Faculdade de Ciências Exatas e da Engenharia

Mestrado em Engenharia Informática

Empowering Users' Privacy Rights in the Internet of Things

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

Wednesday 9th August, 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

List of Tables

List of Acronyms

CCPA California consumer privacy act

GDPR General data protection regulation

IOT Internet of Things

IP Internet Protocol

IT Information technology

M2M Machine-to-machine

MIT Massachusetts Institute of Technology

PA Privacy assistant

PPA Personalized privacy assistant

RFID Radio-frequency identification

SLR Systematic literature review

SaaS Software as a service

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 idea that privacy is a crucial component of personal security, in contrast to the undeniable necessity of public security, including the requirement for guarded walls and closed doors. Long seen as a luxury, privacy is still usually regarded as a good to have rather than an essential requirement, 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". Privacy can be defined [4, 5] 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 [6] 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 [7–9] 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 5000000, 650000 and 80000 documents respectively, including articles, books, conference papers etc.

Most research has focused on the web, while privacy in IoT systems has not been explored as much. Because IoT devices are becoming more prevalent, new methods of communicating, gathering, and analyzing data emerge. Because there is already a substantial quantity of research focusing on web privacy rather than IoT privacy, it is a lot more fertile ground to explore the issue of privacy in the context of the IoT.

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 [10] 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 [11], 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 [12] with thermostats, fridges, microwaves, etc, moving on to smart cars [13], the educational system [14], our clothes and our watches [15] and even into outer space [16]. 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 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 [17] 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 [18], was demonstrated at the Interop Internet Networking show.

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].

IoT is one of the fastest growing technologies [20], it is predicted that it will grow into the trillions of devices by 2030 [21], and with this expansion new security vulnerabilities and data gathering dangers appear, the lack of security in these devices makes them ideal targets for privacy violations and inadequate customer disclosure of device capabilities and data practices aggravates privacy and security issues.

Privacy in IoT systems in not seen as a crucial factor in development [22]. 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 [23–27].

1.1 Título subsecção

ABC

1.2 Título subsecção

ABC

1.3 Estrutura do Documento

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. [28] 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 [29–33], 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 [34–36] apply the theoretical concept of the *homo economicus* [37], 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. Different cognitive biases and heuristics can influence how consumers make decisions, according to several studies on consumer choice behavior [38–41]. According to several articles [42, 43], 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 secondhand accounts. Taddicken's study [44] argues that peer pressure is the reason people have this contradictory behavior, Norberg et al. [45] explains this paradox by suggesting that while perceived risk affects reported attitudes and behavioral intentions, trust has a direct impact on privacy behavior, while others [41, 46] rely on quantum theory. Brandimarte et al. [47] 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 [42, 48–50], online privacy practices are founded on separate privacy mindsets and so they are not inherently paradoxical.

2.2 Privacy in IoT: Approaches

There have been a number of systematic literature reviews (SLR) [51–54] and systematic mapping reviews [55, 56] done to study privacy and security issues in IoT.

In Gupta and Ghanavati's [51] 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 following categories: 'Ontological Modeling and Semantic-based Approaches', 'Data-Driven Approaches', 'Source Code Analysis-based Approaches', 'User Studies and Survey-based Approaches', 'Blockchain-based Approaches' and 'Architectural and Framework-based Approaches'. They then examine current literature on these three prerequisites. The findings show that: most works concentrate on single IoT devices when addressing privacy threats; When analyzing 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 consumers with real-time privacy notices. However, this SLR has the following limitations: the authors only chose articles and not thesis or books and from the selected papers, only the ones written in english were considered.

Kühtreiber et al. [52] 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 establishing the significance of the articles chosen.

Sicari et al. [53] 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, however, has flaws: the prior research analysis focuses primarily on security needs 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 field of IoT privacy that require more attention.

Lin et al. [54] 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, as was already said, consider privacy to be one of the most crucial security aspects, rather than viewing privacy as a distinct concern. Furthermore, the research does not conduct a thorough gap analysis to discover the weaknesses of prior efforts.

Based on Ziegeldorf's [57] 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.

Another concept worth analyzing is differential privacy 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.

2.3 Differential Privacy

The notion of differential privacy, according to Michael Kearns [58], 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". 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".

Differential privacy 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 differential privacy algorithms are no different, they can still help in regards to addressing technology's inherent moral quandaries.

There exist other algorithms that aim to preserve privacy in the same way as differential privacy such as Google's box blurring algorithm [59] that is used in the Google Map's street view, Microsoft's Visor [60] which is a video-analytics-as-a-service tool and Shokri and Shmatikov's [61] system for collaborative deep learning, however, in general, these algorithms struggle with high computational cost, internal attacks, or non-provable privacy.

Zhao et al. [62] conduct a SLR on differential privacy for unstructured data. The authors present differential privacy 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 differential privacy 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.

2.4 Título subsecção

ABC

2.5 Título subsecção

ABC

2.6 Título subsecção

ABC

3 Proposed solutions

This section list seven solutions that emerged from the structured literature review to improve the gap between privacy and security concepts among systems and users.

3.1 Creating new ways for user awareness

There has been some work done to determine the users awareness of their actions online regarding their privacy. Skirpan et al. [63] built an interactive theatre experience, 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 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.

3.2 Legislation

Some papers seek to improve legislation [64, 65] 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. But better protection laws for the user 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.

3.3 Privacy through security

Sun et al. [66] 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.

3.4 Architecture / Framework Proposals

Antunes et al. [67] do a SLR on federated learning in the area of healthcare and make an architecture proposal. The technique known as federated learning allows for the distributed training of machine learning models using remotely hosted datasets without the requirement for data amplification. 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 [68]. 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 that 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. [69] 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 modeling 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 modeling 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.

3.5 Blockchain

Blockchain is an option to guarantee privacy in IoT because of zero-knowledge proofs, ring signatures and mixing [70].

A zero-knowledge proof is a cryptographic technique that enables one party (the prover) to demonstrate to another (the verifier) that a certain claim is true without disclosing any information other than the validity of that claim. Completeness, soundness, and zero-knowledge are the three requirements that must be satisfied by a zero-knowledge proof method. Completeness means that the verifier must be able to confirm that the prover is stating the truth if the information supplied by the prover is true. Soundness indicates that the verifier must be given the opportunity to contradict the prover's claims of speaking the truth if the information provided by the prover is untrue. Zero-knowledge refers to the need that the method only reveal to the verifier whether the prover is speaking the truth or not.

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.

Yu et al. [71] shows various implementations of blockchain that provide privacy through security, based on different categories like data integrity, data sharing and authentication and access control.

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. [72] 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.

3.6 Other proposals

Zhu et al. [73] 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 in IoT sensing, 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.

3.7 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) [74], 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 consumers advice 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 consumers but also take away their ability to influence the process. They found that the participants' attitudes regarding the various implementations were generally favorable, 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) [75], 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-centered 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 consumers 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 consumers 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 [76] 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 [76] 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.

3.8 Sniffers

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

The LTEye project [77] 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.

3.9 Main Takeaways

There are two main ways to provide privacy in IoT systems, through security or using privacy notices, other ways like through legislation or with the creation/usage of a framework that provides privacy fall into these two categories. Most of the literature assumes that security and privacy are synonyms, for example [65, 66, 69], and so most of the proposed solutions fall under privacy through security. The proposed solutions that use privacy notices, like [75], 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 [78] 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 [79]. Hitchhiking [80] is a method of protecting the privacy of users who divulge their location, applications regard locations as the object of their attention. The fidelity tradeoff is removed as it is not important to know who is in a certain location. The introspection [81] method examines 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 analyzed. The privacy of consumers is jeopardized if an IoT device loses integrity due to a hostile assault.

3.10 Título subsecção

3.10.1 Título subsubsecção

4 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.

According to Qu et al. [82], several significant barriers remain, including the lack of a theoretical foundation, the trade-off optimization between privacy and data value, and system isomerism over-complexity. Because there are no mathematical foundations for IoT structure design, IoT system designs are planned and executed using empirical approaches, which have limitations in IoT development. Scientific theory and quantitative analysis must enable trade-off optimization, yet, there are multiple parties with diverse characteristics and requirements, making this optimization highly challenging. A plethora of standards and protocols add to the unneeded complexity of system isomerism. Ensuring effective IoT applications with as few resources as possible implies fewer resources available for privacy protection; however, lightweight privacy protection cannot meet all of the criteria, and attackers can exploit structural information to launch multiple concurrent attacks.

5 Methodology

The overall work will be comprised of two phases which will be described in the following paragraphs. Phase one mainly described throughout this paper, focuses on collecting the state of the art in terms of the most relevant topics, from which main privacy concepts were selected to be explored in the stage 1 of Phase 2 with the preparation of a questionnaire to collect user perceptions regarding privacy and topics collected in the systematic literature review. The second stage of Phase 2 consists 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 Phase 1 Systematic Literature Review gathered the most relevant papers discussing methodologies and techniques for the protection of users' privacy data with special focus on IoT systems. For this SLR, this paper considered focusing only on papers from the last 12 years, from 2010 until 2022, since papers before then become out of date with the evolution of technology. In this SLR, it was reviewed 54 papers published in top computer science, security, privacy and software engineering outlets.

This paper followed Keshav's three-pass approach [83] when choosing which papers to read fully and which ones to ignore, 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, 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? Only 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. Having collected the major findings, this work then aims to conduct a throughout study split in several stages and around the specific research questions which will be explored in each phase. For that matter, the research questions listed are:

Phase 1:

RQ1: What approaches are being considered for privacy issues in IoT in the currently available literature?

RQ2: What are user perceptions on online privacy?

Phase 2:

RQ3: How to empower users to protect their privacy rights?

RQ4: What issues are prevalent in IoT that make it difficult to address privacy and security problems?

The second phase will be evaluated on two stages, the first one consists on doing a study on people's general privacy concerns, while using and interacting with IoT devices. This study will abide on preparing a 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 is to both understand the privacy paradox and collect data on their proposal to address privacy issues with regard to IoT devices.

5.1 Stage 1: User perceptions

This study aims to understand people's perception of IoT and their privacy practices online. It also serves to demystify the privacy paradox and also to help provide a solution to the privacy issue in IoT. The questionnaire consists of 92 questions divided into 7 sections to access users' knowledge, it follows a kind of narrative, the first section being general privacy questions then about the predisposition to data sharing, to concerns with privacy then about daily digital routines, then about profile identification, and then about IoT general knowledge before a section about non-identifiable demographic data. The scale that is used in the questionnaire is based on the work of Philip K. Masur [84]. Great care is taken when it comes to this survey'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.

This survey was partially based in a study done in the Philippines by the government in the context of their privacy act of 2012 [85], this was the second survey done on the country's population. It was also inspired by Alves's master's thesis [86], which was about citizen's perception about privacy in the wake of GDPR.

This survey was done through the internet, it was created in Google Forms, this way it is guaranteed to reach the most people possible, besides Google Forms itself, it will be used other online venues for distribution and even printing.

The questionnaire was available for completion until August 30, 2023, and during the time that it was open 45 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) and 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.

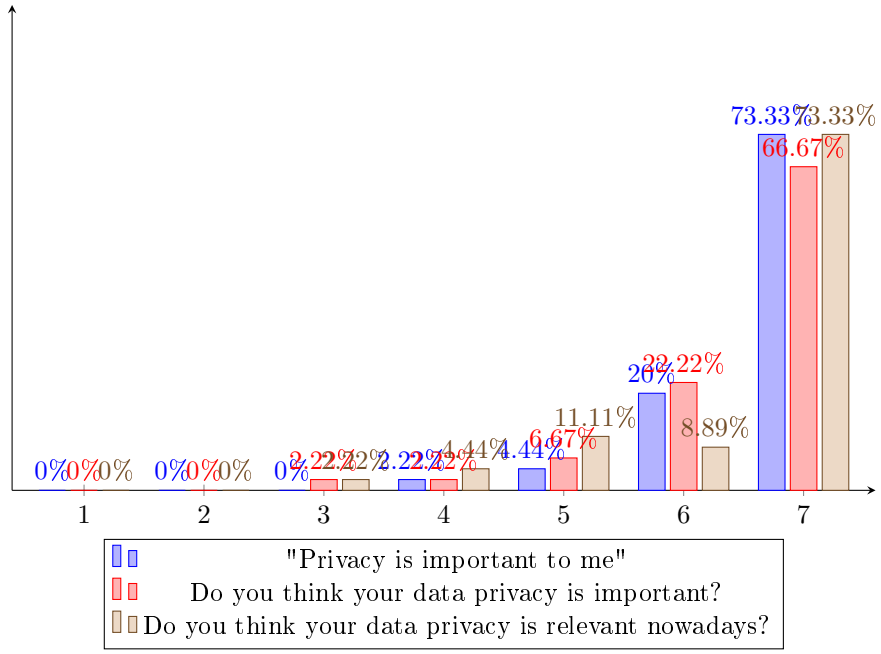


Fig. 1: Responses related to the importance of personal data privacy.

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 1. Most participants also have some knowledge of behaviours or techniques to do in an online environment, as shown on Figure 2, 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.

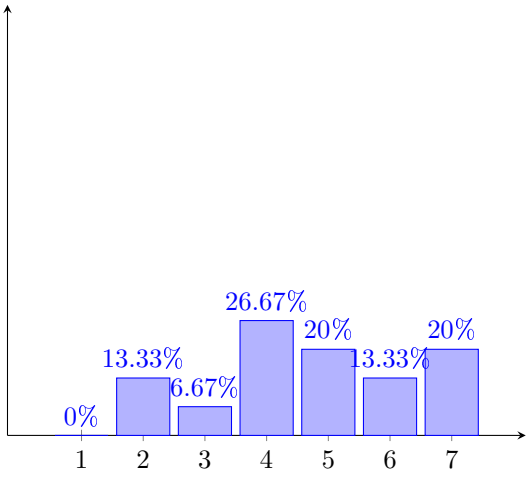


Fig. 2: Participant responses indicating whether they know techniques to guarantee privacy and the protection of their data when using the internet.

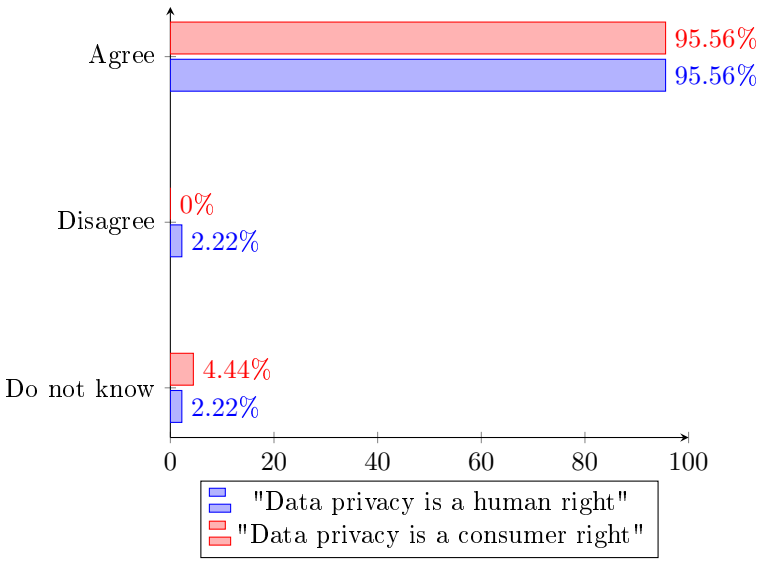


Fig. 3: Participant responses indicating whether they agree or disagree with these statements.

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

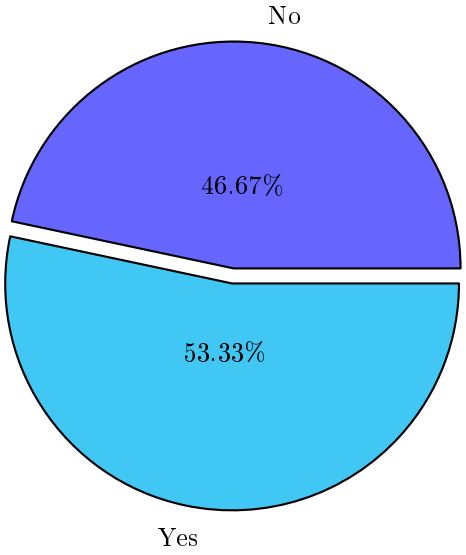


Fig. 4: Responses to security and privacy being synonymous.

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 in Figure 4, where 53.33% of participants believe that privacy and security are synonymous.

Participants have some digital literacy of IT terms, as shown on Figure 5, most know the more popular terms like *wi-fi*, *cookie* and *data protection*, but as the terms become more esoteric the

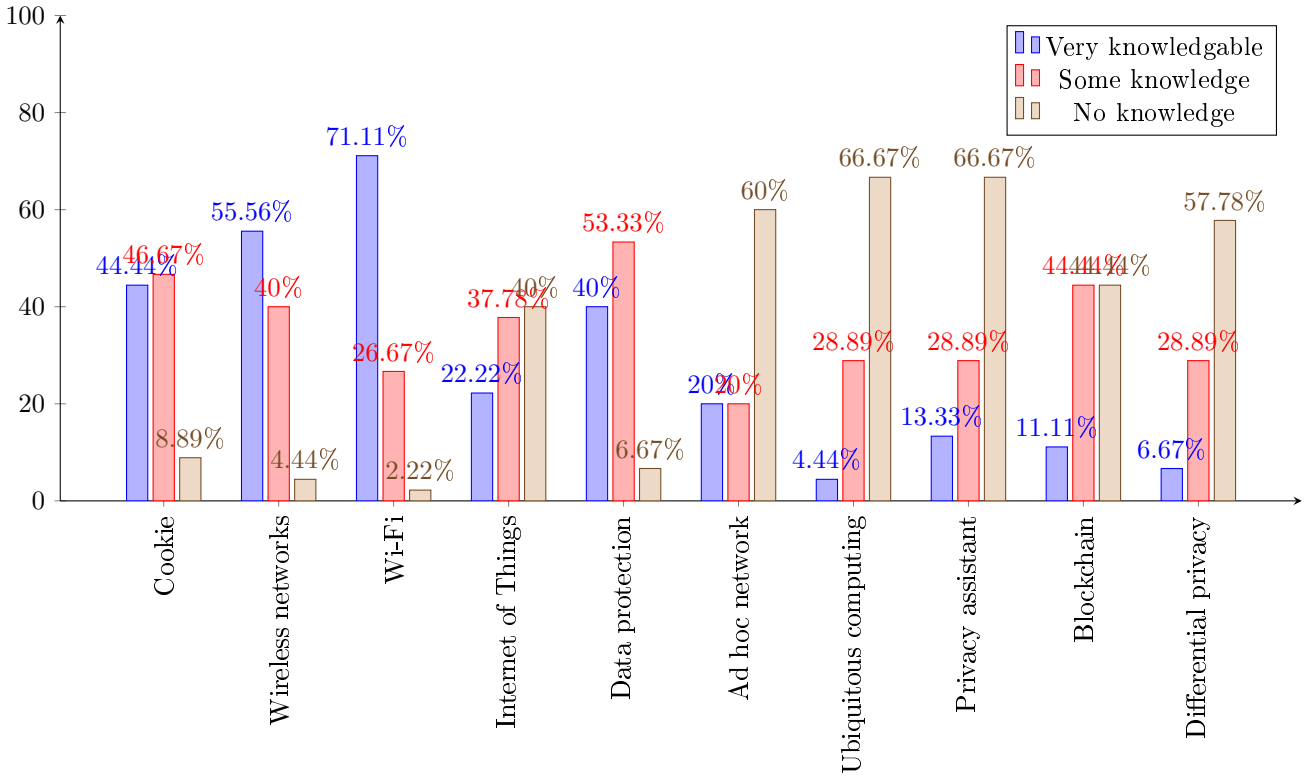


Fig. 5: Familiarity with general IT terms.

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*.

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 9, and are somewhat worried but do not actively try to protect their data privacy, which comes as a bit of a contrast with their early answers. When asked if they accept cookies, respondents occasionally do but are unsure what they do or their importance on the user experience.

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 7.

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 9, and are somewhat worried but do not actively try to protect their data privacy, which comes as a bit of a contrast with their early answers. When asked if they accept cookies, respondents occasionally do but are unsure what they do or their importance on the user experience.

5.2 Stage 2: From theory to practice, an Application

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

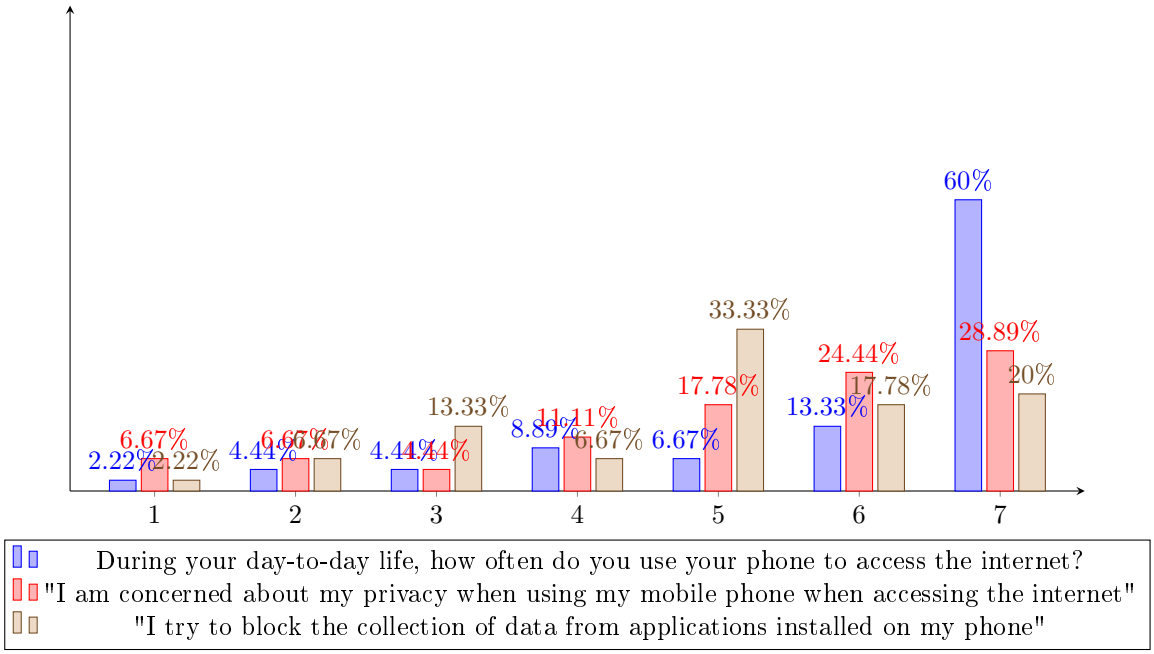


Fig. 6: Responses related to phone usage.

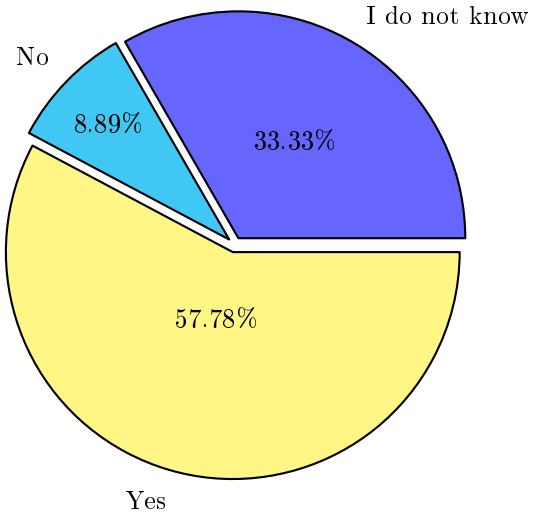


Fig. 7: Responses to the question: Do you consider that your internet activity contributes to the development of profiling?

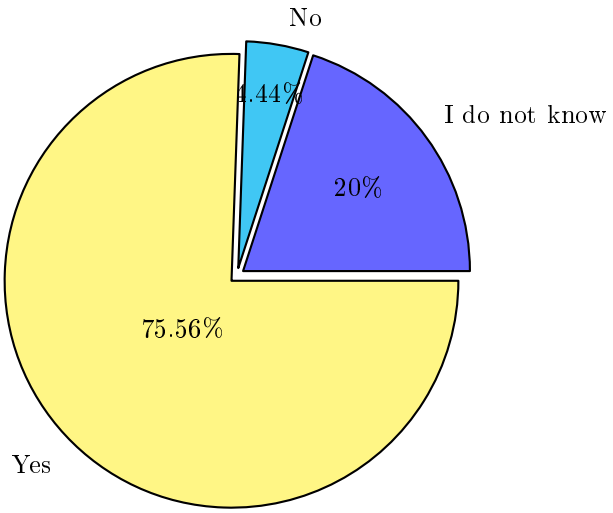


Fig. 8: Responses to the question: "The information I disclose on the internet can serve to identify me". Do you agree with this statement?

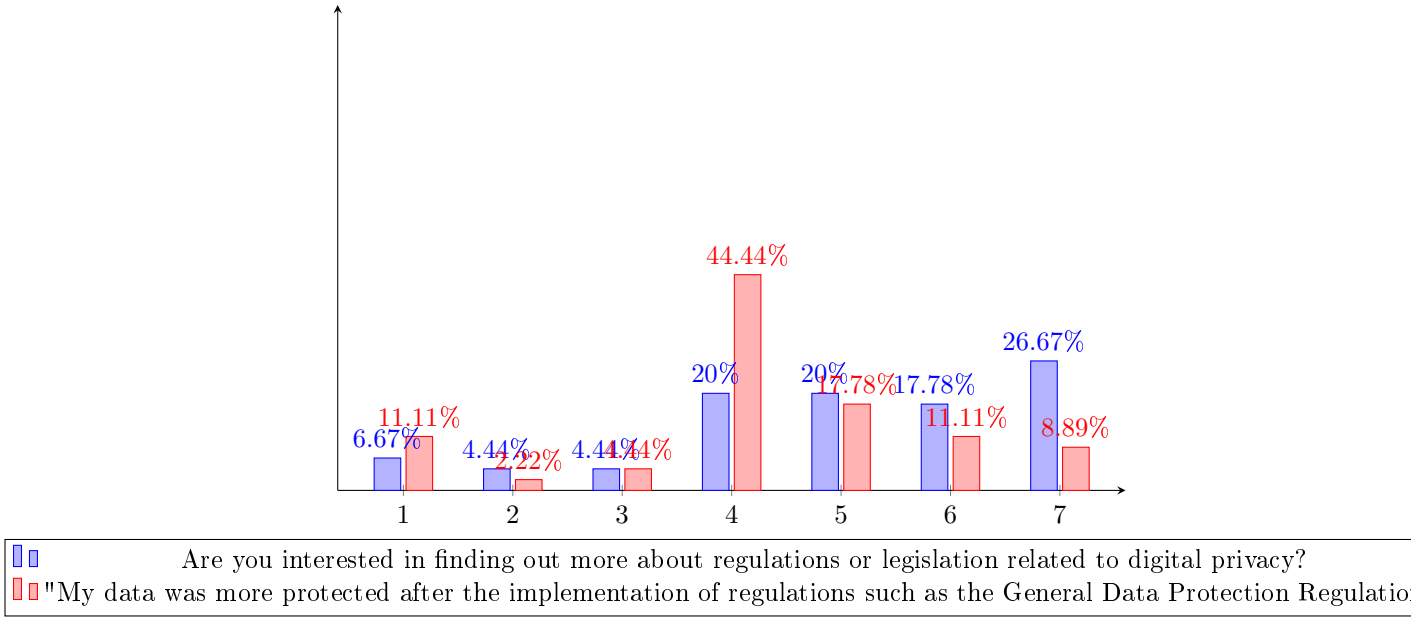


Fig. 9: Responses related to phone usage.

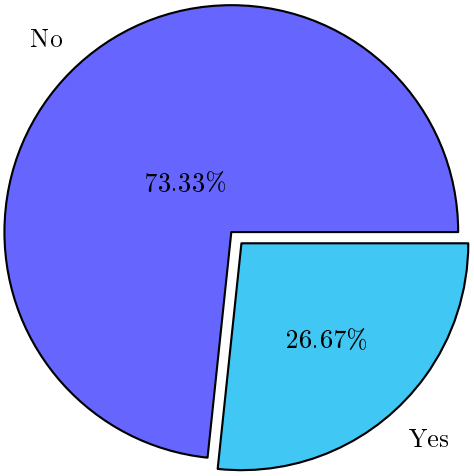


Fig. 10: Responses to the question: Are you aware of the duties of a Data Protection Officer (DPO)?

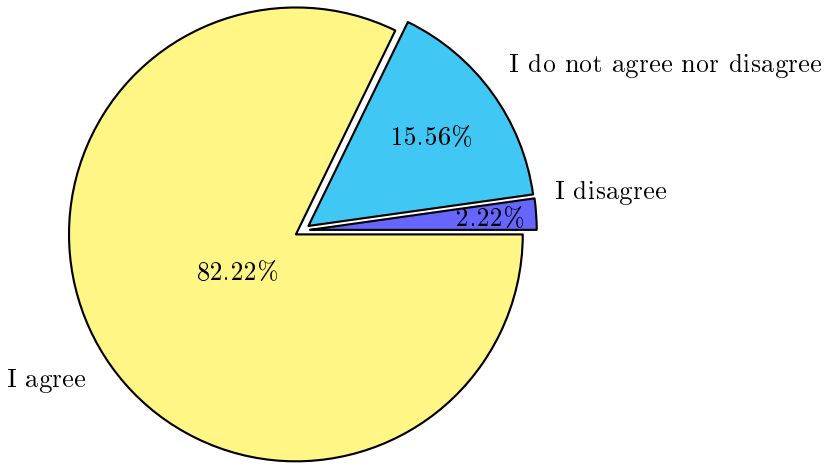


Fig. 11: Responses to the question: "I am interested in knowing where and how my personal information is used." Do you agree with this statement?

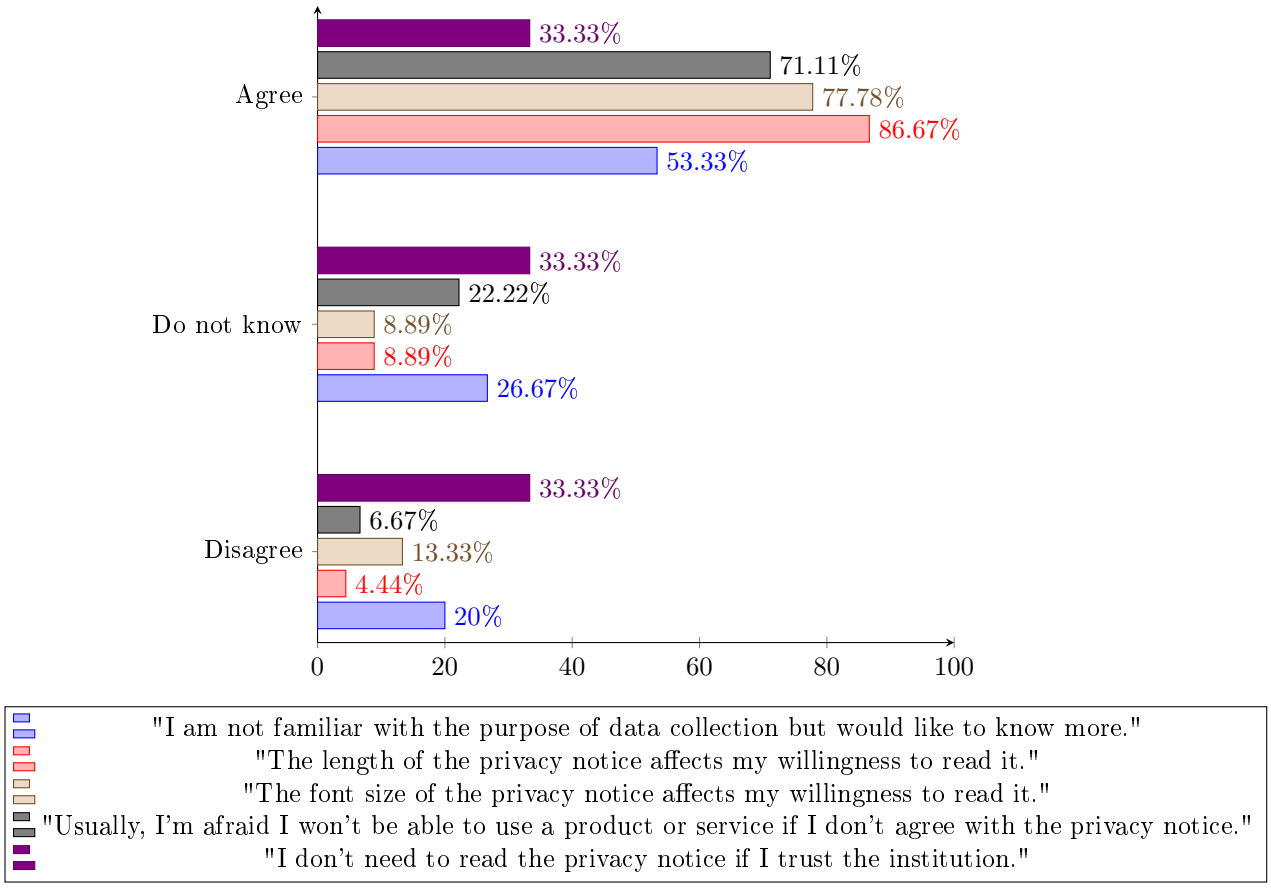


Fig. 12: Participant responses indicating whether they agree or disagree with these statements.

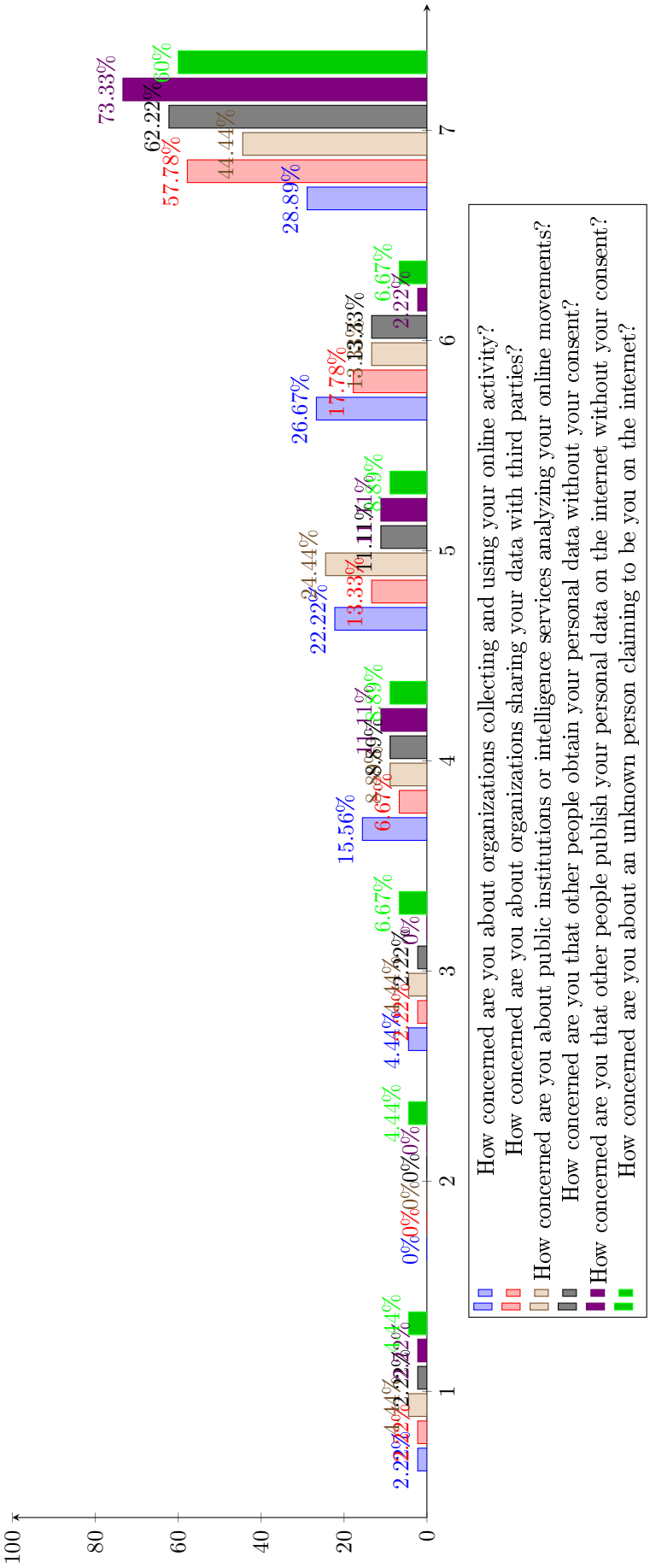


Fig. 13: Responses related to concerns of organizations and individuals handling of private data.

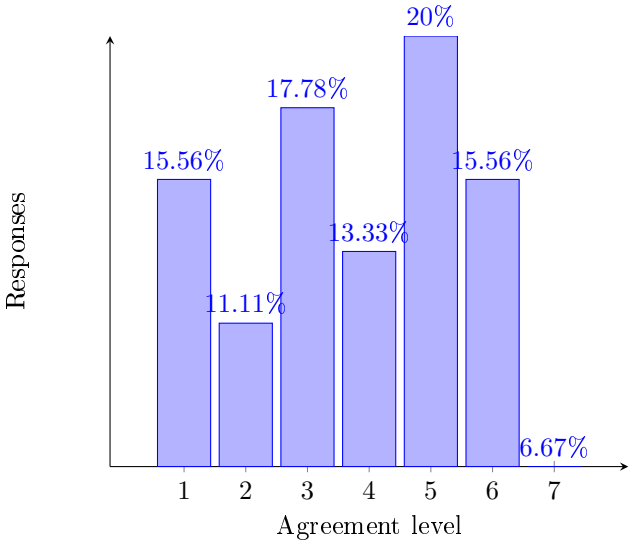


Fig. 14: Responses to the question: When creating an account on an online platform, have you ever entered false personal data?

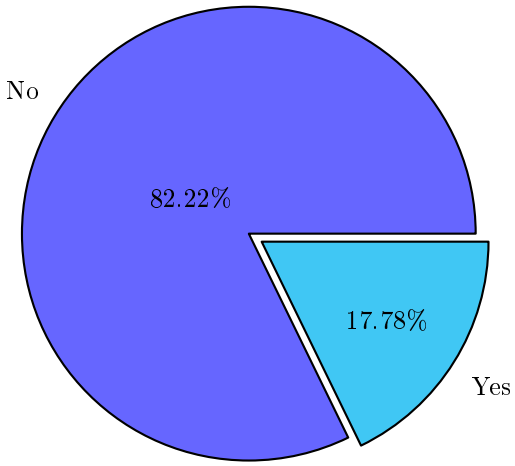


Fig. 15: Responses to the question: Are you aware of data brokers?

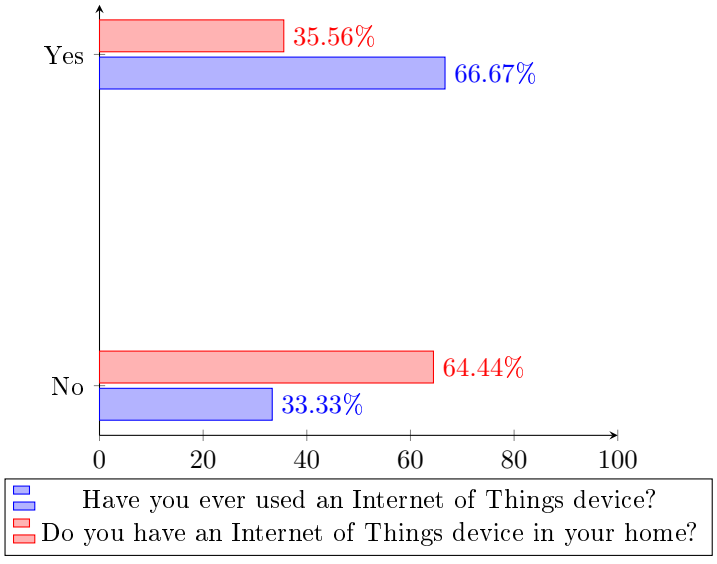


Fig. 16: Participant responses indicating whether they agree or disagree with these statements.

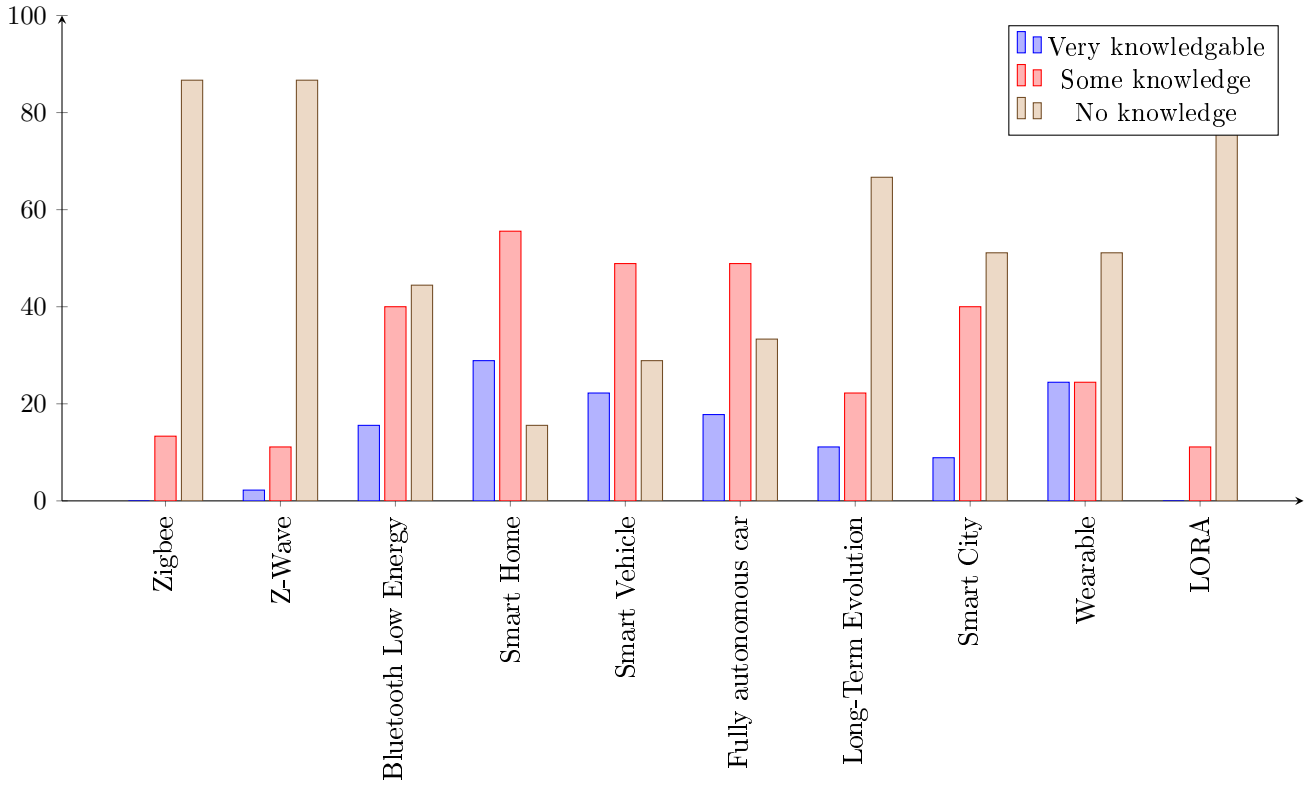


Fig. 17: Familiarity with IoT terms.

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.

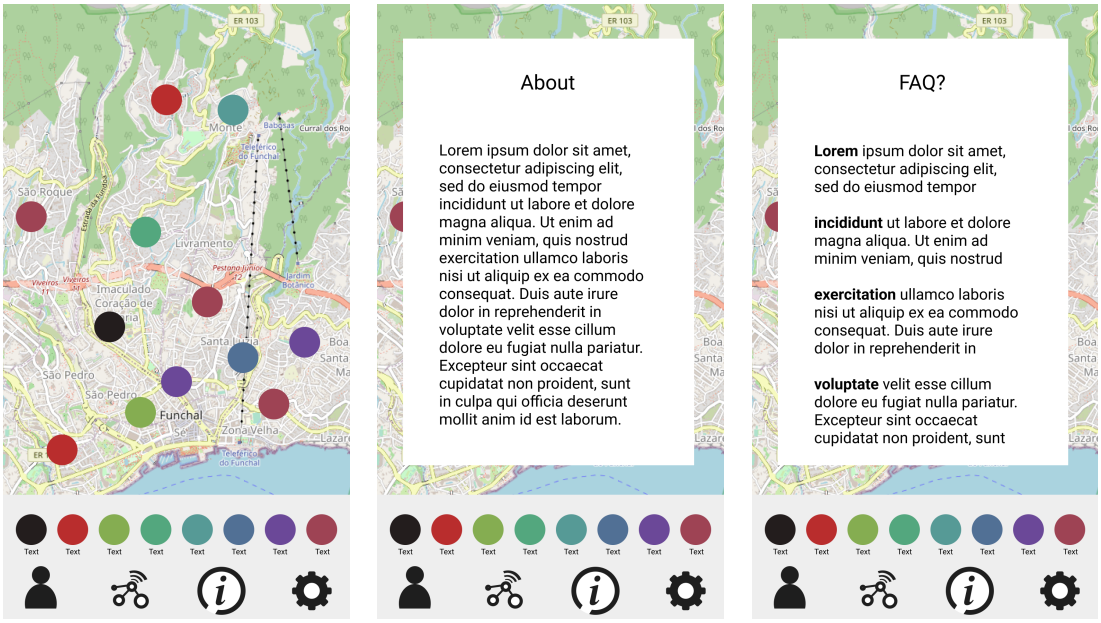
- The **purpose for the data collection**: Defines what is the purpose for the collection of the data, if a device collects temperature and humidity data and is used by a weather based company or government agency then the purpose for the data collection is for weather forecasting.
- **Who has access 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.

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.

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 18 shows tree pages of the low level prototype, these are the



(a) (b) (c)

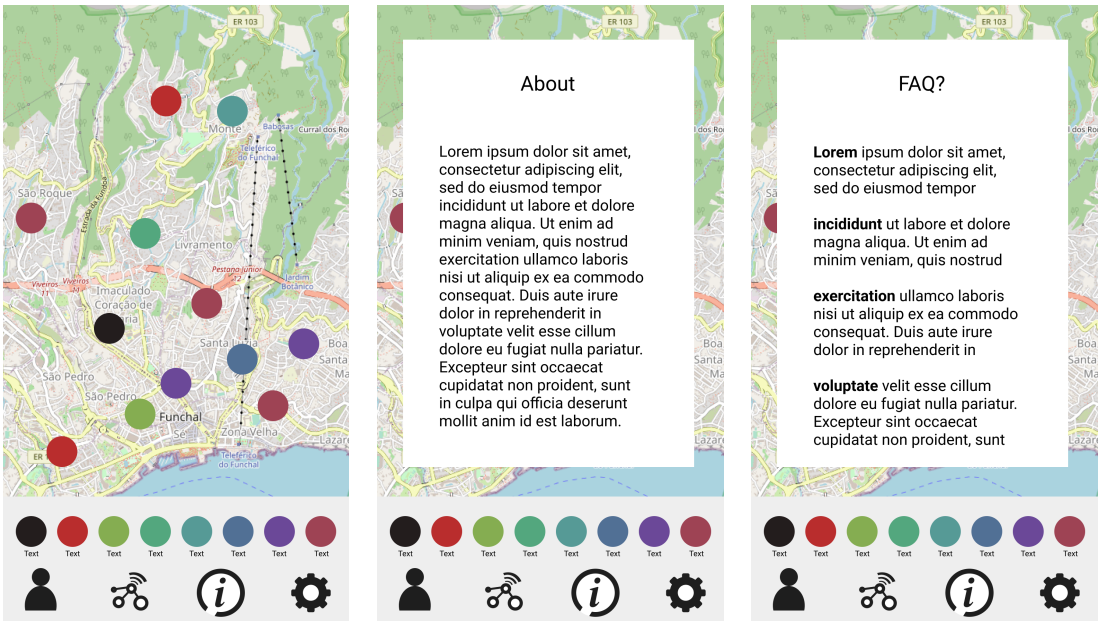
Fig. 18: Low level prototype of (a) homepage, (b) about and (c) FAQ pages.

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 the 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.

Usability tests were conducted in person with X 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 yet done, 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, using this system a score of X was achieved.

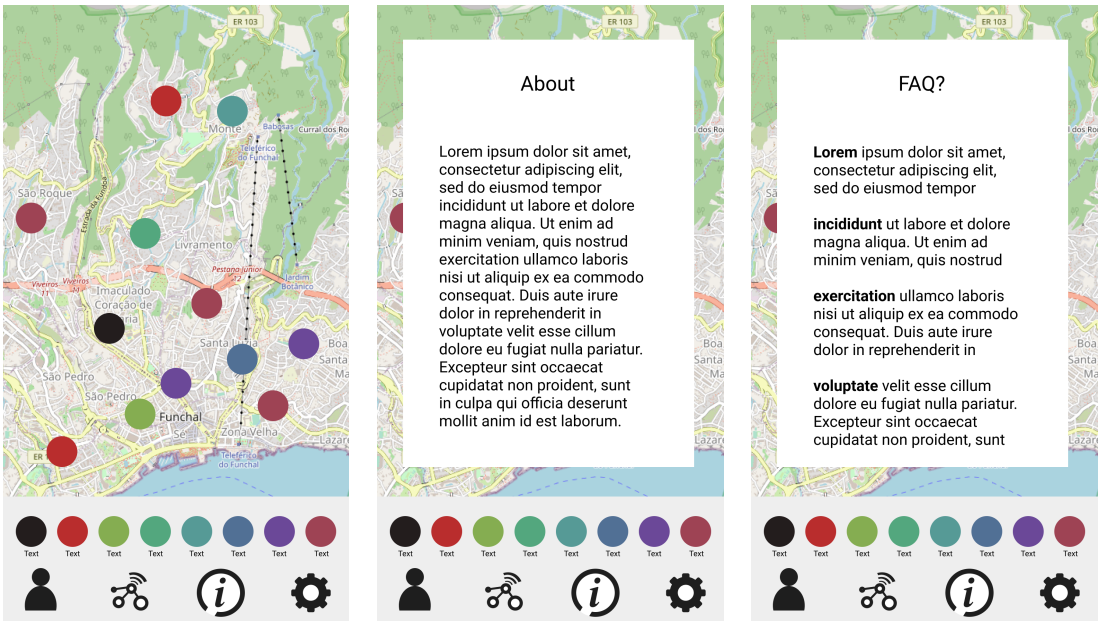
6 Challenges

One of the most difficult points to accomplish in this thesis was the questionnaire, not the fact of constructing the questionnaire but of getting participants. Besides being difficult in itself to get a relatively high number of participants of participants (a few hundred at least) to be able to draw conclusions with any high degree of confidence, it was difficult to get 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 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



(a) (b) (c)

Fig. 19: Medium level prototype of (a) homepage, (b) about and (c) FAQ pages.



(a) (b) (c)

Fig. 20: High level prototype of (a) homepage, (b) about and (c) FAQ pages.

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

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

[6] 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

[7] 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.

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

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

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

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

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

[17] 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.

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

[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] 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/

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

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

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

[28] 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.

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

[30] 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.

[31] 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.

[32] 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.

[33] 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.

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

[35] 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.

[36] 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.

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

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

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

[44] 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, 2014.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

[49] 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

[50] 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.

[51] 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

[52] 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

[53] 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.

[54] 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.

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

[56] 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.

[57] 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.

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

[59] 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

[60] 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/

[61] 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.

[62] 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

[63] 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

[64] 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

[65] 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.

[66] 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

[67] 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

[68] 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.

[69] 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.

[70] W. contributors. (2022) Privacy and blockchain. Accessed: 2023-01-02. [Online]. Available: https://en.wikipedia.org/wiki/Privacy_and_blockchain

[71] 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.

[72] 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

[73] 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

[74] 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

[75] 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

[76] 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.

[77] 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

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

[79] 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.

[80] 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.

[81] 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.

[82] 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.

[83] 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

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

[85] 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

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

Appendix

Appendix content...

Survey

Content...

Software Requirements Specification

Content...

Usability Tests

Content...