiarity with IoT terms.

4.2 Stage 2: Application

The usability test was given to each participant before giving a basic debrief of what the application is about. Participants would score the overall difficulty after performing each task. When participants first used the application they found it to be similar to other mapping applications, they had some difficulty adjusting to the application's function but were quick to adapt. The tasks consisted of navigating the application and performing certain actions such as creating a user account or adding an IoT device, these tasks include navigating to the IoT devices page; looking up information about any one particular device; navigating to the homepage and performing the same action as the previous step; looking up general information about the application itself; creating an account; looking up information about privacy, the Internet of Things, and how to add a device to the application; adding an IoT device; and finally, updating an IoT device. The first tests revealed to be somewhat difficult because of a language barrier, the application, at the time, was only available in english and the participants had a rough grasp of it. It was still possible to conduct the tests after translating everything the application did. Some participants struggled to find the IoT devices on the application at the start, but eventually it became second nature. There was also some difficulty in adding an IoT device, mostly from older participants. Participants who were already comfortable with technology or who had worked in fields connected to IT did not have any trouble completing each task, although only one of these participants had some knowledge of IoT.

The system usability scale was used after the participants performed all tasks, using this system a median score of 92,33% was achieved. This is a high score which would mean that the application provides a good user experience, even though some users experienced some hardships when using it when the first tests where being performed. The feedback given by these participants helped in improving the application, as is the example of when the application when only available in english, now it is possible to use it in english and in portuguese, in the future more language might be added.

An improvement to the usability tests would be to have the participation of an expert software developer, preferably with experience in IoT systems.

5 Discussion

This section serves to discusses the main takeaways found in the literature review and address the results of the questionnaire and the usability tests, and any other interesting remarks found during the course of this work. The research questions that were posed in Chapter 1 will also be addressed

5.1 Systematic literature review Findings

There as been a rising interest by research papers in addressing privacy issues or challenges in the Internet of Things, at least since 2015-2016 where the number or papers jumped from 11600 in 2015 to 19400 in 2016 and rising since then until reaching a peak of 74400 in 2020 and slightly declining since then, and at least until August of 2023 there have been around 36000 papers published, according to the data seen on Figure 24.

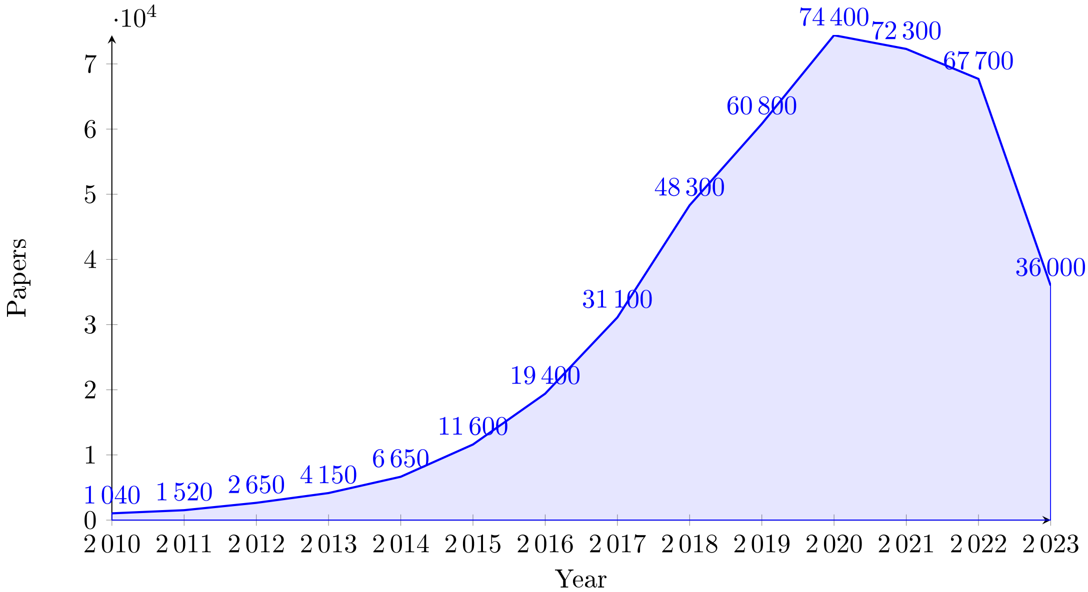


Fig. 24: Distribution of papers, per year since 2010, on privacy in the Internet of Things, based on Google Scholar database results.

Figure 25 depicts the number of publications per topic in multiple databases, all of which are still related to the main theme of privacy in the IoT, particularly Google Scholar, BASE, and CORE. This data is from papers since 2020, the database with the most papers is Google Scholar followed by the distant second BASE and finally CORE with the smallest number of papers. Around 249000 papers exist discussing privacy in the IoT in Google Scholar, 20641 papers on CORE and 8014 on BASE, from these papers the most researched topic has been security with around 196000 papers on Google Scholar, 17738 on CORE and 5600 on BASE. This data cannot be accepted at face value because searches on these databases are conducted using keywords, these keywords may occur in the title of the work or on the content, but this does not imply that the work is about keywords in general, rather, it simply indicates that the keywords appear on the work. But it can represent a rough estimate of what is being researched. Other databases were used in course of the research like Semantic Scholar, Baidu Scholar and RefSeek but these databases do not have good filters for advanced searching and/or have a small number of papers.

There is a big focus on security, AI and blockchain in recent literature, specially from 2020 onward, these finding are in accordance with other similar studies [123].

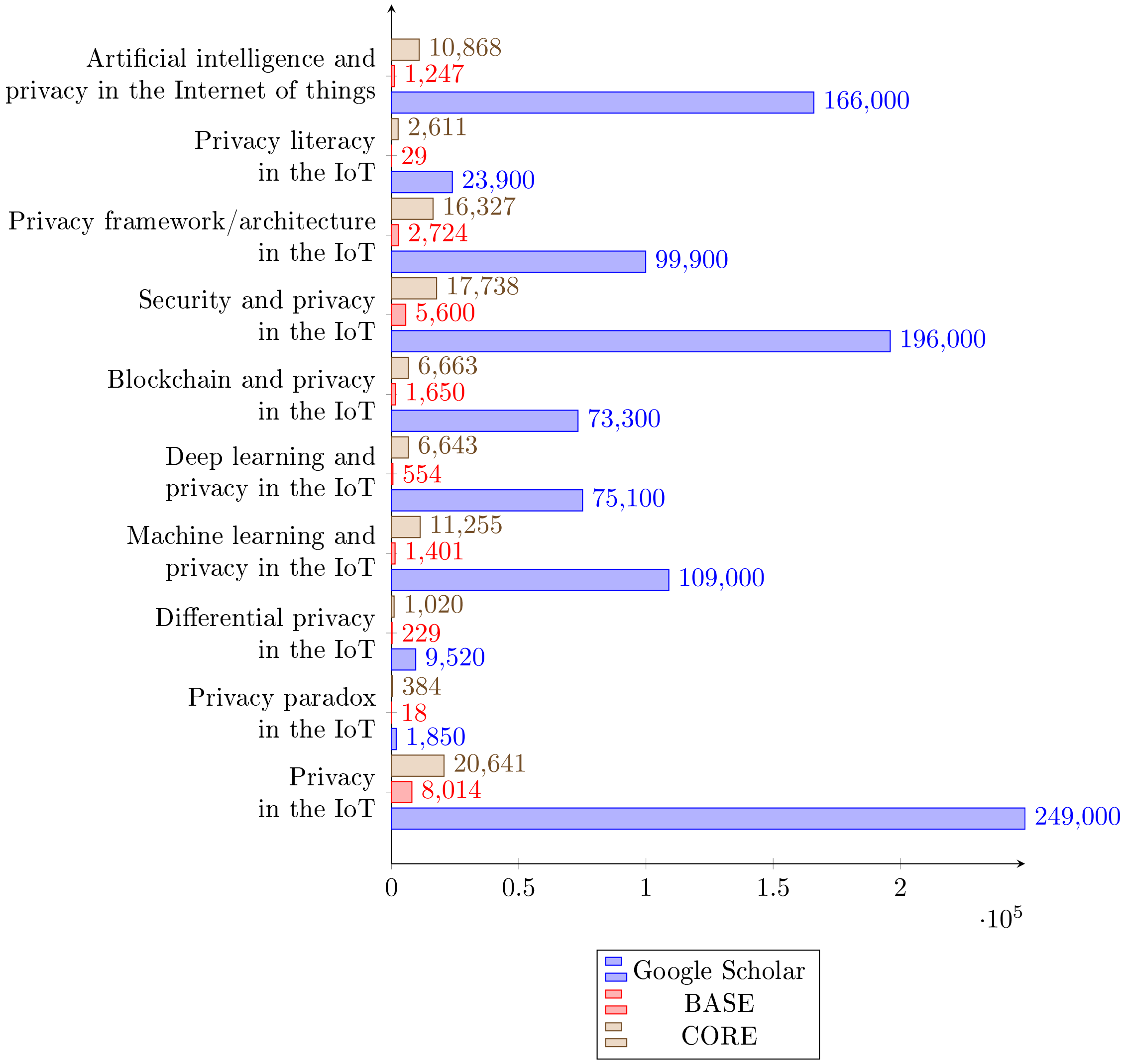


Fig. 25: Distribution of papers on privacy in IoT, since 2020, by keywords searches on Google Scholar, BASE and CORE databases.

Chapter 2 discusses many works that addressed privacy challenges in the IoT and some that proposed approaches to address these issues. Different aspects of the included papers are now assessed, including their general information, purpose, intended application, issues that still exist, and any other comments that should be taken into account.

Table 2 displays general information about the papers that were included [43–45, 47–58, 60, 61, 65–70, 73–77, 80–84, 86–88, 90, 91, 96, 104–112, 124]. It is discernible that 22.92% of the total papers were published in 2022 [60, 65, 66, 69, 70, 77, 80, 86, 88, 90, 110], while 12.5% of the papers were published in 2017 [54, 68, 76, 82, 96, 106], in both 2018 [58, 74, 87, 105, 109] and 2015 [44, 45, 50, 67, 81], there were 10.42% of papers published, in 2013 [47, 48, 55, 57] 8.33% of papers were published, with the same percentage in 2014 [51, 52, 75, 112] and in 2020 [61, 73, 91, 107], the remaining 16.67% of papers were published between the years 2010 [83], 2012 [43, 49], 2019 [56, 104], 2021 [108] and

2023 [84, 111], with the exception of a small percentage of 2.08% being published in 2007 [53]. Overall, 39.58% of the literature originate from the United States of America (USA) [43, 44, 47, 48, 53, 55–58, 65, 76, 77, 80, 87, 90, 107–109, 112], 10.42% originate from China [68, 86, 96, 104, 105], 10.42% are from Germany [49, 50, 52, 66, 75] and 62.79% are spread across 13 different countries [45,51,54,60,61,67,69,70,73,74,81–84,88,91,106,110,111]. The majority of the included papers, 72.92% [45, 47, 48, 50–57, 60, 61, 65–70, 73–75, 77, 81, 83, 84, 86, 88, 91, 96, 104, 105, 109–111], are published in journals, while 25% are published in conferences [43, 44, 49, 58, 76, 82, 87, 90, 106–108, 112], and 2.08% in a magazine [80]. The publishers' information is provided so that the database of origin can be easily accessed. The ACM database [44, 58, 60, 80, 86, 88, 90, 104, 106–108, 110, 112] provides 27.08% of all publications, followed by Elsevier [45, 47, 48, 51, 54, 57, 66, 67, 77, 81, 83, 91] with 25%, IEEE [61, 65, 68, 82, 84, 87, 96, 105, 109] with 18.75%, MDPI [70, 74, 111] with 6.25% and Wiley [50, 53, 75] with the same percentage, and 16.66% are divided between the other 7 publishers.

Ref #	Year	Country	Publication Type	Publisher	Scope
[43]	2012	USA	Conference	**AIS** 33rd International Conference on Information Systems	Privacy paradox
[44]	2015	USA	Conference	**ACM** Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems	Privacy paradox
[45]	2015	South Korea	Journal	**Elsevier** Expert Systems with Applications	Privacy paradox and healthcare
[47]	2013	USA	Journal	**Elsevier** International Journal of Human-Computer Studies	Privacy paradox
[48]	2013	USA	Journal	**Elsevier** The Journal of Strategic Information Systems	Privacy paradox
[49]	2012	Germany	Conference	**Springer** International Symposium on Quantum Interaction	Privacy paradox
[50]	2015	Germany	Journal	**Wiley** European Journal of Social Psychology	Privacy paradox
[51]	2014	South Korea	Journal	**Elsevier** Computers in Human Behavior	Privacy paradox
[52]	2014	Germany	Journal	**Oxford University Press** Journal of Computer-Mediated Communication	Privacy paradox
[53]	2007	USA	Journal	**Wiley** Journal of Consumer Affairs	Privacy paradox
[54]	2017	Greece	Journal	**Elsevier** Computers & security	Privacy paradox
[55]	2013	USA	Journal	**SAGE Publications** Social Psychological and Personality Science	Privacy paradox
[56]	2019	USA	Journal	**Taylor & Francis** Journal of Interactive Advertising	Privacy paradox
[57]	2013	USA	Journal	**Elsevier** Information & Management	Online behaviour

continues ...

...continued

Ref #	Year	Country	Publication Type	Publisher	Scope
[58]	2018	USA	Conference	**ACM** Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems	Privacy paradox
[60]	2022	Australia	Journal	**ACM** Computing Surveys	Differential Privacy
[61]	2020	Singapore	Journal	**IEEE** Internet of Things Journal	Differential Privacy, Federated Learning
[65]	2022	USA	Journal	**IEEE** Internet of Things Journal	Systematic Literature Review
[66]	2022	Germany	Journal	**Elsevier** Pervasive and Mobile Computing	Systematic Literature Review
[67]	2015	Italy	Journal	**Elsevier** Computer Networks	Systematic Literature Review
[68]	2017	China	Journal	**IEEE** Internet of Things Journal	Systematic Literature Review
[69]	2022	Norway	Journal	**Frontiers** Frontiers in Artificial Intelligence	Systematic Literature Review
[70]	2022	Hungary	Journal	**MDPI** Sensors	Systematic Literature Review
[73]	2020	India	Journal	**Springer** Wireless Personal Communications	Systematic Literature Review
[74]	2018	Greece	Journal	**MDPI** Smart cities	Literature Review
[75]	2014	Germany	Journal	**Wiley** Security and Communication Networks	Systematic Literature Review
[76]	2017	USA	Conference	**USENIX** Proceedings of the Thirteenth Symposium on Usable Privacy and Security	User awareness
[77]	2022	USA	Journal	**Elsevier** International Journal of Information Management	User awareness
[80]	2022	USA	Magazine	**ACM** Communications of the ACM	User awareness
[81]	2015	Switzerland	Journal	**Elsevier** Computer Law & Security Review	Legislation
[82]	2017	Italy	Conference	**IEEE** International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)	Legislation
[83]	2010	Switzerland	Journal	**Elsevier** Computer Law & Security Review	Legislation
[84]	2023	Bosnia and Herzegovina	Journal	**IEEE** Communications Magazine	Legislation
[86]	2022	China	Journal	**ACM** Transactions on Sensor Networks	Security

continues ...

. . . continued

Ref #	Year	Country	Publication Type	Publisher	Scope
[87]	2018	USA	Conference	**IEEE** 2018 IEEE International Conference on Acoustics, Speech and Signal	Security
[88]	2022	Brazil	Journal	**ACM** Transactions on Intelligent Systems and Technology	Healthcare
[90]	2022	USA	Conference	**ACM** Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing	Framework
[91]	2020	United Kingdom (UK)	Journal	**Elsevier** Information Sciences	Framework
[96]	2017	China	Journal	**IEEE** Internet of Things Journal	Framework
[104]	2019	China	Journal	**ACM** Computing Surveys	Blockchain
[105]	2018	China	Journal	**IEEE** Wireless Communications	Blockchain
[106]	2017	Italy	Conference	**ACM** Proceedings of the Seventh International Conference on the Internet of Things	Blockchain
[107]	2020	USA	Conference	**ACM** Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems	Privacy Assistant
[108]	2021	USA	Conference	**ACM** Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems	Privacy Assistant
[109]	2018	USA	Journal	**IEEE** Pervasive Computing	Privacy Assistant
[110]	2022	Australia	Journal	**ACM** Transactions on Internet of Things	Smart City
[111]	2023	Portugal	Journal	**MDPI** Electronics	Framework
[112]	2014	USA	Conference	**ACM** Proceedings of the 2014 ACM Conference on SIGCOMM	Sniffer

Table 2: General information of papers.

The focus of the first research question **RQ1** is on identifying current approaches for dealing with privacy issues in IoT. The data shown in Figure 26 provides an overview of the literature per IoT domain, however, it does not include works that are specifically devoted to system literature reviews or the privacy paradox, for instance, only works that propose an approach to IoT privacy issues. Based on this data, it is shown that networking was the main topic of 25% of the studies included [86,106,111,112], while IoT system design [90,91] and mobile IoT applications [108,109] each accounted for 12.5% of the studies. Smart vehicles [61], IoT research [77], digital literacy [80], IoT regulations [84], smart homes [87], healthcare [88], AI [96] and smart cities [110] were the topics of 6.25% of the studies each.

There are two main ways to provide privacy in IoT systems, through security or providing in some way user awareness like the in the case of using privacy notices, other ways like through legislation or with the creation/usage of a framework or architecture that provides privacy mainly fall into one these two categories, for instance, Weber [81] argued for better and more precise

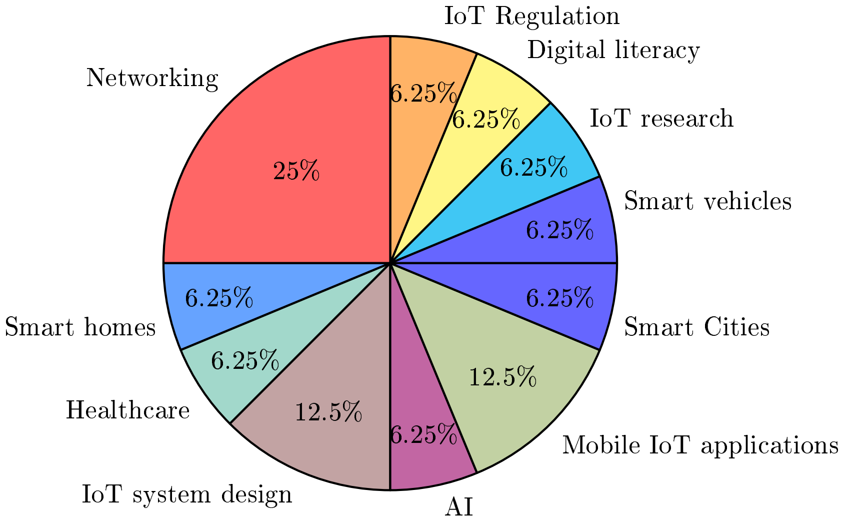


Fig. 26: Distribution of literature per IoT domain.

regulations and later came regulations like the GDPR and CCPA in the EU and USA, respectively, and with this more privacy notices appeared and more security rules were imposed by organisations. Literature that address any AI field [61, 88, 96] or blockchain [106] also fall under privacy through security. Most of the literature assumes that security and privacy are synonyms, for example [82,86,90], and so most of the proposed solutions fall under privacy through security. The proposed solutions that use privacy notices, like [108], are implemented in a way that use other devices like smartphones that provide the notices themselves, it is hard to provide privacy notices on the IoT devices themselves because many of these devices do not have a screen or the screen is too small to provide the necessary information to the user. Because there are still no standards for implementing privacy notices, and best practices are scattered throughout the literature, they are mostly implemented haphazardly, little guidance is given to designers and developers on how to make a privacy notice design that is sufficient and acceptable for their particular system and its features. Designers may be unaware of the numerous possibilities for creating acceptable privacy notifications and, as a result, do not systematically explore them.

Aleisa and Renaud [125] also identify security and privacy awareness as potential solutions to privacy issues in IoT, but also identify data minimization, hitchhiking and introspection. Data minimization entails limiting the collecting of personal information to what is absolutely central and retaining the data just for as long as is required to satisfy the goal of the technology's services [126]. Hitchhiking [127] is a method of protecting the privacy of users who divulge their location, applications regard locations as the object of their attention. The fidelity trade-off is removed as it is not important to know who is in a certain location. The introspection [128] method examines Virtual Machine (VM) actions to adequately safeguard users' private information. Every VM's CPU status, memory contents, network information provided by the hypervisor, and any malicious software that may be present on the VM are all collected and analysed. The privacy of individuals is jeopardized if an IoT device loses integrity due to a hostile assault.

Despite the fact that IoT usage has increased over the previous decade and the number of research publications has also increased, there is still a lack of focus on privacy as opposed to security or artificial intelligence, for example. IoT poses certain privacy concerns that are difficult to handle, and as a result, achieving certain goals will necessitate further research and financing.

The second research question **RQ2** is concerned with detecting frequent challenges that make addressing IoT privacy difficult. IoT complexity of architectures, applications and technologies make it hard to address these problems. The cacophony of networks make interoperability hard to achieve between IoT without the use of intermediaries. The challenges identified by both Qu et al. [113] and Ranjan et al. [115] are also hard to address and have not yet been elegantly overcome. The receptiveness of private organizations to embrace better privacy procedures can also be a challenge, due to the capitalistic nature of the majority of the world's companies. Many organizations in the IT area make most of their profits from advertising, or otherwise from user's private data, as such they are incentivized to keep harvesting data, and even with regulations in place and privacy frameworks that safeguard personal data it might make financial sense for most organizations to not invest in privacy or security, as paying fines for data breaches does not disrupt their operations. Public organizations might behave differently.

5.2 Study analysis and an application of privacy literacy

To answer the research question **RQ3**, the questionnaire makes it clear that there is a general lack of digital literacy, especially when it comes to IoT. This still being a new terminology/technology and only quickly expanding on the last decade, the people that have the most knowledge are the ones working in areas related to it. This survey also helps to demystify the privacy paradox.

Participants seemed eager to learn more about IoT, many had no knowledge of the term, even if some of them knew some devices that belong to the IoT. Some participants took the time to search online about terms that appeared on the questionnaire to get some idea of what they are about, and spend some time discussing them.

The usability tests were very important in improving certain aspects of the application. As was mentioned in the previous chapter, when the first usability tests were being conducted, the application was only available in english, but this soon changed and a portuguese translation was added. Other improvements added based on participant's feedback was a button to add an IoT device in the page where is possible to see all available devices. The icons were also altered, because participants complained that they were not sure how to read them, even after using for some time.

As concluded in the survey and users do not have a great deal of literacy regarding IoT, they do have some privacy literacy though, so to answer the research question **RQ4**, this application servers to increase users privacy literacy by giving them tools to know what kind of devices are around them, what these devices do and give them more information regarding IoT in general and IoT privacy, so that users may make well-informed choices.

The application by itself does not provide any formal privacy protection on IoT devices, but users can use it to better their understanding of IoT and in some cases make privacy choices regarding a Iot device.

One problem this application faces, and other applications where there is some kind of user interaction also face, is the fact that some bad actors will abuse the system by creating many fake accounts, by data scrapping, by adding bogus data or by replacing existing information with bogus data.

Some level of theoretical saturation [129] was reached with the use of the questionnaire and the usability tests, i.e. it was extracted the most amount of information possible from the participants on this topic of privacy on IoT systems, since when doing the usability tests most participants also completed the questionnaire.

This application does not claim to be the best in addressing IoT privacy issues but it improves on the previous works by not only focusing on privacy choices but also providing privacy literacy to the end user, which is the ultimate goal of the application. While existing works are fragmented on their approach, this application offers a more centralized way by allowing users to complete all tasks using it. Users can also contribute to the application itself through various methods, the most straightforward way is using it and adding new IoT devices to the application's database, they can also leave feedback for improving the application on distribution platforms. Because this application is distributed as free and open source software, it is possible for users to the development of the application itself by contributing with code or raising questions about features or bugs, it also can demystify privacy concerns about the application itself because anyone can inspect the source code.

6 Challenges

One of the most difficult points to accomplish in this thesis was the questionnaire, not the construction of the questionnaire itself but getting participants. Besides being difficult in itself to get a relatively high number of participants (a few hundred at least) to be able to draw conclusions with any high degree of confidence, it was difficult to get the potential participants interested in the topic at hand, because although it seems that many people value their privacy very highly and think they should protect it, in practice it seems they are not very interested. This may even be because many people do not have much knowledge about the Internet of Things, and thus feel that they cannot answer the questionnaire because it is out of their field of knowledge, another reason may be that the questionnaire seems a little long, because it takes on average 15 to 20 minutes to answer, and despite being a topic of interest the time investment in the questionnaire may be considered too high. Another point to take into consideration regarding the low number of participants is the way the questionnaire is written and how it was advertised, i.e., a very formal or technical language may have been used both in the construction of the questionnaire and in its dissemination, and the fact that this is a very niche topic may have "scared" possible participants. However, it should be noted that also in the literature that has been carried out there is not a great focus on conducting questionnaires and the ones that have been conducted have not only focused on the Internet of Things and also have some monetary incentive for the participants.

7 Future work

Although there are existing hardware solutions that can detect some devices on particular networks, like ZigBee or Bluetooth LE, namely IoT sniffers and there exist some georeferencing applications that try to pinpoint certain IoT devices, there is still a need for some kind of device or framework that is network agnostic and can detect where the devices are located and what kind of data the IoT devices that are around it are collecting. This gadget should also be capable of informing users about the privacy notices of the devices and what can the users do to safeguard their personal data. The IoT sniffers that are available are primarily used in the detection of problems in the communication of devices in the network or to solve problems of interoperability between different IoT networks. There are many obstacles that impede the creation of such a device and the fact that it still does not exist anything like it may be related ton either there iss not enough interest from users or researchers to focus on such an endeavour or the complexity of such a task is greater than the rewards.

8 Conclusion

This project aims to do an exploratory analysis of privacy in IoT systems. It proposes a survey to better understand user's knowledge on this subject and an application that aims to create more users awareness and better inform about their environment, as well as the IoT devices that inhabit it and how they can respond accordingly.

Hopefully the work conducted on this project will be useful to further support researchers and the application that will be developed will be able to provide greater visibility, thus allowing users to acquire knowledge about the data being collected and how they can adjust their behavior or respond more effectively to protect their privacy rights.

References

[1] D. Vincent, *Privacy: A short history.* John Wiley & Sons, 2016.

[2] B. Moore, *Privacy: Studies in social and cultural history.* Routledge, 2017.

[3] E. Roosevelt, P. C. Chang, C. Malik, W. R. Hodgson, H. S. Cruz, R. Cassin, A. E. Bogomolov, C. D. 1st Baron Dukeston, and J. P. Humphrey. (1948) Universal declaration of human rights. Accessed: 2022-11-05. [Online]. Available: https://www.un.org/en/about-us/universal-declaration-of-human-rights

[4] Constitute. (2023) Constitutions - constitute. Accessed: 2023-08-22. [Online]. Available: https://www.constituteproject.org/constitutions?lang=en&q=privacy&status=in_force&status=is_draft

[5] I. A. for Privacy Professionals. (2021) What does privacy mean? Accessed: 2022-11-04. [Online]. Available: https://iapp.org/about/what-is-privacy/

[6] S. Spiekermann and L. F. Cranor, "Engineering privacy," *IEEE Transactions on Software Engineering*, vol. 35, no. 1, pp. 67–82, 2009.

[7] HIV.gov. (2018) The difference between security and privacy and why it matters to your program. Accessed: 2022-11-04. [Online]. Available: https://www.hiv.gov/blog/difference-between-security-and-privacy-and-why-it-matters-your-program

[8] P. Emami-Naeini, H. Dixon, Y. Agarwal, and L. F. Cranor, "Exploring how privacy and security factor into iot device purchase behavior," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[9] Y. J. Park, "Personal data concern, behavioral puzzle and uncertainty in the age of digital surveillance," *Telematics and Informatics*, vol. 66, p. 101748, 2022.

[10] N. A. Zhang, C. A. Wang, E. Karahanna, and Y. Xu, "Peer privacy concern: Conceptualization and measurement." *MIS Quarterly*, vol. 46, no. 1, 2022.

[11] Y. Duan, P. Liu, and Y. Feng, "Pricing strategies of two-sided platforms considering privacy concerns," *Journal of Retailing and Consumer Services*, vol. 64, p. 102781, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0969698921003477

[12] M. Weiser, "The computer for the 21 st century," *Scientific american*, vol. 265, no. 3, pp. 94–105, 1991.

[13] K. Ashton. (2009) That "internet of things" thing. Accessed: 2022-11-05. [Online]. Available: https://www.rfidjournal.com/that-internet-of-things-thing

[14] D. Marikyan, S. Papagiannidis, and E. Alamanos, "A systematic review of the smart home literature: A user perspective," *Technological Forecasting and Social Change*, vol. 138, pp. 139–154, 2019.

[15] F. Arena, G. Pau, and A. Severino, "An overview on the current status and future perspectives of smart cars," *Infrastructures*, vol. 5, no. 7, p. 53, 2020.

[16] M. Al-Emran, S. I. Malik, and M. N. Al-Kabi, "A survey of internet of things (iot) in education: Opportunities and challenges," *Toward social internet of things (SIoT): enabling technologies, architectures and applications*, pp. 197–209, 2020.

[17] N. Niknejad, W. B. Ismail, A. Mardani, H. Liao, and I. Ghani, "A comprehensive overview of smart wearables: The state of the art literature, recent advances, and future challenges," *Engineering Applications of Artificial Intelligence*, vol. 90, p. 103529, 2020.

[18] I. F. Akyildiz and A. Kak, "The internet of space things/cubesats," *IEEE Network*, vol. 33, no. 5, pp. 212–218, 2019.

[19] S. Madakam, V. Lake, V. Lake, V. Lake *et al.*, "Internet of things (iot): A literature review," *Journal of Computer and Communications*, vol. 3, no. 05, p. 164, 2015.

[20] C. Everhart, E. Caplan, and D. Nichols, "Re: Interesting uses of networking," https://cseweb.ucsd.edu/~bsy/coke.history.txt, 1990, accessed: 2022-12-14.

[21] J. Romkey, "Toast of the iot: The 1990 interop internet toaster," *IEEE Consumer Electronics Magazine*, vol. 6, no. 1, pp. 116–119, 2017.

[22] M. Hasan. (2022) State of iot 2022: Number of connected iot devices growing 18% to 14.4 billion globally. Accessed: 2022-11-04. [Online]. Available: https://iot-analytics.com/number-connected-iot-devices/

[23] S. Al-Sarawi, M. Anbar, R. Abdullah, and A. B. Al Hawari, "Internet of things market analysis forecasts, 2020-2030," in *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, 2020, pp. 449–453.

[24] L. J. Hoffman, "Computers and privacy: A survey," *ACM Computing Surveys (CSUR)*, vol. 1, no. 2, pp. 85–103, 1969.

[25] E. US Department of Health and Welfare, "Records, computers, and the rights of citizens: Report of the secretary's advisory committee on automated personal data systems," US Government Printing Office Washington, DC, 1973.

[26] U. S. P. P. S. Commission, *Personal Privacy in an Information Society: The Report of the Privacy Protection Study Commission*. The Commission, 1977, vol. 2.

[27] N. Alhirabi, O. Rana, and C. Perera, "Security and privacy requirements for the internet of things: A survey," *ACM Transactions on Internet of Things*, vol. 2, no. 1, pp. 1–37, 2021.

[28] D. Peloquin, M. DiMaio, B. Bierer, and M. Barnes, "Disruptive and avoidable: Gdpr challenges to secondary research uses of data," *European Journal of Human Genetics*, vol. 28, no. 6, pp. 697–705, 2020.

[29] A. Gladis, N. J. Hartwich, and O. Salge, "Weaponizing the gdpr: How flawed implementations turn the gold standard for privacy laws into fool's gold," 2022.

[30] G. Gentile and O. Lynskey, "Deficient by design? the transnational enforcement of the gdpr," *International & Comparative Law Quarterly*, vol. 71, no. 4, pp. 799–830, 2022.

[31] B. Green, "The flaws of policies requiring human oversight of government algorithms," *Computer Law & Security Review*, vol. 45, p. 105681, 2022.

[32] D. Y. Byun, “Privacy or protection: The catch-22 of the ccpa,” *Loy. Consumer L. Rev.*, vol. 32, p. 246, 2019.

[33] J. Caltrider. (2017) 10 fascinating things we learned when we asked the world ‘how connected are you?’. Accessed: 2023-09-05. [Online]. Available: https://blog.mozilla.org/en/mozilla/10-fascinating-things-we-learned-when-we-asked-the-world-how-connected-are-you/

[34] H. Khan and P. Singh, “Issues and challenges of internet of things: A survey,” *Journal of Informatics Electrical and Electronics Engineering (JIEEE)*, vol. 2, no. 3, pp. 1–8, Nov. 2021. [Online]. Available: https://jieee.a2zjournals.com/index.php/ieee/article/view/17

[35] A. K. Mohammadzadeh, S. Ghafoori, A. Mohammadian, R. Mohammadkazemi, B. Mahbanooei, and R. Ghasemi, “A fuzzy analytic network process (fanp) approach for prioritizing internet of things challenges in iran,” *Technology in Society*, vol. 53, pp. 124–134, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160791X1730194X

[36] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, “Industrial internet of things: Challenges, opportunities, and directions,” *IEEE transactions on industrial informatics*, vol. 14, no. 11, pp. 4724–4734, 2018.

[37] D. Thomson, D. P. Cochrane, I. Chantzos, P. Carter, S. John, P. U. Helmbrecht, and S. Room, “State of privacy report 2015,” Symantec, Tech. Rep., 2015.

[38] D. J. Solove, “The myth of the privacy paradox,” *Geo. Wash. L. Rev.*, vol. 89, p. 1, 2021.

[39] M. Williams, J. R. C. Nurse, and S. Creese, “Privacy is the boring bit: User perceptions and behaviour in the internet-of-things,” in *2017 15th Annual Conference on Privacy, Security and Trust (PST)*, 2017, pp. 181–18 109.

[40] A.-R. Lee, “Investigating the personalization–privacy paradox in internet of things (iot) based on dual-factor theory: Moderating effects of type of iot service and user value,” *Sustainability*, vol. 13, no. 19, p. 10679, 2021.

[41] D. Goad, A. T. Collins, and U. Gal, “Privacy and the internet of things- an experiment in discrete choice,” *Information & Management*, vol. 58, no. 2, p. 103292, 2021.

[42] N. Gerber, P. Gerber, and M. Volkamer, “Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior,” *Computers & security*, vol. 77, pp. 226–261, 2018.

[43] D. Wilson and J. S. Valacich, “Unpacking the privacy paradox: Irrational decision-making within the privacy calculus,” 2012.

[44] J. Warshaw, T. Matthews, S. Whittaker, C. Kau, M. Bengualid, and B. A. Smith, “Can an algorithm know the" real you"? understanding people’s reactions to hyper-personal analytics systems,” in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 797–806.

[45] N. Lee and O. Kwon, “A privacy-aware feature selection method for solving the personalization–privacy paradox in mobile wellness healthcare services,” *Expert systems with applications*, vol. 42, no. 5, pp. 2764–2771, 2015.

[46] P. J. Zak, *Moral markets: The critical role of values in the economy.* Princeton University Press, 2008.

[47] B. P. Knijnenburg, A. Kobsa, and H. Jin, "Dimensionality of information disclosure behavior," *International Journal of Human-Computer Studies*, vol. 71, no. 12, pp. 1144–1162, 2013.

[48] R. Wakefield, "The influence of user affect in online information disclosure," *The Journal of Strategic Information Systems*, vol. 22, no. 2, pp. 157–174, 2013.

[49] C. Flender and G. Müller, "Type indeterminacy in privacy decisions: the privacy paradox revisited," in *International Symposium on Quantum Interaction.* Springer, 2012, pp. 148–159.

[50] T. Dienlin and S. Trepte, "Is the privacy paradox a relic of the past? an in-depth analysis of privacy attitudes and privacy behaviors," *European journal of social psychology*, vol. 45, no. 3, pp. 285–297, 2015.

[51] Y. M. Baek, "Solving the privacy paradox: A counter-argument experimental approach," *Computers in human behavior*, vol. 38, pp. 33–42, 2014.

[52] M. Taddicken, "The 'Privacy Paradox' in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure*," *Journal of Computer-Mediated Communication*, vol. 19, no. 2, pp. 248–273, 01 2014. [Online]. Available: https://doi.org/10.1111/jcc4.12052

[53] P. A. Norberg, D. R. Horne, and D. A. Horne, "The privacy paradox: Personal information disclosure intentions versus behaviors," *Journal of consumer affairs*, vol. 41, no. 1, pp. 100–126, 2007.

[54] S. Kokolakis, "Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon," *Computers & security*, vol. 64, pp. 122–134, 2017.

[55] L. Brandimarte, A. Acquisti, and G. Loewenstein, "Misplaced confidences: Privacy and the control paradox," *Social psychological and personality science*, vol. 4, no. 3, pp. 340–347, 2013.

[56] W. Xie and K. Karan, "Consumers' privacy concern and privacy protection on social network sites in the era of big data: empirical evidence from college students," *Journal of Interactive Advertising*, vol. 19, no. 3, pp. 187–201, 2019.

[57] K. S. Schwaig, A. H. Segars, V. Grover, and K. D. Fiedler, "A model of consumers' perceptions of the invasion of information privacy," *Information & Management*, vol. 50, no. 1, pp. 1–12, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378720612000754

[58] S. Sannon, N. N. Bazarova, and D. Cosley, "Privacy lies: Understanding how, when, and why people lie to protect their privacy in multiple online contexts," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–13.

[59] M. Kearns and A. Roth, *The ethical algorithm: The science of socially aware algorithm design.* Oxford University Press, 2019.

[60] Y. Zhao and J. Chen, "A survey on differential privacy for unstructured data content," *ACM Comput. Surv.*, vol. 54, no. 10s, sep 2022. [Online]. Available: https://doi.org/10.1145/3490237

[61] Y. Zhao, J. Zhao, M. Yang, T. Wang, N. Wang, L. Lyu, D. Niyato, and K.-Y. Lam, "Local differential privacy-based federated learning for internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8836–8853, 2020.

[62] A. Frome, G. Cheung, A. Abdulkader, M. Zennaro, B. Wu, A. Bissacco, H. Adam, H. Neven, and L. Vincent, "Large-scale privacy protection in google street view," in *IEEE International Conference on Computer Vision*, 2009. [Online]. Available: https://research.google.com/archive/papers/cbprivacy_iccv09.pdf

[63] R. Poddar, G. Ananthanarayanan, S. Setty, S. Volos, and R. A. Popa, "Visor: Privacy-preserving video analytics as a cloud service," in *USENIX Security Symposium*, August 2020. [Online]. Available: https://www.microsoft.com/en-us/research/publication/visor-privacy-preserving-video-analytics-as-a-cloud-service/

[64] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2015, pp. 909–910.

[65] S. Gupta and S. Ghanavati, "Privacy in the internet of things: Where do we stand? a systematic literature review," 5 2022. [Online]. Available: https://www.techrxiv.org/articles/preprint/Privacy_in_the_Internet_of_Things_Where_do_We_Stand_A_Systematic_Literature_Review/19874329

[66] P. Kühtreiber, V. Pak, and D. Reinhardt, "A survey on solutions to support developers in privacy-preserving iot development," *Pervasive and Mobile Computing*, vol. 85, p. 101656, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574119222000785

[67] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in internet of things: The road ahead," *Computer networks*, vol. 76, pp. 146–164, 2015.

[68] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1125–1142, 2017.

[69] G. Yang, "An overview of current solutions for privacy in the internet of things," *Frontiers in Artificial Intelligence*, vol. 5, p. 812732, 2022.

[70] H. D. Zubaydi, P. Varga, and S. Molnár, "Leveraging blockchain technology for ensuring security and privacy aspects in internet of things: A systematic literature review," *Sensors*, vol. 23, no. 2, p. 788, 2023.

[71] J. Porras, J. Pänkäläinen, A. Knutas, and J. Khakurel, "Security in the internet of things-a systematic mapping study," 2018.

[72] B. S. Ahmed, M. Bures, K. Frajtak, and T. Cerny, "Aspects of quality in internet of things (iot) solutions: A systematic mapping study," *IEEE Access*, vol. 7, pp. 13 758–13 780, 2019.

[73] A. Khanna and S. Kaur, "Internet of things (iot), applications and challenges: a comprehensive review," *Wireless Personal Communications*, vol. 114, pp. 1687–1762, 2020.

[74] S. G. Tzafestas, "Ethics and law in the internet of things world," *Smart cities*, vol. 1, no. 1, pp. 98–120, 2018.

[75] J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the internet of things: threats and challenges," *Security and Communication Networks*, vol. 7, no. 12, pp. 2728–2742, 2014.

[76] P. E. Naeini, S. Bhagavatula, H. Habib, M. Degeling, L. Bauer, L. F. Cranor, and N. Sadeh, "Privacy expectations and preferences in an {IoT} world," in *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, 2017, pp. 399–412.

[77] A. Koohang, C. S. Sargent, J. H. Nord, and J. Paliszkiewicz, "Internet of things (iot): From awareness to continued use," *International Journal of Information Management*, vol. 62, p. 102442, 2022.

[78] M. Tsourela and D.-M. Nerantzaki, "An internet of things (iot) acceptance model. assessing consumer's behavior toward iot products and applications," *Future Internet*, vol. 12, no. 11, p. 191, 2020.

[79] B. P. Knijnenburg, X. Page, P. Wisniewski, H. R. Lipford, N. Proferes, and J. Romano, *Modern socio-technical perspectives on privacy.* Springer Nature, 2022.

[80] M. Skirpan, M. Oates, D. Byrne, R. Cunningham, and L. F. Cranor, "Is a privacy crisis experienced, a privacy crisis avoided?" *Commun. ACM*, vol. 65, no. 3, pp. 26–29, feb 2022. [Online]. Available: https://doi.org/10.1145/3512325

[81] R. H. Weber, "Internet of things: Privacy issues revisited," *Computer Law & Security Review*, vol. 31, no. 5, pp. 618–627, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0267364915001156

[82] N. Fabiano, "Internet of things and blockchain: Legal issues and privacy. the challenge for a privacy standard," in *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 2017, pp. 727–734.

[83] R. H. Weber, "Internet of things–new security and privacy challenges," *Computer law & security review*, vol. 26, no. 1, pp. 23–30, 2010.

[84] S. Hadzovic, S. Mrdovic, and M. Radonjic, "A path towards an internet of things and artificial intelligence regulatory framework," *IEEE Communications Magazine*, 2023.

[85] E. Commission. (2021) Europe fit for the digital age: Commission proposes new rules and actions for excellence and trust in artificial intelligence. Accessed: 2023-09-05. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/IP_21_1682

[86] Q. Sun, K. Lin, C. Si, Y. Xu, S. Li, and P. Gope, "A secure and anonymous communicate scheme over the internet of things," *ACM Trans. Sen. Netw.*, vol. 18, no. 3, apr 2022. [Online]. Available: https://doi.org/10.1145/3508392

[87] S. Xiong, A. D. Sarwate, and N. B. Mandayam, "Defending against packet-size side-channel attacks in iot networks," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2018, pp. 2027–2031.

[88] R. S. Antunes, C. André da Costa, A. Küderle, I. A. Yari, and B. Eskofier, "Federated learning for healthcare: Systematic review and architecture proposal," *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 4, may 2022. [Online]. Available: https://doi.org/10.1145/3501813

[89] B. Chen, S. Qiao, J. Zhao, D. Liu, X. Shi, M. Lyu, H. Chen, H. Lu, and Y. Zhai, "A security awareness and protection system for 5g smart healthcare based on zero-trust architecture," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10 248–10 263, 2021.

[90] A. Opara, H. Johng, T. Hill, and L. Chung, "A framework for representing internet of things security and privacy policies and detecting potential problems," in *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*, 2022, pp. 198–201.

[91] C. Perera, M. Barhamgi, A. K. Bandara, M. Ajmal, B. Price, and B. Nuseibeh, "Designing privacy-aware internet of things applications," *Information Sciences*, vol. 512, pp. 238–257, 2020.

[92] Y. O'Connor, W. Rowan, L. Lynch, and C. Heavin, "Privacy by design: informed consent and internet of things for smart health," *Procedia computer science*, vol. 113, pp. 653–658, 2017.

[93] A. Cavoukian and C. Popa, "Embedding privacy into what's next: Privacy by design for the internet of things," *Ryerson University Privacy & Big Data Institute*, pp. 1–10, 2016.

[94] L. Alkhariji, S. De, O. Rana, and C. Perera, "Semantics-based privacy by design for internet of things applications," *Future Generation Computer Systems*, vol. 138, pp. 280–295, 2023.

[95] A. Aljeraisy, M. Barati, O. Rana, and C. Perera, "Privacy laws and privacy by design schemes for the internet of things: A developer's perspective," *ACM Computing Surveys (Csur)*, vol. 54, no. 5, pp. 1–38, 2021.

[96] Q. Zhang, L. T. Yang, Z. Chen, P. Li, and M. J. Deen, "Privacy-preserving double-projection deep computation model with crowdsourcing on cloud for big data feature learning," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2896–2903, 2017.

[97] M. Barrett, "Framework for improving critical infrastructure cybersecurity version 1.1," Apr. 2018.

[98] "Cybersecurity – IoT security and privacy – Guidelines," International Organization for Standardization, Geneva, CH, Standard ISO/IEC 27400:2022, Jun. 2022. [Online]. Available: https://www.iso.org/standard/44373.html

[99] ISO Central Secretary. (2023) ISO - search: I'm looking for privacy. Accessed: 2023-08-22. [Online]. Available: https://www.iso.org/search.html?q=privacy&hPP=10&idx=all_en&p=0&hFR%5Bcategory%5D%5B0%5D=standard

[100] ——. (2023) ISO - search: I'm looking for security. Accessed: 2023-08-22. [Online]. Available: https://www.iso.org/search.html?q=security&hPP=10&idx=all_en&p=0&hFR%5Bcategory%5D%5B0%5D=standard

[101] X. Sun, F. R. Yu, P. Zhang, Z. Sun, W. Xie, and X. Peng, "A survey on zero-knowledge proof in blockchain," *IEEE network*, vol. 35, no. 4, pp. 198–205, 2021.

[102] R. Mercer, "Privacy on the blockchain: Unique ring signatures," *arXiv preprint arXiv:1612.01188*, 2016.

[103] D. Stone, “Trustless, privacy-preserving blockchain bridges,” *arXiv preprint arXiv:2102.04660*, 2021.

[104] R. Zhang, R. Xue, and L. Liu, “Security and privacy on blockchain,” *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, pp. 1–34, 2019.

[105] Y. Yu, Y. Li, J. Tian, and J. Liu, “Blockchain-based solutions to security and privacy issues in the internet of things,” *IEEE Wireless Communications*, vol. 25, no. 6, pp. 12–18, 2018.

[106] M. S. Ali, K. Dolui, and F. Antonelli, “Iot data privacy via blockchains and ipfs,” in *Proceedings of the Seventh International Conference on the Internet of Things*, ser. IoT ’17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3131542.3131563

[107] J. Colnago, Y. Feng, T. Palanivel, S. Pearman, M. Ung, A. Acquisti, L. F. Cranor, and N. Sadeh, “Informing the design of a personalized privacy assistant for the internet of things,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–13. [Online]. Available: https://doi.org/10.1145/3313831.3376389

[108] Y. Feng, Y. Yao, and N. Sadeh, “A design space for privacy choices: Towards meaningful privacy control in the internet of things,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445148

[109] A. Das, M. Degeling, D. Smullen, and N. Sadeh, “Personalized privacy assistants for the internet of things: Providing users with notice and choice,” *IEEE Pervasive Computing*, vol. 17, pp. 35–46, 07 2018.

[110] H. Zhu, S. C.-K. Chau, G. Guarddin, and W. Liang, “Integrating iot-sensing and crowdsensing with privacy: Privacy-preserving hybrid sensing for smart cities,” *ACM Trans. Internet Things*, vol. 3, no. 4, sep 2022. [Online]. Available: https://doi.org/10.1145/3549550

[111] J. Lola, C. Serrão, and J. Casal, “Towards transparent and secure iot: Improving the security and privacy through a user-centric rules-based system,” *Electronics*, vol. 12, no. 12, 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/12/2589

[112] S. Kumar, E. Hamed, D. Katabi, and L. Erran Li, “Lte radio analytics made easy and accessible,” in *Proceedings of the 2014 ACM Conference on SIGCOMM*, ser. SIGCOMM ’14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 211–222. [Online]. Available: https://doi.org/10.1145/2619239.2626320

[113] Y. Qu, S. Yu, W. Zhou, S. Peng, G. Wang, and K. Xiao, “Privacy of things: Emerging challenges and opportunities in wireless internet of things,” *IEEE Wireless Communications*, vol. 25, no. 6, pp. 91–97, 2018.

[114] R. H. Petrucci, F. G. Herring, J. D. Madura, and C. Bissonnette, *General Chemistry: Modern Principles and Applications, 12th edition.* Pearson, 2023.

[115] C. Perera, R. Ranjan, L. Wang, S. U. Khan, and A. Y. Zomaya, “Big data privacy in the internet of things era,” *IT professional*, vol. 17, no. 3, pp. 32–39, 2015.

[116] S. Keshav, "How to read a paper," *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 3, pp. 83–84, jul 2007. [Online]. Available: https://doi.org/10.1145/1273445.1273458

[117] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The prisma 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, 2021. [Online]. Available: https://www.bmj.com/content/372/bmj.n71

[118] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," vol. 2, 01 2007.

[119] P. K. Masur, *Situational privacy and self-disclosure: Communication processes in online environments*. Springer, 2018.

[120] P. K. Masur, D. Teutsch, and S. Trepte, "Entwicklung und validierung der online-privatheitskompetenzskala (oplis)," *Diagnostica*, 2017.

[121] P. Survey and R. Center. (2022) Conduct of privacy survey. Accessed: 2022-12-19. [Online]. Available: https://www.privacy.gov.ph/wp-content/uploads/2022/01/CONDUCT-OF-PRIVACY-SURVEY-Final-Report-v3.pdf

[122] M. H. d. S. Alves, "Gdpr in portugal: Analysis of citizens' perception about privacy," 2021.

[123] M. Dadkhah, M. Lagzan, F. Rahimnia, and K. Kimiafar, "What do publications say about the internet of things challenges/barriers to uninformed authors?: a bibliometric analysis," *What do Publications say about the Internet of Things Challenges/Barriers to uninformed Authors?: a bibliometric Analysis*, pp. 77–98, 2020.

[124] A. Acquisti, S. Gritzalis, C. Lambrinoudakis, and S. di Vimercati, "What can behavioral economics teach us about privacy?" in *Digital privacy*. Auerbach Publications, 2007, pp. 385–400.

[125] N. Aleisa and K. Renaud, "Privacy of the internet of things: a systematic literature review (extended discussion)," *arXiv preprint arXiv:1611.03340*, 2016.

[126] E. D. P. Supervisor, "Directive 95/46/ec of the european parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data," Official Journal of the European Union, 1995.

[127] K. P. Tang, P. Keyani, J. Fogarty, and J. I. Hong, "Putting people in their place: an anonymous and privacy-sensitive approach to collecting sensed data in location-based applications," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2006, pp. 93–102.

[128] C. Kang, F. Abbas, and H. Oh, "Protection scheme for iot devices using introspection," in *2015 6th International Conference on the Network of the Future (NOF)*. IEEE, 2015, pp. 1–5.

[129] J. Low, "A pragmatic definition of the concept of theoretical saturation," *Sociological Focus*, vol. 52, no. 2, pp. 131–139, 2019.

Appendix

Survey

Section 1: General knowledge and attitudes towards privacy

The purpose of this section is to ask general questions about your knowledge in the area of information privacy.

1. "Privacy is important to me". Do you agree with this statement?

	1	2	3	4	5	6	7	
Disagree	○	○	○	○	○	○	○	Agree

2. "I know techniques to guarantee privacy and the protection of my data when I use the Internet". Do you agree with this statement?

	1	2	3	4	5	6	7	
Disagree	○	○	○	○	○	○	○	Agree

3. Could you give examples of techniques or strategies you use to guarantee the privacy and protection of your data?

Short answer...

4. Do you think your data privacy is important?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very important

5. How do you define data (or digital) privacy?

Long answer...

6. Do you think your data privacy is relevant nowadays?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very relevant

7. "Data privacy is a human right". Do you agree with this statement?

○ I agree

○ I disagree

○ I do not know

8. "Data privacy is a consumer right". Do you agree with this statement?

○ I agree

○ I disagree

○ I do not know

9. In your opinion, are security and privacy synonymous?

○ Yes

○ No

10. How familiar are you with the following terms?

	I Know well	I have some knowledge	I do not know
Cookie	○	○	○
Wireless networks	○	○	○
Wi-Fi	○	○	○
Internet of Things	○	○	○
Data protection	○	○	○
Ad hoc network (wireless)	○	○	○
Ubiquitous computing	○	○	○
Privacy assistant	○	○	○
Blockchain	○	○	○
Differential privacy	○	○	○

11. Do you think your data is stored/collected when you use the internet?

- ○ Yes
- ○ No
- ○ I do not know

12. What means do you think are used to collect your data?

- ☐ Sound (microphones, recorders, etc.)
- ☐ Image (webcams, scanners, etc.)
- ☐ Online Surveys / Questionnaires
- ☐ Hackers (illegal means)
- ☐ Cookies
- ☐ None
- ☐ Another option...

13. What kind of information can be extracted using the collected data?

- ☐ Visited websites information
- ☐ Work productivity
- ☐ Georeferencing data
- ☐ Online habits
- ☐ Online shoppping
- ☐ Another option...

14. For what purposes can organizations use your data?

- ☐ Improve internet usage
- ☐ Costumer profile
- ☐ Targeted advertising
- ☐ Advertising adapted to the nearest locations
- ☐ Selling data to third parties
- ☐ Another option...

15. Which organizations do you consider to have the best privacy behaviour?

- ☐ Amazon
- ☐ Apple
- ☐ Google
- ☐ Meta (Facebook, Whatsapp, Instagram)
- ☐ Microsoft
- ☐ Samsung
- ☐ IBM
- ☐ ScienceSoft
- ☐ PTC
- ☐ Cisco
- ☐ Huawei
- ☐ GE Digital
- ☐ Bosch
- ☐ Siemens
- ☐ Other...

16. Which organizations do you consider to have the worst privacy behaviour?

- ☐ Amazon
- ☐ Apple
- ☐ Google
- ☐ Meta (Facebook, Whatsapp, Instagram)
- ☐ Microsoft
- ☐ Samsung
- ☐ IBM
- ☐ ScienceSoft
- ☐ PTC
- ☐ Cisco
- ☐ Huawei
- ☐ GE Digital
- ☐ Bosch
- ☐ Siemens
- ☐ Other...

17. During your day-to-day life, how often do you use your phone to access the internet?

	1	2	3	4	5	6	7	
Never	○	○	○	○	○	○	○	Very often

18. "I am concerned about my privacy when using my mobile phone when accessing the internet". Do you agree with this statement?

	1	2	3	4	5	6	7	
Disagree	○	○	○	○	○	○	○	Agree

19. "I consider that accessing the internet through my phone is safer than through a computer". Do you agree with this statement?

	1	2	3	4	5	6	7	
Disagree	○	○	○	○	○	○	○	Agree

20. "I try to block the collection of data from applications installed on my phone". Do you agree with this statement?

	1	2	3	4	5	6	7	
Disagree	○	○	○	○	○	○	○	Agree

21. When using a website, do you usually read the Privacy Policy?

○	Yes, always
○	Yes, almost always
○	No, because I am not interested
○	No, but I know what it means or what happens

22. How often do you allow the use of cookies?

	1	2	3	4	5	6	7	
Never	○	○	○	○	○	○	○	Always

23. When you accept / become aware of the cookies policy, why do you take this decision?

- ○ It is the only way to access the website
- ○ I understand what it is about and I agree with the policy presented
- ○ I don't understand what this is about but it doesn't seem relevant to me
- ○ Other...

24. Are you aware of the concept of "profiling" or "automated processing of personal information"?

- ○ Yes
- ○ No

25. Do you consider that your internet activity contributes to the development of profiling?

- ○ Yes
- ○ No
- ○ I do not know

26. Are you aware of regulations such as the General Data Protection Regulation or the California Consumer Privacy Act?

- ○ Yes, I fully understand the regulations
- ○ Yes, but it's not absolutely clear to me what it represents
- ○ No

27. If you answered yes, how did you learn about the 'General Data Protection Regulation' or the 'California Consumer Privacy Act'?

- ☐ Internet
- ☐ Television
- ☐ Newspapers
- ☐ Friends
- ☐ Work
- ☐ Other...

28. Are you interested in finding out more about regulations or legislation related to digital privacy?

	1	2	3	4	5	6	7	
Not at all	O	O	O	O	O	O	O	Very much

29. According to the General Data Protection Regulation, personal data is:

O	Just name, email, date of birth and tax identification number
O	All of the above and bank details
O	All of the above and medical information
O	Any information relating to an identified or identifiable natural person

30. "My data was more protected after the implementation of regulations such as the General Data Protection Regulation". Do you agree with this statement?

	1	2	3	4	5	6	7	
Disagree	O	O	O	O	O	O	O	Agree

Section 2: Disposition for sharing personal information

This section is intended to ask questions to generally understand your willingness to share personal information.

1. What kind of personal information are you willing to share at any time?

- ☐ Age
- ☐ Complete name
- ☐ Date of birth
- ☐ Genre
- ☐ Religion
- ☐ Complete name of parents
- ☐ Marital status
- ☐ Home address
- ☐ Work address
- ☐ Nationality
- ☐ Qualifications
- ☐ Professional experience
- ☐ Height
- ☐ Weight
- ☐ Phone/mobile number
- ☐ Region of birth
- ☐ Ethnic group
- ☐ Signature
- ☐ Spouse's name
- ☐ Email
- ☐ Health condition
- ☐ Face photo
- ☐ Full body photography
- ☐ Username
- ☐ Copy of citizen card
- ☐ Fingerprint
- ☐ Licenses

2. In what situations are you willing to provide more personal information?

- ☐ Renew your citizen card
- ☐ Medical consultation (face to face)
- ☐ Take out health insurance
- ☐ Create new bank account
- ☐ For contact tracing
- ☐ Confined in the hospital
- ☐ Make a new Internet/Mobile plan
- ☐ Online shopping
- ☐ Apply for a bank loan or credit card
- ☐ Online medical consultation
- ☐ Request loyalty/discount cards
- ☐ None

3. Do you agree to share health data that can identify you with health professionals?

- ○ Yes, I agree
- ○ Maybe if asked first
- ○ No, I disagree
- ○ I do not know

4. Do you agree to share health data that cannot identify you with health professionals?

- ○ Yes, I agree
- ○ Maybe if asked first
- ○ No, I disagree
- ○ I do not know

5. What kind of applications do you have installed on your smartphone?

- ☐ Social media
- ☐ Instant messages
- ☐ Email
- ☐ Browser
- ☐ Navigation (ex. GPS)
- ☐ Anti-virus
- ☐ Online shopping
- ☐ Digital Wallet
- ☐ Photo/video editing
- ☐ Contact tracing
- ☐ Online banking
- ☐ Other...

6. Before sharing your data, do you consult any of the following information?

- ☐ Privacy policy
- ☐ Terms and conditions
- ☐ Purpose of data collection
- ☐ Consent Form
- ☐ Privacy notice
- ☐ Reliability of the organization/institution
- ☐ I do not consult any information

7. How often do you find privacy policies?

- ○ Almost everyday
- ○ Once a week
- ○ Once a month
- ○ Very rarely
- ○ Never

8. Are you aware of the duties of a Data Protection Officer (DPO)?

- ○ Yes
- ○ No

9. "I am interested in knowing where and how my personal information is used." Do you agree with this statement?

- ○ I agree
- ○ I do not agree nor disagree
- ○ I disagree

10. "I am not familiar with the purpose of data collection but would like to know more". Do you agree with this statement?

- ○ I agree
- ○ I do not agree nor disagree
- ○ I disagree

11. "The length (or number of words) of the privacy notice affects my willingness to read it." Do you agree with this statement?

- ○ I agree
- ○ I do not agree nor disagree
- ○ I disagree

12. "The font size of the privacy notice affects my willingness to read it." Do you agree with this statement?

- ○ I agree
- ○ I do not agree nor disagree
- ○ I disagree

13. "Usually, I'm afraid I won't be able to use a product or service if I don't agree with the privacy notice." Do you agree with this statement?

- ○ I agree
- ○ I do not agree nor disagree
- ○ I disagree

14. "I don't need to read the privacy notice if I trust the institution". Do you agree with this statement?

- ○ I agree
- ○ I do not agree nor disagree
- ○ I disagree

Section 3: Privacy concerns

This section aims to ask questions related to possible fears associated with sharing personal information.

1. How concerned are you about organizations collecting and using your online activity?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

2. How concerned are you about organizations sharing your data with third parties?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

3. How concerned are you about organizations tracking your online behavior and thus obtaining your personal data?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

4. How concerned are you about public institutions or intelligence services analyzing your online movements?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

5. How concerned are you that other people obtain your personal data without your consent?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

6. How concerned are you that other people find information about you online?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

7. How concerned are you that other people are disclosing information about you without your knowledge?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

8. How concerned are you about other people sharing your personal data (photos, address, mobile phone number, etc.) with others without your consent?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

9. How concerned are you that other people publish your personal data (photos, address, mobile phone number, etc.) on the internet without your consent?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

10. How concerned are you about an unknown person claiming to be you on the internet?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

11. How concerned are you about the possibility that someone may misuse your identity on the internet?

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very concerned

Section 4: Current online habits and practices

This section brings together general questions about using the Internet in your daily life.

1. Do you have internet access at home?

- ☐ By Wi-fi or Ethernet cable
- ☐ By mobile network (eg. 4G or 5G)
- ☐ No

2. How often do you use the internet?

- ○ Everyday
- ○ 2 or 3 days a week
- ○ Once a week
- ○ Once a month
- ○ Never

3. How much time do you spend per day surfing the internet?

- ○ <1 hour
- ○ 1-2 hours
- ○ 2-5 hours
- ○ 5-10 hours
- ○ 10+ hours

4. What device(s) do you use to access the internet?

- ☐ Computer
- ☐ Laptop
- ☐ Tablet
- ☐ Phone
- ☐ Other...

5. For what purposes do you use the internet?

- [ ] Search for information about politics, health, etc.
- [ ] Social media
- [ ] Online shopping
- [ ] Consult email
- [ ] Listen to music
- [ ] Play videogames
- [ ] Watch movies/series
- [ ] Study online
- [ ] Look for a job
- [ ] Work
- [ ] Publish to a blog or online journal
- [ ] Other...

6. What platforms/applications do you use in your day-to-day life?

- [ ] Social media
- [ ] Instant messages
- [ ] Email
- [ ] Browser
- [ ] Online shopping
- [ ] Photo/video editing
- [ ] Online banking
- [ ] Anti-virus
- [ ] Contact tracing
- [ ] Delivery services

7. On social media, do you usually:

- [ ] Use your own photo showing your face as your profile picture
- [ ] Use your real and full name
- [ ] Put the date of birth, age and other information
- [ ] Do not post information that could identify me

8. Your activity on social media goes through:

- ☐ Sharing photos of children/family members who are minors
- ☐ Follow any elected officials, candidates for office or other political figures
- ☐ Follow celebrities, or people with some notoriety
- ☐ Post links to/from business, sport or articles
- ☐ Post links to political stories or other articles
- ☐ Post political or social opinions
- ☐ I don't use social media

Section 5: Profile identification

This section brings together more specific questions about using profiles to make it easier to create personalized experiences.

1. Do you know the concept of profiling?

Long answer...

2. If a law enforcement officer asks for your personal data, are you willing to share it?

- ○ Yes
- ○ No
- ○ I do not know

3. In what situations are you willing to provide personal data when data collection is disclosed?

- ☐ Renew your citizen card
- ☐ Medical consultation (face to face)
- ☐ Take out health insurance
- ☐ Create new bank account
- ☐ For contact tracing
- ☐ Confined in the hospital
- ☐ Make a new Internet/Mobile plan
- ☐ Online shopping
- ☐ Apply for a bank loan or credit card
- ☐ Online medical consultation
- ☐ Request loyalty/discount cards
- ☐ None

4. In what situations are you willing to provide personal data when data collection is not disclosed?

- ☐ Renew your citizen card
- ☐ Medical consultation (face to face)
- ☐ Take out health insurance
- ☐ Create new bank account
- ☐ For contact tracing
- ☐ Confined in the hospital
- ☐ Make a new Internet/Mobile plan
- ☐ Online shopping
- ☐ Apply for a bank loan or credit card
- ☐ Online medical consultation
- ☐ Request loyalty/discount cards
- ☐ None

5. Do you usually answer online questionnaires when they are requested?

	1	2	3	4	5	6	7	
Never	○	○	○	○	○	○	○	Always

6. When answering questionnaires, do you enter any false/incorrect information?

	1	2	3	4	5	6	7	
Never	○	○	○	○	○	○	○	Always

7. When creating an account on an online platform, have you ever entered false personal data?

	1	2	3	4	5	6	7	
Never	○	○	○	○	○	○	○	Always

7.1. If yes, why did you enter this false data?

- ☐ I am creating a temporary account
- ☐ I do not want to disclose any kind of personal data, I want maximum privacy
- ☐ I do not find this type of data(s) relevant
- ☐ I want to access the content of the platform as soon as possible
- ☐ I entered the data by mistake
- ☐ I don't have a specific reason
- ☐ The number of data to be inserted is too large
- ☐ Other...

7.2. When you have to enter false data, what types of false data do you enter?

- ☐ Age
- ☐ Complete name
- ☐ Date of birth
- ☐ Genre
- ☐ Religion
- ☐ Complete name of parents
- ☐ Marital status
- ☐ Home address
- ☐ Work address
- ☐ Nationality
- ☐ Qualifications
- ☐ Professional experience
- ☐ Height
- ☐ Weight
- ☐ Phone/mobile number
- ☐ Region of birth
- ☐ Ethnic group
- ☐ Signature
- ☐ Spouse's name
- ☐ Email
- ☐ Health condition
- ☐ Face photo
- ☐ Full body photography
- ☐ Username
- ☐ Copy of citizen card
- ☐ Fingerprint
- ☐ Licenses
- ☐ Other...

8. In your opinion, does the data you disclose on one platform influence the use of another? Justify.

Long answer...

9. Do you think companies sell your personal data? Justify.

Long answer...

10. Can the data I disclose serve to create a profile of my online habits?

	1	2	3	4	5	6	7	
Disagree	○	○	○	○	○	○	○	Agree

11. "The information I disclose on the internet can serve to identify me". Do you agree with this statement?

○ Yes

○ No

○ I do not know

12. Are you aware of data brokers?

○ Yes

○ No

13. If so, explain what data brokers are and what they do.

Long answer...

Section 6: Knowledge and habits regarding the Internet of Things

1. What do you understand by Internet of Things?

Long answer...

2. Have you ever used an Internet of Things device (smartwatch, contactless cards, air or sea traffic applications)?

- ○ Yes
- ○ No

3. Do you have an Internet of Things device in your home (assistants like Alexa, smart lock, video surveillance, etc.)?

- ○ Yes
- ○ No

3.1. If so, how often do you interact with that device(s)?

- ○ Everyday
- ○ 2-5 days a week
- ○ 1 or 2 days a week
- ○ A few days a month
- ○ A few days a year

3.2. And for what purposes do you use the device(s)?

Long answer...

3.3. Why did you feel the need to have an Internet of Things device?

Long answer...

4. How familiar are you with the following terms?

	I Know well	I have some knowledge	I do not know
Zigbee	○	○	○
Z-Wave	○	○	○
Bluetooth Low Energy	○	○	○
Smart home	○	○	○
Smart vehicle	○	○	○
Fully autonomous car	○	○	○
Long-Term Evolution (LTE)	○	○	○
Smart City	○	○	○
Wearable	○	○	○
LORA	○	○	○

5. In your opinion, how many Internet of Things devices are there today?

Short answer...

Demographic data

We thank you in advance for your participation. And we ask that you fill in the following questions with demographic information. This section intends to collect general information that allows to characterize the users in statistical terms that will be part of this study when sharing their personal perceptions in terms of privacy.

- Age

○	18-25
○	26-35
○	36-45
○	45-65
○	65+

- Genre

- ○ Male
- ○ Female
- ○ Other...

- Country

Short answer...

- District / State

Short answer...

- Level of Education

- ○ Basic education
- ○ High school
- ○ Bachelor's degree
- ○ Master's degree
- ○ Doctorate

- Professional area

- ○ Farmers and skilled workers in agriculture and animal production, market-oriented
- ○ Farmers, livestock breeders, fishermen, hunters and gatherers, subsistence
- ○ Meal preparation assistants
- ○ Vehicle drivers and operators of mobile equipment
- ○ Directors of hotels, restaurants, commerce and other services
- ○ Directors of production and specialized services
- ○ Directors of administrative and commercial services
- ○ Office workers, general secretaries and data processing operators
- ○ Specialists in the physical sciences, mathematics, engineering and related techniques
- ○ Specialists in legal, social, artistic and cultural matters

- ○ Specialists in finance, accounting, administrative organization, public and commercial relations
- ○ Specialists in information and communication technologies (ICT)
- ○ Armed Forces Officers
- ○ Data, accounting, statistical, financial services and registration related operators
- ○ Operators of fixed installations and machines
- ○ Other Armed Forces Personnel
- ○ Other administrative type support staff
- ○ Direct customer support staff
- ○ Protection and security services personnel
- ○ Teachers
- ○ Health professionals
- ○ Representatives of the legislative power and executive bodies, senior leaders of the Public Administration, of specialized organizations, directors and managers of companies
- ○ Sergeants of the Armed Forces
- ○ Technicians of information and communication technologies
- ○ Intermediate level technicians of legal, social, sports, cultural and similar services
- ○ Intermediate level technicians, from the financial, administrative and business areas
- ○ Technicians and professionals, of intermediate level of health
- ○ Intermediate level science and engineering technicians and professions
- ○ Assembly workers
- ○ Workers in food processing, woodworking, clothing and other industries and crafts
- ○ Cleaning workers
- ○ Personal care workers and similar
- ○ Waste workers and other elementary services
- ○ Personal service workers
- ○ Non-qualified workers in agriculture, animal production, fishing and forestry
- ○ Non-qualified workers in extractive industry, construction, manufacturing and transport
- ○ Qualified construction workers and the like, except electricians
- ○ Qualified forestry, fishing and hunting workers, market-oriented

- ○ Qualified printing workers, precision instrument makers, jewelers, artisans and the like
- ○ Qualified metallurgy, metalworking and similar workers
- ○ Qualified workers in electricity and electronics
- ○ Sellers
- ○ Street vendors (except food) and service providers on the street
- ○ Another situation

- Annual income

  - ○ <10.000 €
  - ○ 10.000-20.000 €
  - ○ 20.000-50.000 €
  - ○ 50.000-100.000 €
  - ○ 100.000-200.000 €
  - ○ 200.000+ €
  - ○ Rather not answer

Software Requirements Specification

Introduction

This report aims to define the details and principles of the privacy assisting IoT application with the ultimate goal of assisting in the development of the application. Defining the scope of the project helps to understand how to better implement the application using this document as its base. Stakeholders tell us who communicates with the system directly or indirectly. Business requirements will have to be defined to know in detail the possibilities of value for both sides. The swimlane diagram gives a general understanding of the main case and what cause has each action, the contextual diagram helps to give an understanding of all actions between the system and the active parties. The data flow diagram, lets you know in detail what results from the particular activities. Finally, the technology requirements allow to know how the system will look like and to define the budget of the work.

The scope and vision of the project

This project is carried out in the context of the Master's thesis in Informatics Engineering, which aims to create a mobile application that would provide information about IoT devices in the users' surroundings like the type of information these devices collect and what privacy options are available. The main objective of this application is to give users another option in order to protect their private data. The application will show the geolocation of the IoT devices, what type of device it is, what type of data is being collect by the device. The application will not detect the devices by itself, this will be done by the users themselves. As for competition there are other similar online systems with the same scope as this project. The application offers an easier way to search for information about the IoT devices that are around users' location.

Stakeholders

A stakeholder can be a person, group, or organisation that is involved in the project, is affected by its process and outcome or can influence its process and outcome. Stakeholders can be internal or external to the project team and the organisation.
It is important to identify the stakeholders to make sure to get all the right requirements for the project and to develop a system that can match the proposed problem well.
The following stakeholders are identified in this project:

- **Programmer/designer of the application**: The programmers are the ones who will create the application and even if they do not use it they are directly related to it.
- **IoT Device Owners**: These device owners will be priority stakeholders being that the application in good part will be directed to them, device owners have an indirect influence.
- **Application Users**: The users will be the main focus of this project, they are the ones that provide the information that will be inserted in the application, since they can change the course of the project they have a direct influence.
- **Thesis advisor**: Because this application is implemented in the context of a master's thesis, the advisor has an indirect influence during the implementation and the final product.
- **University of Madeira**: Because this application is implemented in the context of a master's thesis, the university as an organization has an indirect influence on the final product.

- **Legislation**: The legislation in relation to the privacy of the collected data allows to impose rules on the use of the data. It has an indirect influence.

System Relations with Stakeholders

Contextual diagram

The contextual diagram aims to establish links between the system and the other actors that interact with the application.
Identifies the identities external to the application that interact with the system with data and control between the external entities and the application.

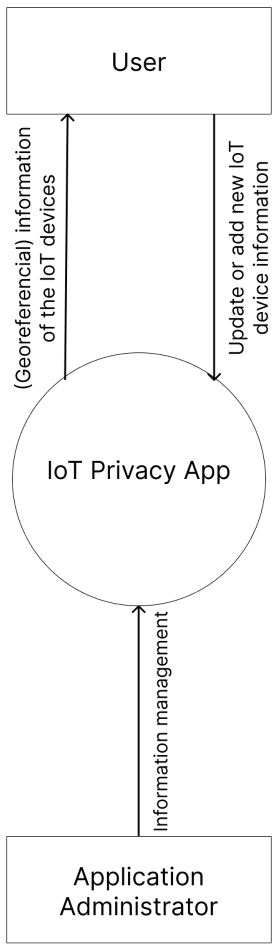


Fig. 27: IoT privacy contextual diagram.

The contextual diagram is based on showing the actions performed with the application.

User:

→ Receives:

- (Georeferencial) information of the IoT devices

→ Sends:

- Update or add new IoT device information

Application Administrator (Programmer):

→ Receives:

- All information related to the application

→ Sends:

- Information management

Data flow diagram

A data flow diagram shows how information flows between the various entities in the system and their relationships.

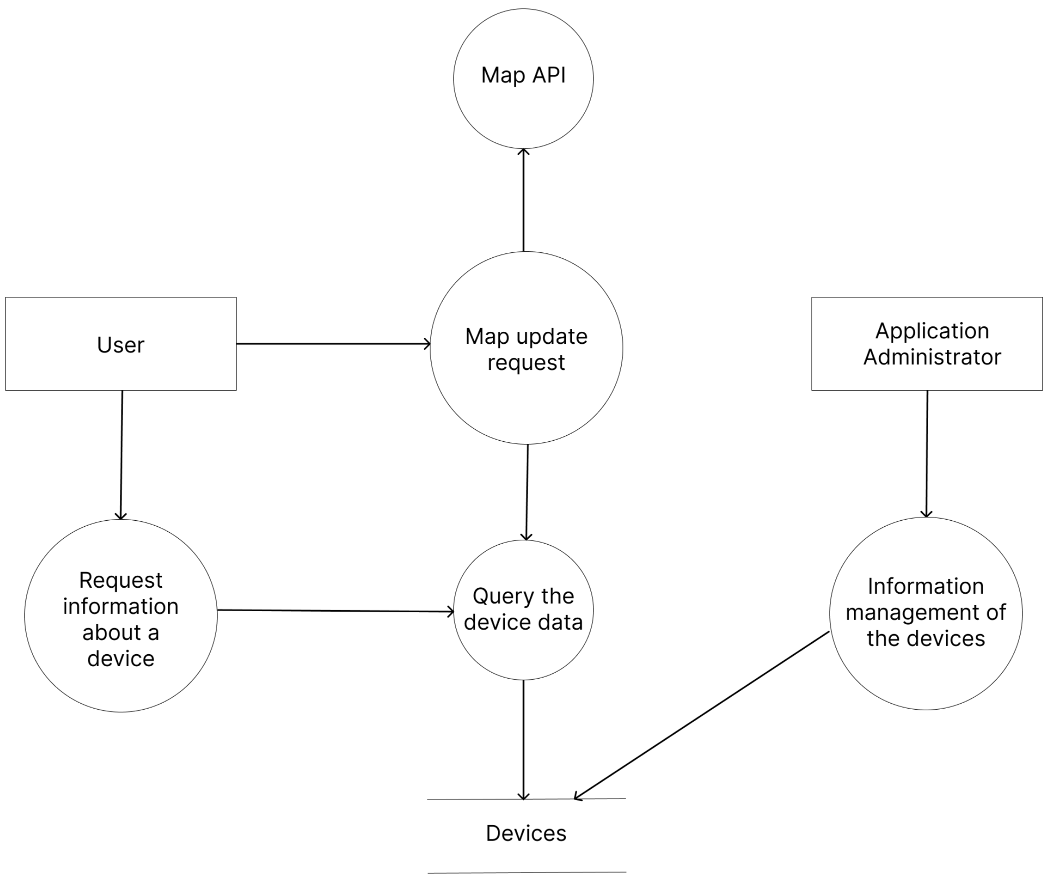


Fig. 28: Data flow diagram.

The user can browse the application map, which will be created with an API, and see the locations of the IoT devices, the user can also look up information about the devices by clicking on a device on the map or by searching for the device in the application. The administrator of the application is responsible for its maintenance (correcting or deleting incorrect data), security and for the veracity of the information.

Swimlane diagram

A swimlane diagram is a type of flowchart in which processes and decisions are grouped into lanes. Parallel lines divide the diagram into lanes, each lane being assigned to people/groups and application.

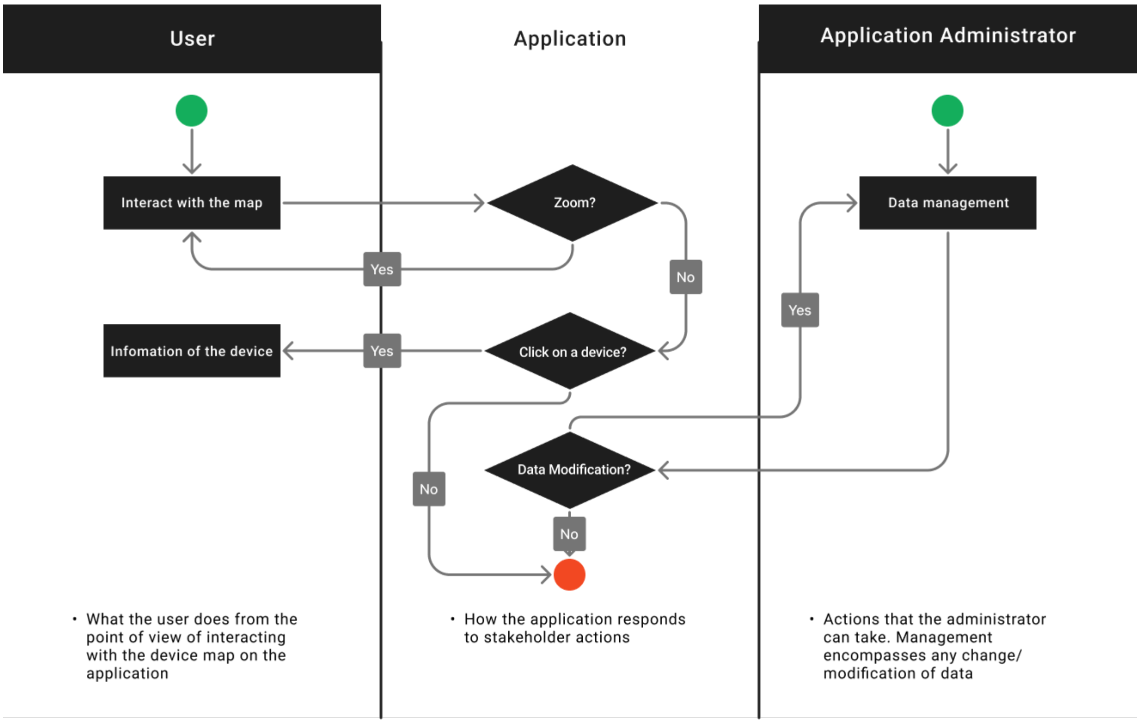


Fig. 29: Swimlane diagram.

This swimlane diagram represents a high level view of a possible user interaction from the application's map, in this map the user can view the location of IoT devices and can see more information about a particular device by selecting it on the map. The application administrator, as mentioned above, can modify the devices' data.

System Features and Requirements

Business requirements

Business requirements describe in business terms what must be delivered or achieved to deliver value. It is what defines the way of doing business, reflecting the internal policy, the defined process and/or the basic rules of conduct. In other words, it is a set of instructions that users already follow and that the system to be developed must contemplate. Restrictions, validations, conditions and process exceptions are classic examples of business rules. A business rule will not necessarily be reflected in the system as a functionality, but it will certainly determine the behaviour of one or

more functionalities of the system.
No business requirements have been identified for this project.

Technology requirements

Technology requirements describe what both hardware and software must be used in order for a system to be realizable. In terms of hardware, it describes what kind of physical components are needed for the software to work. The software to be chosen must take into account the hardware that has been chosen and what is intended by the stakeholders. This has implications for how the system is implemented.
The technology requirements that have been identified for this project are as follows:

- Firestore or similar database server

- Flutter framework with Dart being the main programming language

- Accessible on any smartphone (iOS or Android)
  Note: No web based version is to be available.

These requirements have been chosen so that the system is available to as many users as possible regardless of the hardware they use. The database will allow to store the information that the users provide about the IoT devices. The application will be developed with Flutter since it uses ahead of time and just in time compilation with Dart as its programming language. Flutter has better performance than React Native or a PWA stack and as such it is the chosen framework for this application.

Requirements table

The requirements table identifies each feature and links each feature to an origin.
This is important to make it easier to manage requirements in the future. Knowing the origin of the creation of the requirements makes it easier to refer back to the source to clarify any questions.
A table was created with all the features found. For each feature it was identified the users to which it is applied and the source.
For the features found were described the most appropriate requirements for them.

R#	Feature	Applicable stakeholders	Description	Source
1	Navigate the map	User	**User**: The system should allow the user to scroll through the map of devices	
2	Select device on the map	User	**User**: The system should allow the user to select a device on the map to view more information	
3	Query devices through parameters	User	**User**: The system should allow the user to consult devices of only a certain type, data collected, general location	
4	Query statistics of the devices	User	**User**: The system should allow consulting statistics of devices	
5	Add a device	User	**User**: The system should allow the user to add a new device with name, category, data collected, location, etc.	
6	Edit a device	User	**User**: The system should allow the user to change some data of a device	
7	Delete a device	App Administrator	**App Administrator**: The system should allow the administrator to delete a device	
8	Create account	User	**User**: The system shall allow a user to create an account.	
9	Select privacy choices	User	**User**: The system shall allow the user to select their privacy choices for a certain device if the device allows for it.	
10	See more information about privacy in IoT	User	**User**: The system shall allow the user to check what the terms used in the application mean.	Survey results

Table 4: Requirements table

Functional requirements

Functional requirements define the functions of a system or its components, where functions are specifications or behaviours between system outputs and inputs. These describe what developers have to implement so that users can complete tasks (user requirements), which in turn satisfy business requirements. Functional requirements are essential to the success of a project. After building the tracing table the functional requirements that were needed were extracted for each feature and grouped appropriately according to the following groups:

User Requirements

UR1 - The system shall allow the user to navigate through the devices georeferences;
UR2 - The system shall allow the user to select a device on the map to view more information;
UR4 - The system shall allow the user to consult devices of only a certain category, data collected, etc.;
UR5 - The system shall allow consulting statistics of the devices;
UR6 - The system shall allow the user to create an account;
UR7 - The system shall allow a logged in user to add a new device with name, category, data

collected, location, etc.;
UR8 - The system shall allow a logged in user to change associated data of a device;

Administrator Requirements

AR1 - The system shall allow a logged in administrator to add a new device with name, category, data collected, location, etc.;
AR2 - The system shall allow a logged in administrator to change associated data of a device;
AR3 - The system shall allow a logged in administrator to delete a device;

System requirements

SR1 - The system shall generate statistics related to the IoT devices that reside in the database;

Non-functional requirements

NFR1 - The system shall behave the same in different platforms (Android and iOS);

Use cases

Use cases diagram

The use cases diagram provides a high level visualisation of the user requirements. The box represents the system boundary. An actor's arrow for a use case indicates that he is the primary actor for it. The primary actor initiates the use case and derives the primary value from it.

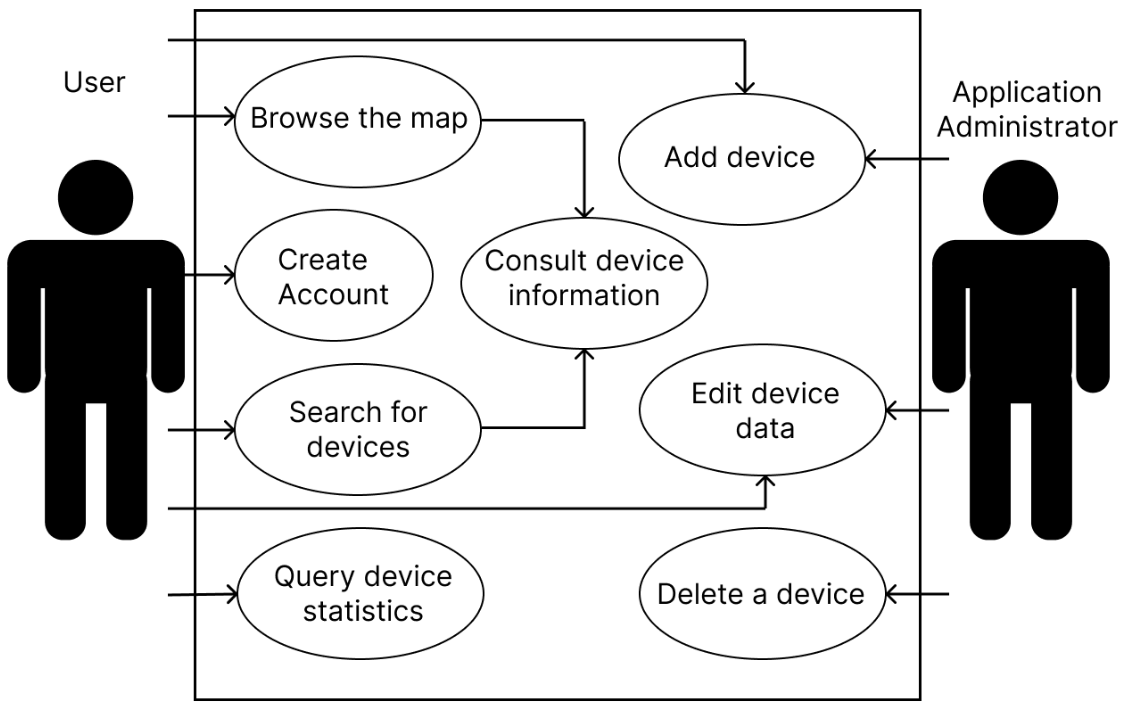


Fig. 30: Use cases diagram.

Use cases

A use case is a type of classifier representing a coherent functional unit provided by the system, subsystem, or class manifested by sequences of interchangeable messages between systems and one or more actors.
This technique describes the tasks that users need to perform with the system or the user-system interaction that may be important to some stakeholders. They also help in testing by checking that the functionality has been implemented correctly. The use cases uses an Unified Modeling Language (UML) notation.

ID and Name:	UC-01 Query devices through certain parameters
Created By:	Nelson Vieira 20/02/2023
Primary Actor:	End User
Description:	The user makes a device information query
Trigger:	The user wants to search device information
Pre-conditions:	N/A
Post-conditions:	POST-1. The user finds device information
Normal Flow:	**1.0 Query information of a device on the map** 1. The user browses the map 2. The user clicks on the icon to show some information about the device 3. The user clicks on the device pop-up
Alternative Flow:	**1.1 Device information search** 1. User enters device name 2. The user chooses the device he wants from a list generated from the search performed
Alternative Flow:	**1.2 Alternative search for information from a device** 1. The user selects one of the parameters: (a) Category (b) Name 2. The user chooses the device he wants from a list generated from the search carried out
Exceptions:	**1.0.E1 The API is not working** 1. The system displays an alert message: “We are having connection problems, please wait for a while”
Priority:	High
Business Requirements:	N/A
Assumptions:	N/A

Table 5: Use case 1 - device information query

ID and Name:	UC-02 Device statistics query
Created By:	Nelson Vieira 20/02/2023
Primary Actor:	End User
Description:	The user queries the statistics of the devices
Trigger:	The user wants to find statistics of devices
Pre-conditions:	N/A
Post-conditions:	POST-1. The system displays the statistics of the devices
Normal Flow:	**2.0 Device statistics query** 1. User selects statistics tab 2. The user can only select certain parameters, such as: (a) Category (b) Location
Alternative Flow:	N/A
Exceptions:	**2.0.E1 The API is not working** 1. The system displays an alert message: "We are having connection problems, please wait for a while"
Priority:	Medium
Business Requirements:	N/A
Assumptions:	N/A

Table 6: Use case 2 - statistics query

ID and Name:	UC-03 Add a device
Created By:	Nelson Vieira 22/02/2023
Primary Actor:	End User
Description:	Addition of a new IoT device in the application
Trigger:	The user wants to add a new IoT device
Pre-conditions:	N/A
Post-conditions:	POST-1. A new IoT device is added to the application
Normal Flow:	**3.0 Add a device** 1. The user enters the following data of a new IoT device: (a) Name (b) Type of data collected (c) Category (d) Photos 2. The user clicks submit 3. The user adds the location of the IoT device on the map
Alternative Flow:	N/A
Exceptions:	**3.0.E1 The device is already in the database** 1. The system displays an error message
Priority:	High
Business Requirements:	N/A
Assumptions:	N/A

Table 7: Use case 3 - add a device

ID and Name:	UC-04 Edit a device's data
Created By:	Nelson Vieira 22/02/2023
Primary Actor:	End User
Description:	Editing the data of an IoT device in the application
Trigger:	The user wants to edit an IoT device's data
Pre-conditions:	N/A
Post-conditions:	POST-1. The data that has been changed appears in the application
Normal Flow:	**4.0 Edit a device's data** 1. The user can change any of the following device data: (a) Name (b) Category (c) Photos 2. The user clicks on submit
Alternative Flow:	N/A
Exceptions:	**4.0.E1 The device to be edited has been deleted in the meantime** 1. The system displays an error message 2. The system prohibits editing
Priority:	High
Business Requirements:	N/A
Assumptions:	N/A

Table 8: Use case 4 - edit a device's data

ID and Name:	UC-05 Delete a device
Created By:	Nelson Vieira 22/02/2023
Primary Actor:	App Administrator
Description:	Deleting a device in the application
Trigger:	The administrator wants to delete a device
Pre-conditions:	PRE-1. The device to be deleted must be in the application's database
Post-conditions:	POST-1. The device is deleted from the application
Normal Flow:	**5.0 Delete a device** 1. The administrator deletes a device, through: (a) ID of device (b) Name of device 2. The administrator confirms the deletion 3. The system deletes the device
Alternative Flow:	N/A
Exceptions:	**5.0.E1 The device to be deleted no longer exists in the database** 1. The system displays an error message 2. The system prohibits deletion
Priority:	High
Business Requirements:	N/A
Assumptions:	It is assumed that the administrator has database access

Table 9: Use case 5 - delete a device

ID and Name:	UC-06 Create an account
Created By:	Nelson Vieira 22/02/2023
Primary Actor:	End User
Description:	Create a new end user account
Trigger:	The user wants to create an account
Pre-conditions:	PRE-1. The user wants to add a new device
Post-conditions:	POST-1. The account is created in the application
Normal Flow:	**6.0 Account creation process** 1. The user enter the following data in the register screen: (a) Username (b) Email (c) Password 2. The system adds the account data to the database 3. An account confirmation email is sent to the email provided by the user 4. The user confirms the account
Alternative Flow:	N/A
Exceptions:	**6.0.E1 The username already exists in the database** 1. The system displays an error message 2. The system allows the user to recover the account **6.0.E2 The email already exists in the database** 1. The system displays an error message 2. The system allows the user to recover the account
Priority:	High
Business Requirements:	N/A
Assumptions:	N/A

Table 10: Use case 6 - create an account

Requirements prioritisation

Regarding the prioritization of requirements, the Quality Function Deployment technique proposed by Cohen in 1995 is used to estimate the priority of a group of requirements. This is based on the benefit of including a feature/requirement, the penalty of not including it, and also the cost and risks associated with implementation. With the MoSCoW method, the initial features are reduced to facilitate the use of the Quality Function Deployment table.
In this approach the values 0 and 1 are used. In case of 1 it means that the column requirement/feature is a higher priority than the row one and if it is 0 the opposite is true.

	R#1	R#2	R#3	R#4	R#5	R#6	R#7	R#8	R#9	R#10
R#1		0	0	0	1	0	0	0	0	0
R#2	1		0	0	1	0	0	0	0	0
R#3	1	1		0	1	0	0	0	0	0
R#4	1	1	1		1	0	0	0	1	1
R#5	0	0	0	0		0	0	1	1	1
R#6	1	1	1	1	1		0	0	1	1
R#7	1	1	1	1	1	1		1	1	1
R#8	1	1	1	1	0	1	0		1	1
R#9	1	1	1	0	0	0	0	0		1
R#10	1	1	1	0	0	0	0	0	0	
Total	8	7	6	3	6	2	0	2	5	6

Table 11: Prioritisation table using the MoSCoW technique

After this initial selection, a prioritisation table is created where it is measured, on a scale of 1 to 9, to rank the benefit and penalty of each requirement. The cost and implementation risk associated to each feature is also estimated.

Feature		**Relative benefit**	**Relative penalty**	**Total value**	**Value %**	**Relative cost**	**Cost %**	**Relative risk**	**Risk %**	**Priority**
Navigate the map	1	9	9	27	13,24	5	10,42	5	10,00	0,65
Select device on the map	2	9	9	27	13,24	5	10,42	5	10,00	0,65
Add a device	5	9	9	27	13,24	3	6,25	4	8,00	0,93
See more information about privacy in IoT	10	5	6	24	11,76	6	12,50	2	4,00	0,71
Query devices through parameters	3	6	8	20	9,80	7	14,58	6	12,00	0,37
Select privacy choices	9	5	7	19	9,31	6	12,50	8	16,00	0,33
Query statistics of the devices	4	3	5	11	5,39	5	10,42	7	14,00	0,22
Create account	8	8	9	15	7,35	5	10,42	5	10,00	0,36
Edit a device	6	7	8	22	10,78	3	6,25	4	8,00	0,76
Delete a device	7	4	4	12	5,88	3	6,25	4	8,00	0,41
Total		67	62	**198**	100,00	**39**	100,00	**37**	100,00	

Table 12: Features prioritisation table

Using this method we get the requirements sorted by priority:

Rank	**Feature**	**# Feature**	**Priority**
1	Add a device	5	0,93
2	Edit a device	6	0,76
3	See more information about privacy in IoT	10	0,71
4	Navigate the map	1	0,65
5	Select device on the map	2	0,65
6	Delete a device	7	0,41
7	Query devices through parameters	3	0,37
8	Create an account	8	0,36
9	Select privacy choices	9	0,33
10	Query statistics of the devices	4	0,22

Table 13: Highest priority requirements ordered

Acceptance Criteria

To make it easier to test whether the highest priority features that were chosen previously were well implemented, these acceptance criteria were created for each of them. These criteria help us understand the minimum conditions for this application to be considered an MVP, i.e., for this project to have the minimum possible requirements in order for it to be considered production ready.
For these acceptance criteria the following was considered:

- High-level functionality that must be present for the system to be usable
- Non-functional criteria and quality metrics that have to be satisfied
- Possibility of open problems or defects (we can guarantee that no defects or TBD is present for the system to be accepted)
- Legal or contractual restrictions (that have to be met for the system to be accepted)

Features

Add a device

- The system allows the user to add a new device that is not yet present in the database

Edit a device

- The system allows the editing of an existing device
- The system saves in the database the changes that have been made

Delete a device

- The system allows the deletion of an existing device
- The system deletes the device from its database

Navigate the map

- The system can represent the devices on the map
- The system allows the user to navigate throughout the map and view the devices

Query devices through parameters

- The system allows searching devices by the certain parameters like the category, the type of data collected

Select device on map

- The system allows the user to select a device on the map

See more information about privacy in IoT

- The system allows the user to see more information about privacy in IoT

Select privacy choices

- The system allows the user to select a device and view its details
- The system allows the user to select privacy choices for that device (if that option is available)

Consult devices' statistics

- The system allows the user to consult statistics concerning the devices

Create account

- The system allows the creation of an account
- The user has to enter its username, email and a password
- The system can detect whether the email is already in use
- The system can send a profile creation confirmation email
- The user can confirm the profile creation
- The system allows the user to sign in

Prototype

For the prototype several drafts are first made with design tools, such as Figma or . Then, with the tools previously mentioned in **Technology Requirements**, an application is created.

Usability Test

Tasks

	Very difficult						Very easy
1. Go to devices page							
(a) Overall this task was?	1	2	3	4	5	6	7
2. Look up more information about one device							
(a) Overall this task was?	1	2	3	4	5	6	7
3. On the homepage look up more information about one device							
(a) Overall this task was?	1	2	3	4	5	6	7
4. Look up more information about this application							
(a) Overall this task was?	1	2	3	4	5	6	7
5. Create an account							
(a) Overall this task was?	1	2	3	4	5	6	7
6. Look up more information about privacy and Internet of Things on the app							
(a) Overall this task was?	1	2	3	4	5	6	7
7. Look up a tutorial on how to add a device							
(a) Overall this task was?	1	2	3	4	5	6	7
8. Add a device to the application							
(a) Overall this task was?	1	2	3	4	5	6	7
9. Update a device on the application							
(a) Overall this task was?	1	2	3	4	5	6	7

Usability Assessment

	Strongly Disagree						Strongly Agree
1. I think that I would like to use this system frequently	1	2	3	4	5	6	7
2. I found the system unnecessarily complex	1	2	3	4	5	6	7
3. I thought the system was easy to use	1	2	3	4	5	6	7
4. I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5	6	7
5. I found the various functions in this system were well integrated	1	2	3	4	5	6	7
6. I thought there was too much inconsistency in this system	1	2	3	4	5	6	7

7. I would imagine that most people would learn to use this system very quickly

1	2	3	4	5	6	7

8. I found the system very cumbersome to use

1	2	3	4	5	6	7

9. I felt very confident using the system

1	2	3	4	5	6	7

10. I needed to learn a lot of things before I could get going with this system

1	2	3	4	5	6	7