Ethics of Personal Data in IoT

Janik Müsse
Jonas Due Rosenzweig
Kanza Ahmad
Namra Imtiaz Jabeen
Sara Holst Winther

With supervisor Henning Christiansen

December 2019

Figure 1: [1]

Abstract

This paper is an investigation of the ethics of personal data handling specifically within the trend of the Internet of Things. With the advent of this technological trend, network connectivity is becoming a more common attribute in commercial products, that might traditionally not have such capabilities. Personal data are becoming more valuable due to amount of data that the data aggregators and re-sellers have in their possession today. The trend of IoT increases the amount of data collected, due to the fact that devices and products that are not traditionally connected to networks are now connected. There can be ethical consequences for consumers, businesses and society due to the way personal data are handled. This report investigates this technological trend, outlines the ethical consequences of the handling or mishandling of personal data in general and looks at a few real-world cases of the IoT products to analyze the way personal data are handled. The preliminary goal of this investigation is to see whether the IoT products that force the collection of personal data through forced internet connectivity present more ethical problems than those that do not. Some of the IoT devices require internet connectivity to function, while others do not. Some that require this connectivity only do so to ensure network security and make efforts to ensure the protection of personal data. The second goal is to outline technologies that can enable the IoT products to connect to the internet while guaranteeing the protection of personal data.

Contents

1 Introduction

1.1 Motivation

The Internet of Things (IoT) is a transformational technological trend in today's world. The fundamental idea behind IoT is to enable different devices and "things" to be able to establish a wireless connection and communicate with other "things" and humans over networks. In addition to that, the IoT either adds functionality through wireless connectivity or automates decision-making without human interaction[2]. A common example of this is the smart fridge. Smart fridge is an IoT enabled refrigerator which can restock food automatically. The advent of this technological trend gives rise to many opportunities and risks for consumers, companies and society[3].

Lately, the term IoT has seen widespread use as a commercial term with little technical meaning, a marketing buzzword of sorts. If a device, appliance or product is being labelled "IoT enabled" it is understood as "capable of connecting to networks". The IoT as a technical term does not refer to individual devices or their capabilities due to the term's overuse in marketing. IoT, as the name suggests, refers to a system of interrelated things connected together over networks. These things can be anything: machines, computers, devices, appliances, sensors to keep track of resources, livestock, autonomous machines or even people. In an IoT setup each of these things is given a unique identifier so it can be recognized over such networks. For the IoT networks to work properly networks depend on these identifiers and in particular the ability of each thing to transfer data over a network without requiring any human involvement[2].

Opportunities of the trend of the IoT for consumers lie primarily in improving the end user experience of the IoT enabled products. This can either be through better data and information acquisition to aid in decision making, or by automating decision making such that a user profile dynamically lets the system make decisions instead of requesting human input at each step. For example, smart irrigation systems can check the weather forecast and set watering schedules in accordance with weather forecasts. In addition to that they can inform the user of plant, crop and soil nutrient levels through an array of sensors and let the user make dynamic decisions, or the system can make these decisions automatically[4].

This helps in reducing information overload for the consumer, which according to Basel Solaiman and Eloi Bosse in the book "Information Fusion and Analytics for Big Data and IoT", is a growing problem in modern society[5].

Opportunities for companies exist primarily in increasing their products' marketability by improving the user experience by adding functionality. Other opportunities for companies lie in taking advantage of the greater range and amount of personal data collected. Societally, personal data acquisition and usage is an equally important opportunity: for example, healthcare applications such as the CPR app (hjerteløber) uses geolocation from devices to find individuals trained in first aid to respond to people suffering heart attacks near their location [6]. Other opportunities of this tech trend for society involve streamlining and efficiency. For example, for public transport applications or in better managing resources such as in irrigation.

These opportunities also present risks. According to "The Internet of Things: Foundational ethical issues"(2018) by Fritz Allhoff and Adam Henschke, the misuse of personal data can have lasting consequences for privacy as well as unintended consequences for society[7], as is outlined in the ethics section. This misuse does not only come from malicious intent from within a company, but can also come from outside sources. This paper argues that heightened cybersecurity should be the goal of the IoT product manufacturers due to these risks. If part of the utility of this trend is the automation of some forms of decision making, product manufacturers and designers need to ensure that the measures needed to enable good cybersecurity in such devices are strong and redundant. Since the nature of the IoT requires personal data to be transmitted over networks to enable functionality of their products, devices, and services, there is, according to the RSA conference from 2019, a greater risk of such data being compromised. This paper would therefore argue that larger investments of time and resources need to be made in adhering to the highest standards and protocols in securing all relevant network communications in the IoT products[8], due to the ethical consequences of data mishandling outlined in Section 2.4 of this paper.

1.2 Relation to Semester Constraints

According to the study regulations for the 3rd semester constraints, the purpose for the 3rd semester project is for the student "to gain experience with scientific-theoretical analysis of natural science (...) working with a representative example"[9]. It is an objective of the project that the student "be able to describe an issue within the Natural Sciences in such a way that it becomes accessible to scientific-theoretical analysis and reflection"[9]. According to the notes on the third semester theme, a discussion on what the third semester project is about, written by previous course coordinators, reflection refers to "well-formed systematic analytical considerations on a clear and logically consistent foundation over a complex problem." These constraints direct the project to ask questions about natural science, shedding light on science as a cultural and social phenomenon, viewed from the outside.

This paper is an investigation into ethical aspects of science, namely computer science. According to the study regulations, "A science-ethics project typically concerns either an internal perspective, (...) or alternatively an external perspective, i.e., ethical aspects of science as viewed by society in general, for example societal consequences of science"[9]. By this definition, the project fits the description of an external science-ethics perspective: in the field of computer science, this paper is looking at the ethics of personal data and data privacy within the IoT. Although the project does not concern the development of a technology, as mentioned in the examples of potential external science-ethics projects found in the past supervisor notes, it concerns itself with a technological trend and its analysis through the lens of various fields of computer and information sciences.

1.3 Personal Data, Research Question, and Hypothesis

According to Gary Allemann, in his book "Data Management, are you seeing the real value?", personal data today is the most valuable commodity in the world, if one were to look at the total value of all personal data, and compare it to the total value of the world's oil reserves, for example[10]. Each person on the planet with access to the internet has on average 52,000 data points – according to Julia Angwin et al.[11], or user 'traits' describing various attributes of their online activities, ranging from metrics such as

time spent online and search history, geolocation data, music tastes, reading habits, level of education, political leanings, nationality, sex and so on. The Netflix documentary The Great Hack makes the hypothetical claim that in the future, the advancement of AI can lead to much more information being extracted from these data[12]. Essentially, this can mean that given enough history on a particular person through access to all the personal data they produce, AI algorithms may be able to predict the behavior of individuals based on lesser amounts of data[12]. This heightens the need for something to be done in the field of personal data protection, as soon such protections may be entirely futile.

How much emphasis is put on securing users' personal data is one central aspects of the investigation in this paper. Another central aspect is how these personal data are used and can be used. This paper aims to outline the ethical problems in certain IoT devices based on these aspects. These ethical problems are presented from a consequentialist perspective, are researched and based in events that have taken place, and are presented in the ethics section. This paper looks at IoT products, devices or services to outline the good and the bad of personal data handling in the IoT, as well as to help answer the following research question:

- To what extent is the handling of personal data within the IoT ethical or unethical, from a consequentialist perspective?

A list of sub-questions is presented here that are useful in this investigation:

- Does the forced or unforced external network connectivity aspect of the IoT device affect the ethical concerns of our research question?
- What ethical considerations exist when handling personal data, generally and in the IoT?
- What are the ethical consequences of mishandling personal data in the IoT, in relation to personal data privacy, security, ownership and the transparency of the handling of data?
- How can different technologies help make handling personal data in the IoT more ethical?

This investigation is expected to show that investment in top of the line network security and other types of cybersecurity leads to less ethically troubled IoT products. Indeed, manufacturers and designers of IoT devices and

systems need to have a good understanding and mastery of technologies that enable personal data to be communicated over networks unchanged, untouched, privately and securely. This is especially critical in the IoT, as personal data drives a lot of the functionality of such systems. More recent technologies, such as blockchain, can help in this task, ensuring data anonymity (removing any personally identifiable data). The pursuit of data privacy and protection in an earnest and transparent way is a good step towards IoT products being more ethical – yet from the consequentialist ethical perspective, intentions do not matter. This paper only cares about the outcome. In that matter, it can be said that these technologies, if they achieve greater data privacy and protections, renders the product they are employed in more ethical. Other aspects of the products, such as the business model, marketing, forced/unforced outside network connectivity, added functionality through network connectivity, among other aspects, are also taken into account here – this paper goes into more detail in the discussion.

1.4 Approach to the Problem

The approach to the problem is the following: first, it is intended to present and describe a well thought out and researched ethical framework that enables the investigation to answer the research question. A technical investigation into the central technological aspects of the IoT systems and devices is then presented, in order to create a foundation of knowledge on which to build a discussion and analysis on the ethical considerations and consequences associated with IoT enabled devices and systems and the data that these handle. These two aspects of research helps ground both aspects of the paper: the ethical side, as well as the technical side.

Following these two sections, different real world cases of the IoT are presented. These cases and their analysis ground the paper to what is currently happening in the world of the IoT. The aim is to pick cases that are representative of both common and well known use cases and lesser known more obscure cases to give a representative picture of the field. The paper focuses on two main cases, one where there is forced connectivity to outside networks to ensure added functionality of the IoT product, and one where there isn't. The forced/unforced debate seems the one that fits most clearly with the paper's hypothesis:

- IoT products that force the collection of personal data that is not used to provide functionality of that product are more unethical than those that don't
- IoT products that force connection to outside networks for either authentication or data collection purposes increase the risk to their users' data and therefore are more unethical than those that don't
- IOT products that collect personal data have a heightened ethical requirement that the data collected be integral to the functionality of the product
- IoT products whose functionality depends on the collection of personal data have a heightened ethical requirement that this functionality improve the user's quality of life substantially compared to similar non-IoT products

To better explain these hypotheses, a good example to look at is that of the smart fridge. Smart fridges work using sensors and a user profile through an app to order and restock depleting food automatically. From a technical point of view, the fridge is contacting its own manufacturer's servers to log user data through the app on the phone where the user can choose which items to automatically reorder, then contacting whichever online store the service uses. The ethical question here becomes, is the added ease of use or time saved from that functionality, worth the added need to trust the user data to an outside actor. Is it worth the added risk to in data being disseminated to outside networks? The alternative being that a smart fridge with no forced connection to outside networks can update its contents through an app that is limited to establishing a connection with the phone over LAN or encrypted GSM only, giving the user the information they need but requiring them to take the extra step to order or purchase the food themselves. Would the added functionality be worth the trade-off of having to trust someone else with acess to this personal data?

For the IoT products that this paper analyzes, the case studies involve a technical investigation into the setup of that product's IoT networks and systems, and any relevant technical knowledge that help follow how the user data in that product or system is tracked, stored and used. These case studies are followed by a discussion to look into the validity of the hypothesis, and help answer the research question. The focus of the discussion is how the

technical aspects of how these IoT products are designed impacts the way the data are or can be handled, and how that affects the ethics of the question, from the chosen framework.

Then, a follow up on this discussion with additional perspectives to the ethical investigation is undertaken to help in answering the research question further, such as looking at questions on topics of economics, personal and corporate responsibility and other areas of discussion. Finally the paper answers the research question and sub-questions.

2 Ethical Framework

As this paper is an ethics investigation, there needs to be a stance taken on what ethical framework to chose. The ethical perspective chosen is the one deemed most appropriate in analyzing the consequences of the mishandling of personal data and ethical considerations that need to be taken(Section 2.4).

2.1 Definition of "Ethical" in this Paper

Ethical can mean a lot of things, even within natural science. Within computer science, computer ethics as defined by James H. Moor a philosophy professor from Dartmouth College is "the analysis of the impact of nature and society on computer technology"[13]. The ethical view represented in computer ethics that this paper focuses on is consequentialism, as opposed to another view from computer ethics: Kant's deontological theory (also known as duty theories)[14]. This paper aims to take a broad consequentialist approach, looking through the lens of utilitarianism when necessary to aid in seeing the bigger picture. This paper aims to look at consequences of personal data mishandling for different actors, such as consumers, businesses, policy makers and so forth. These consequences and their ethical repercussions are not aligned amongst all these different groups, and therefore, a utilitarian lens can aid in tying together the overall ethical issues of data mishandling in the IoT.

Consequentialism is about determining what is right or wrong based on the consequences. An example would be that most people agree on lying is a bad thing, but if the lie was told in order to save a life, a consequentialist would argue that lying in that case was not wrong[15]. An important

precision to make here is that consequentialism concerns itself with real consequences, things that have happened and therefore have precedence. Consequentialism has difficulty making moral or ethical judgment on hypothetical events, though this framework can make determinations on probable potential events that have precedence.

Within consequentialism there lies the branch of utilitarianism. In essence, utilitarianism is about ensuring "the greatest good for the greatest number" in terms of consequences[16].

Deontology (or duty ethics) is an ethical theory that uses universal rules to determine whether something is right or wrong. This theory is often associated with philosopher Kant. These rules are set in stone such as "Do not lie" or "do not steal". It is often referred to as duty theory because it is about people just doing their duty and follow the universal rules for right and wrong. Unlike consequentialism where ethics is weighed on the outcome of an action, a duty theoretical perspective is focused on intention behind the action[17].

2.2 Consequentialist Approach

This paper takes a consequentialist view of ethical considerations. Consequentialism can be divided into two main parts: Act consequentialism, also called utilitarianism, and rule consequentialism. Act consequentialism states that for an action to be right the outcome must at least be as good as alternative outcomes. Rule consequentialism states that actions are right if they conform to a set of rules that if they are observed can reasonably be expected to create an outcome that is at least as good as alternative outcomes[15].

There are a few reasons for choosing this approach. The first and primary argument for a consequentialist ethical perspective is that it makes the task of analyzing the chosen cases simpler. What is meant by this is that a consequentialist approach does not care for intentions. This paper is not looking for less transparent and harder to research facts about the cases. Rather, a consequentialist approach helps look at what and how something is done, and infer ethical nature of the act based on the consequences or outcomes from decisions. These inferences are straightforward and supported in grounded thinking and evidence.

It makes more sense for this paper to analyze concrete and probable outcomes that have real impacts on consumers, businesses and society, and looking at what those impacts are – rather than taking a deontological approach (which attempts to analyze intentions; this type of research can lead to assumptions based on the rationale or intentions of different actors).

2.3 Ethics and Morals

While taking a consequentialist perspective it should be noted that it may be more difficult to cast moral judgments in a complete fashion. Morality tends to concern what someone or some entity ought to do. Morals do not necessarily change based on consequences. Similar to actions, decisions are either moral or immoral. Morality in this sense resembles ethics from a deontological perspective as morals do not immediately concern themselves with consequences. An action or decision can be classified as moral or immoral without the awareness of potential consequences. Once these consequences are known, the morality of the question can change but doesn't necessarily. Yet, consequentialism as an ethical theory puts the focus on consequences. When determining the ethical problems surrounding the IoT and data this paper concerns itself purely with ethics. This investigation aims to outline the ethical pitfalls that are consequences of certain aspects of IoT product design, in order to see how technology might best be used to help lessen these ethical problems.

2.4 Ethical Considerations of IoT and Data

The aim of this section is to outline the potential consequences of the mishandling of personal and private data. Outlining these consequences is integral for this paper as it is required for the consequentialist approach and completes the ethical framework. It is important to note here that these consequences differ for different actors. The way these consequences differ is outlined in this section. The way this impacts the ethical nature of these consequences and the ramifications of those ethical determinations on these actors is covered in the discussion. The actors, this paper focuses on, are the consumer (or the user), businesses, and governments.

In this paper, mishandling of data is defined either as breaches of regulations such as GDPR (General Data Protection Regulation), large scale data breaches from aggregate data hoarders, re-sellers, marketers or other

businesses, small scale breaches or hacks into homes or small-business networks, breaches of terms of service and privacy policies from the businesses, intentional or unintentional deceit of individuals pertaining to the handling of their personal data, or any combination of such events. According the ICO, the UK's independent authority on information rights, a data breach is "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed in connection with the provision of a public electronic communications service"[18].

Personal data are defined as any data users of technology create while using technology. This includes location data, browsing and search data, passwords, bank account details, workplace credentials, social media profile data, metadata from uploaded content, medical data and so on.

This paper considers personal data as a categorization to encompass not only the actual data that users create, but also any information that can be gained from this data through analysis. Browsing data can lead to information about people's consumer habits. Social media data can be analyzed in a similar way and can be used to infer political leanings and biases. Medical data can be used to infer information about medical conditions, mental health as well as genetic predispositions. Bank account details can lead to information about individuals' wealth and income, government credentials can lead to information and knowledge about one's nationality and national identity. Workplace credentials can lead to information about individuals' work history and so on. Terms of service are an important part of the use cases we are going to describe in Section 4. The reason this paper makes this consideration is the fact that any information obtained through analysis of personal data is dependent on access to that data. This data can be gained by the use of, for example, the terms of service of a website or product. These terms of service describe, among others, the privacy police of the company. The privacy policy is a statement which informs the user about how the company collects, stores and releases personal data. The issue with the Terms of services are, that if the user doesn't agree with them, the user wont be able to use the product/website. So, the user can either not use the product/website or has to agree on the way the company uses his/her personal data. As proprietary as the analysis aspect of personal data can be, as much as the analysis adds value to the data, this paper takes the position that information gained through such analysis is to be viewed the same way as raw personal data when looking at ethical considerations (see figure 2).

	Ethical considerations for users/consumers	Ethical considerations for businesses	Ethical considerations for regulators and government
Personal Data Privacy	- Does the user have the ability to control what data are collected, when and how? - Does the user know how their personal data are and can be used?	- What data and information are shared with third parties? - How is data privacy ensured? - Is there a good reason for the data being collected?	- What laws and regulations enhance consumer data privacy? - Do these laws do enough to solve ethical problems? - Are they enforceable?
Personal Data Security	- Does the user have agency in securing their personal data? - Is the user informed about the security risks of their personal data being handled?	- How is the security of the data ensured? - Are the risks of handling personal data and the work of securing it worth harvesting it?	- Who can be held accountable for data and security breaches? - Who can be held accountable for poor designs leading to poor security?
Personal Data Ownership	- Does the user have ownership of their personal data? - Does the user share in the benefits of the use of personal data?	- Where and for how long are data stored? - How are data disposed of and when? - Are user's data rights made clear to them?	- Is the transfer of data ownership regulated? - Is the creation and ownership of data regulated?
Transparency in handling personal data	- Is the user aware of what data is collected, when and how? - Is the user aware of security and privacy measures taken when their data is handled?	- Is the ratio of data privacy to the applications of data balanced? - Are privacy, security and transparency considered in this balance?	- Is transparency with regards to policies ensuring privacy and security of personal data enforced and regulated? - Are users' personal data rights made clear to them by law?
Applications for personal data	- Does the user have a say in how their data is used and analyzed - Does the user have the ability to correct incorrect information obtained through analysis of their data?	- What applications are possible with access to personal data? - Do the data have to be personal for these applications, or can they be anonymized?	- Can the data be monetized, by whom, and for what price? - Is the information gained through various applications derived from personal data regulated?

Figure 2: Table of Ethical Considerations of Personal Data, which gives an outline of the main ethical considerations compiled in this project

2.5 Ethical Consequences of the Mishandling of Personal Data

It is important when outlining ethical considerations for personal data usage and handling to look at the consequences that arise from ignoring or not taking ample enough care about these considerations. A question to pose when it relates to data privacy, is what are the dangers of data privacy being compromised? According to "Four ethical issues of the information age"[19] by Richard O. Mason, the author describes an ethical framework focusing on ethical issues for the information age, "Information privacy is generally understood as a right to seclude information about oneself. Assuring privacy means that one should have a right to determine whether, when, how and to whom one's personal information is to be revealed"[19]. This right to privacy is considered an ethical consideration by Mason. Although, today information and data privacy is widely viewed, especially among the younger generations, as a foregone luxury, the consequences of the lack of privacy can be dire. Some argue that information is power. In a 2019 TED talk from TEDSummit, Carole Cadwalladr, a prominent journalist who was featured in the 2019 Netflix Documentary "The Great Hack", makes this claim. In Carole's own words, "It's not about privacy or data, it's about power. A lot of young people think that privacy does not exist, and do not care about that. They have lived with this lack of privacy all their lives and do not feel there are real consequences. That is a misconception. It's who has the information, who has the data about you, that is where the power now lies. These companies that have that information are now the most powerful companies on earth"[20].

This paper takes the same view – that data privacy is important; mishandling data privacy offers more data to those interested in its analysis in the pursuit of centralizing information and therefore power. The consequences of this have been dire – the examples of "The Great Hack" are that of the analysis of large sets of personal data leading to knowledge of individual's political leanings, enabling political campaigns to target those that are most persuadable, breaking election and data privacy laws in the process[12]. Part of the problem today is that personal data and data in general is being hoarded for later analysis, as data sets get so big, they cannot be feasibly looked at with current techniques. If this paper is to then answer the question, what is private data used for? Other than in serving as a tool to enable product functionality, private data are used for many things today, such as targeted

advertisement, tailoring healthcare to individuals, figuring out individuals' political leanings, the list goes on. This list continues to grow, as emerging analytical techniques help uncover new types of information from huge datasets. From a consequentialist ethical perspective, it is possible to determine the consequences for current uses and analysis of personal data, such as knowing individuals' political leanings (something this paper's authors argue a majority of people agree has to be kept private to uphold the integrity of the political process and of elections). The ethical consequences of leaking personal data are therefore the same as leaking the potential information that can be gained from the analysis of that data.

For businesses, mishandling data can have consequences for the organizations' bottom line. When looking at data breaches, the consequences are financially driven. According to "2016 Cost of a Data Breach Study: Global Analysis", a benchmark research report sponsored by IBM and independently conducted by Ponemon Institute (an industry research center "dedicated to privacy, data protection and information security policy[21]), the average total cost of a data breach globally was 4 million USD in 2015). This study looked at 383 companies in 12 countries. In the US this average cost was 7 million USD and in Germany 5 million USD. Key points from the study: "48% of all breaches were caused by malicious or criminal attacks". By this the study's authors mean hackers as well as criminal insiders. "Incident response teams and extensive use of encryption reduced the cost of data breach. (...) The Loss of customers increased the cost. (...) The more records lost, the higher the cost (...) Time to identify and contain a data breach affects the cost." According to this report, the cost of data breaches is due to regulations enforcing the notification of such events to customers, the cost of losing customer trust or attempt to retain customers, the cost of compensating customers for financial losses incurred, the cost of investigating the breach, and the cost of improving security to prevent future breaches. Some examples of recent large-scale data-breaches, according to a CNBC article, are: Yahoo in 2013 and 2014, with 3.5 billion records stolen; First American Financial Corp. with 885 million; Facebook in 2019 with 540 million records stolen, Mariott in 2018 with 500 million records stolen, and Friend Finder in 2016 with 412.2 million records stolen[22]). According to webopedia, a record, when referring to the field of database management systems, is "a complete set of information (...) composed of fields, each of which contains one item of information"[23].

For businesses and the market ecosystems they exist in, data mishandling of an intentional nature can have positive consequences financially. Sharing and reselling customers' personal data with third parties leads to more revenue. When looking at IoT manufacturers, as well as with "free" online services such as search engines and social media sites, part of the revenue is often derived not from the service that is provided, but rather the data harvested from online activities of the consumers. This can be deduced from the fact that online services are required to maintain functionality of such products and services – these online services and infrastructure depend on constant running costs, yet not all of these function on a subscription basis. One must then ask the question, how is this financially feasible? The answer lies in reselling customer data to third parties for the use of targeted advertising and data hoarding for analysis.

For governments, the ethics of personal data lead to difficult questions and considerations for regulators, as outlined in table 2. A consequence of this difficulty is the lag between technology and regulation, creating a cycle of inadequate laws for the information age, which in turn leads to businesses not seeing impactful consequences for their actions, even when they do break laws. An example cited in Carole Cadwalladr's 2019 talk at TEDSummit[20] is that of Facebook; when the company was fined for it's involvement in the Cambridge Analytica scandal, to the tune of $5B, the stock price actually increased by $6B the same day, effectively rewarding the company for the fact that regulators would not affect its bottom line with fines[20].

In looking at consequential ethical considerations of handling personal data, different actors have different values which lead to different ethical consequences. This is further looked at in the discussion of this paper.

Cybersecurity ties in with these ethical considerations and consequences in a direct way. Indeed, cybersecurity helps protect data privacy in enabling the stronger prevention of data breaches when looking at aggregations of personal data and related leaks. Cybersecurity when related to networks within consumer homes and smaller businesses is important from this perspective as well, to a lesser extent – protecting the networks in such settings helps prevent smaller-scale targeted attacks. Greater cybersecurity measures lead to better data privacy and information protection.

Figure 3 shows a diagram of the ethical framework this paper works with. This framework is based on research and on the ethical considerations table, and aims to show how different actors are covered, what areas of considerations are taken, and what areas the discussion is focused on.

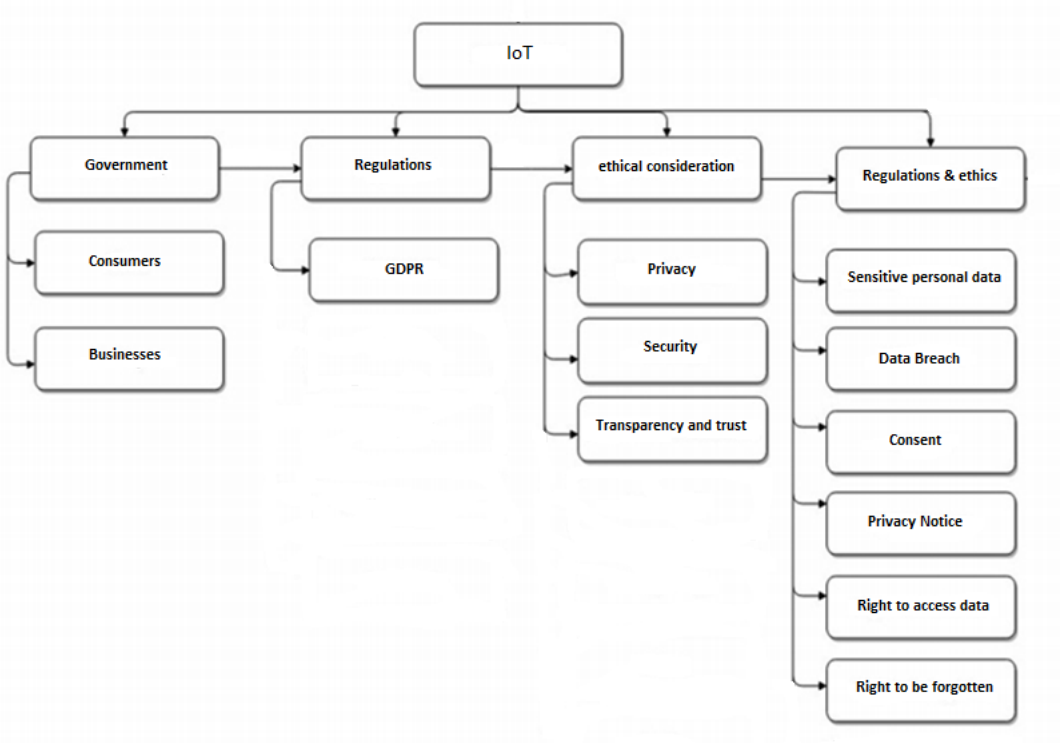

Figure 3: Diagram of the Ethical Framework

3 Technology

In this section the technologies that enable the trend known as the IoT are defined and described. The book "Securing the Internet of Things"[24] by Shancang Li and Li Da Xu is used as a technological reference for the later part of this section.

3.1 Evolution of the IoT

Different sources disagree on the first IoT appliances, but almost all of them agree on the origin of the term "Internet of Things" referring to devices connected over the internet.

The term "Internet of Things" was first used in 1999 by the co-founder of the Auto-ID Center at the Massachusetts Institute of Technology (MIT),

Kevin Ashton[2]. At the time Ashton was working at Procter and Gamble. Ashton termed IoT as a system in which objects in the physical world can be connected to the internet sensors[25]. Ashton put together the term in order to depict the benefits of connecting Radio-frequency Identification (RFID) tags that were used in corporate supply chains to the internet so that the the goods can be tracked without the need for human intervention.

One of the first examples of the IoT is The Trojan Room Coffee Pot. It was created by Quentin Staffor-Fraser and Paul Jardetzky in 1993. The reason behind creating this was to keep an eye on the coffee pot that was located in the computer laboratory while they were busy sitting and working at their workstations. The computer laboratory was equipped with a video camera. They aimed the camera at the coffee pot and connected it to the internet so the people would not have to make aimless trips to the coffee pot and wait for the coffee to be brewed. Instead, they can look at the images produced by the camera and see when the coffee was ready[25].

From when the term was first coined and roughly 17 years later there has been a huge advancement in the IoT. Since and beyond people have connected home devices, connected cars. People now have IoT in solar trackers as well as IoT manufactured plants.

Some of the key factors that has enabled the advancement and increased adoption of IoT are the following[26]:

- Advancement in Connectivity and network capabilities: Today there are numerous technologies that enable wireless connectivity such as Wi-Fi, Bluetooth, GSM.
- Improvement in cloud computing: Cloud computing enables data storage and data processing, that can always be accessed though internet connection.
- Invention of Data analytical tools and rapid improvement in data handling capabilities
- Reduced costs; for the actual devices as well as the low lost sensors. This also goes for the cost for cloud computing, wireless connectivity costs.

All these factors play a major role in the advancement and evolution of IoT[26].

3.2 What is IoT?

The internet of things (IoT) is a system of interrelated devices, objects or people that have the ability to transfer data over networks without requiring human-to-human or human-to-device interaction[7]. From a technical standpoint, a thing, as the name suggests, can be just about any thing. Among other things a device can be a machine, a sensor or a piece of infrastructure. The only defining factor for a thing in IoT is that the thing must have an identifying ID such that any transmission of information from that thing can be identified over a network, and ascribed to that thing. This identifier can be a unique ID such as an IP-address. But, on other types of networks the ID can be anything (such as on bluetooth constructed networks, or GSM networks, etc..).

Figure 4: Stylized depiction of a IoT network [27]

The other defining factor of IoT is that it describes a system, not a single object or device. For a device to be considered as an IoT device it needs to either be inherently part of a system (for example sensors for resource management sold as individual parts of a packaged IoT solution which comprises a controller on top of those sensors), or it needs to have the functionality to do so (a wireless webcam is an IoT device even if it is not sold as part of a "system" – the LAN network becomes a part of the IoT for that webcam, when connected). In this sense, any wireless enabled device can be a part of the IoT. Yet it is only an IoT product if that connectivity affects the added functionality of the product: smart fridges have the added functionality of being able to order food automatically – if such products would be unable to connect to networks, they would no longer become IoT products, as a regular fridge is not IoT. Another example to better illustrate this point is with smart watches: both a watch and a smart watch fulfill the watch functionality of displaying the time; the smartwatch has an added functionality of being able to connect to networks to complete other tasks such as messaging or calling.

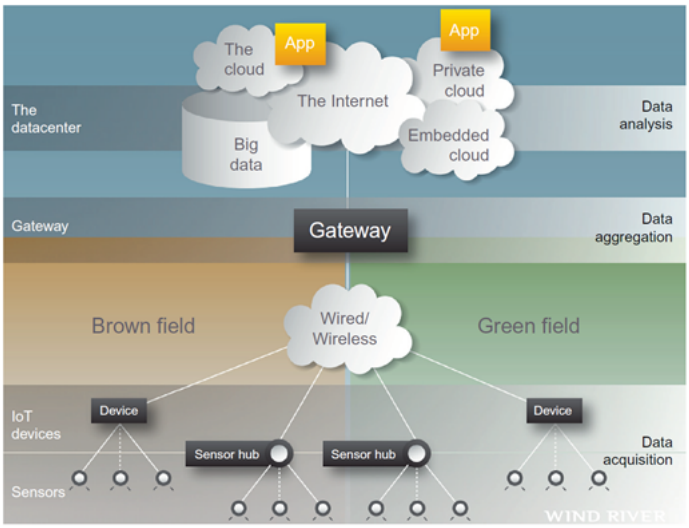

Figure 5: Diagram of a typical IoT setup[24].

3.3 How is Personal Data Collected / Stored / Used?

3.3.1 How is Data Collected

On an everyday basis users of the internet generate a lot of information as they surf the web. They are constantly being tracked on what kind of web pages they visit, how long they spend on those pages etc. In order to do so, websites use cookies, which is the most common and best-known technology to track users. A cookie is stored as a text file on a hard drive to store and transfer information to the server of the website[28]. Invented back in 1994, the purpose of cookies was to provide websites with a "memory" to, for example save items in an online shopping cart. Nowadays, cookies are still used to serve this purpose, but they are also used to monitor users, which can give detailed understanding of the users behaviour. It is common for websites to gather personal data about the users merely by asking their age, gender, income and geographic location by claiming to make the user experience better[29]. Websites use two different types of cookies, First-party cookies and third-party cookies[30].

First-Party cookies
These cookies are saved by the websites you are visiting directly. With these cookies, the website owners collect analytics data, remember language settings and more they need to provide a good user experience[30].

Third-Party cookies
Third-party cookies are cookies from other websites than the website the user is visiting. These can be websites which provide advertisements on the website the user is browsing on[31].

After the GDPR law was enforced in 2018, websites need the user to enable the cookies, by giving permission to use cookies as soon as the user enters the website[32]. Through this, they also indulge in sharing their data voluntarily. Personal data that is given voluntarily by users is transformed into a precious commodity, as mentioned in Section 1.3.

The user has also the possibility to disable cookies, so websites cannot save them on the users hard drive, but that might reduce the functionality of the website, for example no more advertising is displayed. Instead of not getting any data, websites have other tracking technologies such as tracking pixels and the digital fingerprint[29].

Digital Fingerprint
The digital fingerprint, also referred to as the browser fingerprint, is the uniqueness of a users computer, device or browser. A computer or device provides a website with information about its system and settings, every time a user visits a website. The fingerprint can consist of for example a user's particular configuration, location, time zone and language settings or browser plugin details. Individually, this information seems unnecessary or useless, but combined this fingerprint can stand out as one among millions of other, which makes this technique to 99% successful[29].

Javascript on websites
When building a website it is possible to add a javascript SDK, this SDK(software devolopment kit) makes it possible to link a website to a social media platform – like facebook. According to Mark Alan Richards facebooks javascript SDK can be illigal since it does not always uphold the GDPR. This is because these SDKs often give facebook access to a number of private data points such as reading the users user details and session cookies[33].

3.3.2 What Data are Collected and why ?

Location Data
One of the justifications for collecting location data is the necessity of the map application to know where the users is, to lead the individual to the desired location. Therefore the user agrees that his or her location can be tracked, so that the GPS and other tracking technologies can determine the devices geographical location.

Multiple systems and devices track their users location frequently. With agreeing to the terms of service of applications(apps) like Google, Facebook, Amazon etc, the users also agree that the apps are allowed to track the users location at any time. Besides this justification, there is a secondary market which has the purpose to make conclusions and predictions about the tracked individual. These predictions can be used for targeted advertising and marketing for retail and other business purposes[34].

Account Data
Account data are the data a website or application collects about the user, as soon as the individual creates an account. Facebook is a good example to clarify this. In the Facebook help centre everyone can see which information they collect and people can even download the information Facebook has for their account.

Facebook collects everything a user ever posted in its timeline, like relationships, work, education, etc. Facebook saves all the users movements on the website like which events the user went to, which people he or she is following, who likes the individuals posts/photos, which posts/photos the user likes. If someone purchases something on Facebook, they save their credit card number. All of that is saved as a part of a users account[35].

Browsing and Searching Data
Browsing and searching data are all data that are collected while one is on a web browser. If one is browsing in Google Chrome, they save information like what you search for,which videos you are looking at,what ads the user clicks on, one's location and websites you visit.

If an individual creates an account on Google Chrome, to for example use Gmail, Googles email service, Google starts collecting private information like names, birthday, gender, phone number, password, mails you write or calendar events.

Furthermore is it possible to browse privately on the internet, which is called the Incognito mode. While browsing, cookies and site data are remembered, but these data are deleted as soon as one leaves the Incognito mode. In the example of Google Chrome they delete the browsing history, one's cookies and data, information you entered in forms and permissions you give websites, but one's location might not be invisible to websites you visit, including ads[36].

IoT provides ways for human activities to be monitored in public places[37]. Google searches show detailed things about humanity: unending curiosity and inquisitiveness. They also show darker sides: unending bigotry, ignorance and depravity. Allowing a website or app to know one's location allows it to track one's movement throughout the day. There are cameras everywhere with the sole purpose of monitoring people's activities and behaviours, the best example for that might be China's surveillance state[38]. But, one does not have to be in public to be monitored. Several in-home appliances come with microphones and cameras to the spaces that were once considered private. Even if one does not have such in-home appliance it does not mean that parts of their daily routine and behaviours are not being observed and tracked. Their movements, behaviours and routines are still being observed through things like web-connected surveillance cameras, smart billboards and other public technologies. These seemingly harmless activities generate loads of data.

IoT systems are generally made up of of several components. The primary and most fundamental component is Wireless Sensor Networks (WSNs). WSN are primarily used for collecting data from the surroundings and then transferring them to central controllers so the generated data can be further processed[39]. Unlike the usual wireless sensor networks, the sensors in IoT are designed to be smarter. Along with obtaining information from the surrounding, they are equipped to make decisions with little to no human interaction.

Data in IoT is made up of data-sets that are generated by sensors. The data that the sensors gather from their environment is then analyzed and or combined with data from other sensors to help establish a pattern. Since the data collected in IoT does not originate from one single device rather it is a collection of data received through the different IoT devices, it must be processed before it can be used. The reason for this is that by processing the different formats from different devices can be turned into one uniform format. Another reason includes filtering out data that is unwanted or outdated for the sake of improving accuracy[40].

3.4 Opportunities and Risks of IoT and Personal Data

According to Bernard Mar from Forbes.com, people worldwide create about 2.5 quintillion (10^{30}) bytes of data every day, but with the expansion of the Internet of things, the pace of creating personal data is accelerating[41]. In 2018, 90% of the data in the world, was generated within the past two years[41]. Personal data are extremely valuable; as a whole and as a commodity, it has become more valuable than oil (if one were to extrapolate the price of a complete data set multiplied by the earth's population, the number is greater than the combined value of all of earth's oil reserves)[10]. But accordding to Liza Agrba from industryandbusiness.ca most people neither know how their data are being collected or sold, nor how valuable it is[42]. Companies use personal data in four main ways to make money: advertising, marketing, product development and data management[43].

Agrba states that With all the data companies collect, they build up a profile for every single person, so they can provide advertisement based on the data, which the users want to see[42]. Those ads include products or services that might make the lives better or easier. With the introduction of smartphones, hyper-localized advertising became a major opportunity for digital advertisers. The combination of social data and location data can be used to send advertisements offering discounts to the user in real time, which might lead them through the shops doors – according to "The impact of Big Data on the Digita Advertising Industry" on Qubole.com[44].

Johanna Rivard from Marketing Insider Group states that since the 1960s, marketing evolved from mass targeting, through direct mail and advertising, to direct marketing with the help of computer processing to target individual parts of the population through direct mail and telemarketing[45]. The old approach was to work off assumptions or instinct, but nowadays companies

can predict the needs, desires and future behaviours of their customer, based on their personal data – according to Eugen Knippels article "What is Data-Driven Marketing? The Definitive Guide" on adverity[46].

According to Andrijana Horvat et al. in their article "Understanding consumer data use in new product development and the product life cycle in European food firms – An empirical study", companies use personal data to improve their product development. The aim of this improvement is to better fulfill the consumer's needs and answer market demand for product functionality[47].

Knippel states that conventional data management systems need to handle the storage, retrieval and update of basic data items, records and files. The biggest advantage of a good data management is that it improves the business performance and provide data breaches or data privacy issues. In addition, can well-executed strategy in managing data give competitive benefits to business rivals?[46].

3.4.1 Opportunities for the end user

These risks may be worthwhile for the consumer, as examples of improved end-user experience thanks to IoT as a technological trend are endless. The first example, historically, is that of self-stocking vending machines(a coca cola machine). These machines can automatically know, thanks to sensors, when their stocks are low, and contact the drink supplier through a network in order to resupply[48]. This is the same concept as smart fridges. Home security systems that are IoT enabled create the added functionality of the system being able to automatically contact authorities in the case of a break-in or fire, as well as helping monitor the home from a distance[49]. Systems using bluetooth, RFID, wireless protocols and other shorter range network technologies can be used to keep track of items such as keys, devices, or anything else that is easily lost and can have an RFID or bluetooth chip fitted onto it. Smart irrigation systems and other types of resource management, sensor driven, IoT products, can help save and manage resources, and optimize their use in the most efficient ways, either letting the consumer directly make decisions, or basing such decisions off sensor information or weather forecasts. Vehicle to vehicle communication, another application of IoT, can theoretically, in self-driving cars, improve efficiency of traffic through coordination of vehicle movements and traffic as a whole[50]. Connected health IoT systems can help doctors and hospitals monitor consumer

health remotely and without being invasive, through sensor technology such as heartrate monitors, breathing monitors, activity monitors among other healthcare monitors. This can help personalize medicine and give doctors much more information and insight into patient health metrics[51]. Applications are also present in industries such as retail, for example in automatically tracking consumer trends and behavior in the store and changing the layouts and inventory automatically[52]. In farming, the applications resemble the smart irrigation system, but can go beyond, for example deciding which crops to plant when, based on weather and soil sensor metrics – even deciding harvest timing optimally based on the same data. Supply chains throughout the world can also be further streamlined, automatically sending orders much like smart fridges do, but on a much larger scale. These applications all rely on access to personal data from users[53].

3.4.2 Risks and Cybersecurity

All the new applications and home appliances that help interact with the lives run software, and no software can be "perfect". Code will always contain an expected rate of error, a bug rate. Every device or system is hackable, but good cybersecurity can be the difference of a hacker spending few days, months or even years to find the vulnerability in the code and the more code associated with an app or device, the bigger is the risk of cybersecurity threats. With the exponential growth of IoT devices and the amount of code used within these devices, the risk of cyberattacks grows as well[54]. An example is, when a man in 2018 hacked a baby monitor. The hacker can speak directly to the parents and even see what they did through the camera[55].

Another issue that arises, is that the more data are collected, even for the right reasons and to serve useful functionality of popular products, the amount of personal data can explode, and in the wrong hands this data can be extremely useful. A basic example is that of the thief who, based on someone scheduling heating, can infer when they are home or not. With the growing amount of data points, such inferences become much more precise and their predictions are much more accurate[56].

Google has recently patented a mirror that collects health metrics, and through AI it can tell if one is likely to have a heart attack or stroke the same day[57]. This information can be life saving but, in the wrong hands can be dangerous; this paper argues that the more personal data there exists about individuals, the more vulnerable people can be to its misuse in ways that can be unethical.

3.5 Technologies of the IoT

In this part different technologies central to the IoT are described. These technologies can be divided into different categories: network architecture, network protocols, encryption, and authentication. None of these technologies are specific to IoT – as IoT is a technological trend, it is a new way of using existing technology. Technical terms that are used in the analysis are defined here.

3.5.1 Network Architecture

LAN
LAN (or Local Area Network) is a small geographical network that is connected within a home, a school etc. Each device that is connected on this network is then able to access and share data with other devices on this LAN. Devices include; Computers, printers, scanners and data storage devices. Can also be wireless, referred to as WLAN. LAN and WLAN networks tend to be private and password protected[58].

VPN
VPNs (or Virtual Private Networks) are private networks accessible remotely. Like LAN networks they require authorization to be accessed[59].

3.5.2 Network Protocols

Network protocols define how data are transmitted over networks from one machine to the next. This section presents two types of wireless network technologies that differ from the most common wireless network protocols (the IEEE 802.11 family[60]). The two mentioned are the most widely used non-standard wireless protocols, but there are many more (such as ZigBee, Z-wave, Thread, NFC, RFID)[61].

TCP/IP, OSI, and IPv4-IPv6

An internet protocol (IP) is the primary protocol in the Internet Layer of the Internet Protocol Suite, which is a set of communications protocols consisting of four layers: the link layer, network layer, transport layer and application layer. The internet protocol suite (referred to as TCP/IP) is a suite of protocols designed to establish a network of networks to provide a host with access to the internet. TCP (Transmission Control Protocol) is a transport layer protocol, which creates a connection between two nodes on a network. IPv4 and IPv6 are link layer protocols, responsible for assigning IP addresses. Network layer protocols are found in internet protocols (IP) and are the backbone of Open Systems Interconnection Model (OSI Model), the model on which this 4 layer description is based[63],[64].

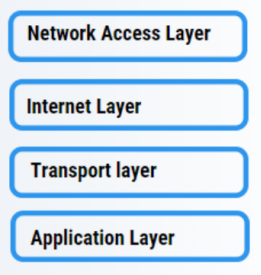

Figure 6: TCI/IP model[62].

HTTP

Hypertext Transfer Protocol (http) is an application layer protocol which uses a client-server protocol where the web browser is the client and communicates with the webserver that hosts the website. The browser uses HTTP, which is carried over TCP/IP to communicate to the server and retrieve Web content for the user. Http is the most widely used protocol on the internet due to its simplicity, but it lacks security[65].

HTTPS

Hypertext Transfer Protocol Secure(https) is a more secure version of http, what makes this more secure is the fact that it makes it possible for encrypted communication to occur. The version of http was made to secure sensitive data that would not have been safe otherwise, data such as credit card transactions, and user logins etc. These data are secured via either SSL (secure sockets layer) or TLS (transport layer security) encryption[66].

Bluetooth

Bluetooth is a encryption-free wireless network technology which works over small distances. Bluetooth networks tend to not have connectivity to the internet inherently[67].

GSM
GSM stands for Global System for Mobile Communication. It is a cellular network that can provide wireless communication using cells, primarily macro, micro, pico and femto cells where macro and micro are usually used outdoors and pico and femto are usually used indoors. If a phone has a SIM card then it is a GSM phone. GSM networks tend to have connectivity to the internet inherently, but this access can be modified by network admins[68].

DNS
Domain name system (DNS), is a naming system built to translate domain names to IP addresses using a distributed database. DNS makes it possible to assign domain names to groups of Internet resources and users[69].

3.5.3 Authentication

According to Shancang Li, "In IoT, authentication is the process of identifying users, devices, applications, and restricting access to authorized users and nonmanipulated devices or services"[24]. According to M. A. Ferrag et al. in "Authentication Protocols for Internet of Things: A Comprehensive Survey", Digital authentication is notoriously difficult to achieve[70]. It is quite easy to impersonate others electronically, a good example of this is phishing emails. Authentication is important for the IoT – such networks rely on the ability for the nodes and things on the network to have their identity authenticated. The way authentication is handled in the IoT products has a large impact on the security of those networks and therefore the ease of manipulating or intercepting data over such networks, if one were to circumvent or subvert the network's authentication protocols. Authentication can happen through authentication servers, or through cryptographic authentication protocols. In the IoT, it is important, especially for products that do not force connectivity to outside networks, for each thing on the network to be able to identify itself to every other thing on the network, as well as authenticate every other thing [70].

3.5.4 Encryption

There exist many different protocols and algorithms for encryption of data. Encryption is used to transform data into meaningless noise primarily to ensure that if it were to be intercepted by an unintended party, that they can not understand and read it. Encrypted data requires the right key to be decrypted and become legible and useful again. Encryption is relevant to the IoT in many different aspects: WPA and WPA2 are wireless network security standards that employ encryption. Certain authentication protocols rely on encryption algorithms, such as PKI (Public Key Infrastructure). HTTPS is a secure application networking protocol that relies on encryption for its security[71].

Encryption is particularly noteworthy to this paper and to the IoT due to the fact that authentication of things in the IoT relies on encryption schemes. Indeed, according to Shancang Li, in the process of authenticatio, "cryptographic schemes are used to provide a robust secure operation over the IoT"[24]. Yet, public-key based authentication (public key encryption, or PKE, being the most common form of encryption used on the internet according to Li[24]), is " for constrained environment such as IoT due to expensive cryptographic operations"[24]. What the author means by expensive here is in terms of computational resources. The IoT relies on lightweight nodes with varying degrees of computational abilities to all work together – expecting each of these nodes in the IoT to undertake these computations in a timely manner is not realistic. Therefore, there need to be more lightweight encryption schemes available to enable authentication in such computationally limited environments. Luckily, there exist lightweight cryptographic methods for PKE (based in complex mathematics involving elliptical curves [24]).

This paper does not go into detail about the mechanism of these cryptographic primitives, but deems it important to mention that this type of encryption makes it possible for authentication to take place within LAN networks and limited IoT setups, making the need for central authority authentication through external servers unnecessary.

4 Case Studies and Related Analysis

In this section, different real-world IoT products are observed, with one of the goals of identifying aspects of these products and their technology that are either "good" or "bad" in terms of how they treat consumers' personal data, how easy they are to hack, and whether or not their added functionality depends on outside network connectivity. Added functionality in this case is defined as the main purpose of the particular device. For instance the smart fridge's added functionality is not to keep one's food cold. The smart fridge's added "smart" functionality is having the ability to register what needs to be restocked and in most advanced case, to do so without human interaction. The analysis aims to set-up evidence for the discussion of ethical consequences which follows. This aim entails a systematic approach to analysis, though this is not always possible due to available research and documentation on the chosen cases.

4.1 Distinction Between Forced/Unforced in this Paper

The IoT products this paper analyses are divided into two categories: those with forced web connectivity, and those with unforced web connectivity. We refer to forced web connectivity as a forced use case, and similarly with unforced use cases and unforced web connectivity. Web connectivity refers to connecting to networks outside local networks. This means forced or unforced web connectivity, specifically connectivity to outside networks. Forced use cases, or products that force web connectivity tend to collect personal data by default. An IoT product is an unforced use case if the added functionality of the product does not rely on connectivity to the internet, or to outside networks. Outside networks are defined in this paper as networks external to the user controlled networks, such as LANs or VPNs. When an IoT product is required to connect to an outside server and the server collects personal data, it is deemed a forced use case in this report. An IoT product is unforced if it does not connect to outside networks by default. A product which retains its added functionality when disconnected from outside networks is not an unforced use case if it attempted that connectivity by default. For example, a camera with wireless upload functionality over WLAN is still a forced use case if it demands a connection to outside networks and servers to be made to access that functionality, through a proprietary application

perhaps. If the user is capable of restricting external network access and maintaining functionality, the product remains a forced use case in this paper's view, as that lack of outside connectivity was not inherent. A device that connects to Bluetooth, without demanding or requesting outside network connectivity, would on the other hand be an unforced use case. This question of forced/unforced is central to this paper due to the hypothesis that there are greater and more negative ethical consequences to an IoT that forces the collection of personal data.

The two types of forced IoT devices, this paper will focus on are a smart home security system, and smart irrigation systems. The unforced use case in this paper is the location device Tile. There were many different types of forced IoT devices that could have been chosen, and very few unforced. This paper only presents one case of unforced IoT device, because the few that were found were very similar. The reason for the less numerous unforced IoT devices, as it has been defined in this paper, is due to the fact that forcing outside network connectivity and collecting personal data is financially beneficial to the IoT manufacturers, as the private data collected through this forced connection can then be exploited and sold. The lack of the IoT products that do not do this that were found in this paper's research leads to the conclusion that such products , whose business models do not involve the collection and exploitation of personal data, are rare.

4.2 Forced Web-connectivity IoT Products: Cases

4.2.1 Ring Home Security

Product Description and Business Model

Ring is a home security system, which encapsulates many products, they have doorbells that let the user see if a person is approaching the front door, the user will then be notified if someone approaches the front door. Besides doorbells they also have a general security system, such as; Motion detectors, Contact sensors(notifies the user if windows or doors are opened in the house) and smoke and CO2 listeners – These are all setup through an internet connection, all of this is set up using an app. Ring also has an app called Neighbors App, this app lets the user share their surveillance videos with people in their neighborhood[72].

Figure 7: The Ring Alarm systems are built upon a number of devices such as Motion detectors, contact sensors and smoke and CO2 listeners[72]

Ring's business model relies on enabling connectivity to outside networks in order to contact emergency services for their customers paying a monthly subscription, in the event of a break-in. Seeing as this subscription offer does not make any hardware changes to the product, it is a software switch of sorts that allows the system to contact outside servers. The capability is there even without the subscription service[72].

Data collection

In the case of the home security system Ring, according to the terms of service given on their website Ring states that the content i.e video surveillance or similar forms of content is owned by the individual. However, they later state that they do have the right to view the content of its users "for the limited purposes of providing Services to you, protecting you, improving the Products and Services, developing new Products and Services"[72]. They also state that if one shares this content on for example their neighbor app then they give Ring complete consent to use that content as they see fit,

and therefore grant them "unlimited, irrevocable, fee free and royalty-free, perpetual, worldwide right to use, distribute, store, delete, translate, copy, modify, display, and create derivative works"[72]. Apart from this they also state that "Deleted Content and User Recordings may be stored by Ring in order to comply with certain legal obligations and are not retrievable without a valid court order"[72].

Privacy Policy

From Ring's own Terms of Service: "Equipment that relies on wireless or internet connections or are connected to a network of any kind (...) may not be secure and may be exploited or hacked by malware and spyware (...)[which] may provide a gateway for a person with malicious intent the capability to arm or disarm your system or related equipment; view, extract, change, destroy, steal, disclose or alter your data, or the data of others; monitor and/or spy on your activities and the activities of others; cause internet and network outages; provide for unintended or unauthorized access by others to your network, or the network of others; and otherwise place people, property or data at risk. RING MAKES NO WARRANTY OR REPRESENTATION THAT THE ALARM PRODUCTS, ALARM SERVICES OR MONITORING SERVICE IS SECURE, DOES NOT HAVE, OR IS NOT SUSCEPTIBLE TO, MALWARE VULNERABILITIES. Ring assumes no liability whatsoever for any Malware Vulnerabilities and, to the fullest extent permitted by applicable law, you agree to release and hold Ring harmless from any Malware Vulnerabilities and any related loss or damage of any kind or sort, even if caused by any breach of contract or negligence of any kind or degree of Ring"[72].

This does not instill confidence. Essentially this part of the Terms of Service releases Ring from liability from anything that may happen due to malware or spyware exploiting vulnerabilities in their program.

Cybersecurity

Why is this type of IoT product interesting to attackers? The answer for this is that as a result of attacking the home security system potential attackers can get private and intimate details about households and families. According to Harmon Leon from the observer, this includes getting their Wi-Fi passwords and gaining surveillance over their personal lives[54]. Attacks on such systems can cause a lot of damage. By getting a hold of the system, third party attackers can gain access to the whole network and to other devices that are connected by that network and can launch larger at-

tacks – according to A.J. Dellinger in the article "A Security Flaw Leaves Ring Doorbells and Cameras Vulnerable to Spying"[73]. One of the devices of the Ring home security system is the Ring video doorbell. Earlier this year, security researchers tried to hack the Ring video doorbell on stage at the Mobile World Congress. They were able to successfully hack the system which showed that audio and video transmissions can be exposed to third-parties[73]. The way this would work is that the attacker would hack the WiFi network of a household either by guessing the password or by using another smart home device that is connected to the same network[73]. The way the attack on the Ring can be carried out, the attacker needs to be on the same WiFi network as the device. Therefore, once the attacker hacks the WiFi, they can see the audio and video recordings the same way on the Ring app as the owner of the device. During the process of the audio or video recording being transferred to the app, the content is unencrypted which further makes it easier to intercept once the attacker has gained access[74] – according to Alfred Ng. "Ring doorbells had vulnerability leaking Wi-Fi login info, researchers find" from cnet.com.

How were these attacks performed by security researchers? Pen Test Partners is a limited liability partnership who excels in assessing devices, apps and more for potential vulnerabilities that their network security is exposed to. When they assessed Ring Video Doorbell they found a serious vulnerability that can easily allow the attackers to exploit the device and by extension, other devices that were on that network[75].

Internal or External Data Breaches

Before Amazons purchase of Ring in 2018, Ring has been criticized for poor security and abuse of the users' personal data. According to reports by the information, Ring has granted their Ukraine-based R & D (research and development) team full access to unencrypted video files, as well as live feed from customer cameras, regardless of this access being necessary for their work[76].

Figure 8: Ring[75]

In 2016, during a meeting with the R & D team in Ukraine and the founder of Ring Jamie Siminoff, an engineer requested access to private customer video feeds in order to improve the AI of their video cameras, and the founder agreed to it, in the hopes that this would speed up the capabilities of imagine recognition. This is according to employees present at the meeting or briefed shortly after[77].

When asked about it in an interview, Siminoff did not recall personally giving permission to this, but also argues against the concern that customers' personal data are in an increased state of vulnerability in Ukraine.

According to The Information, Ukraine is a hotbed of cybercrime, and to have Ring customer personal data shared in Ukraine is putting it at higher risk of falling into the wrong hands. Joshua Motta, a former CIA analyst who now runs the cybersecurity insurance provider Coalition, was interviewed by The Information and said the following about cybersecurity in Ukraine: "I would certainly place sending data to and from Ukraine as higher risk than operating elsewhere. Ukraine would be on a list of countries where I'd advise people to be more careful about what it is they share, and who it's shared with"[77].

A Ring Spokesperson confirmed in December 2018 that they now encrypt the videos, but does not speak further on when they began the encryption. She also added that it is only data that the customer has given consent to sharing, that is being shared to the R&D team in Ukraine.

A former employee said that this was not necessarily always the case. In 2016 the customer videos were widely shared to Ukraine. Ring's terms of service do not inform that customers videos are used for image recognition research and AI advancement. When asked about this, Siminoff answered that Ring's terms of service were sufficient[77].

Since Amazon acquired Ring in 2018, employees in Ukraine are no longer allowed to download and store customer videos on their computers. Ring has definitely added security measures since then[77].

When researching internal or external data breaches of Tile, Ring, GreenIQ and Rainmachine, there was not found any previous breaches of Tile, GreenIQ or Rainmachine.

If Tile or a third party of Tile had a potential data breach, it would not put the user at risk, because the data Tile shares with third parties, has been pseudonyminized. This means that any data is not identifiable with the actual user.

In terms of the internal data breach of Ring, there are some points to be mentioned. For Ring to have an Research and development team in Ukraine is from a user's perspective putting their personal data at risk. As mentioned, Ukraine is a place where significant amounts of cybercrime takes place. Had the Ring user known this information, this would likely compromise their trust in Ring. From the perspective of the business actor, it is financially beneficial for the company to have this department in Ukraine rather than in USA because of cheaper foreign labor.

The founder of Ring Siminoff, who gave the Ukraine research and development team access to the users personal data, argues that this was not in violation to Ring's Terms of Service. A well-aware user would not have been able to realize that by agreeing to the Terms of Service they allowed their videos to be used for image recognition research and AI development. This again compromises the trust of the customer towards the company.

Ring initially did not invest in encryption as a cybersecurity solution. A reason some companies don't invest in encryption and other cybersecurity solutions is that such initiatives are investments with no clear return on investment. When a developer mentions a need for spending on cybersecurity, an uninformed CFO (Chief Financial Officer) might dismiss this request based on the lack of direct returns. What this CFO would fail to realize is that such an investment is a preventative one - on the long term the costs of bad cybersecurity far potentially outweigh any savings from lack of investment. This turned out to be the case for Ring, who now encrypt their customer's data when they share them with third parties.

4.2.2 Smart Irrigation Systems

Product Description and Business Model

RainMachine, BlueSpray, GreenIQ are three companies offering smart irrigation IoT products. Each of these has a different approach to the product concept. The concept is using GSM and Wifi enabled sensors to automate irrigation and save money through increased efficiency, better resource management, automatic adaptation and monitoring of plants' consumption. Some of these companies offer the functionality of being able to automate irrigation based on weather forecasts, others take the approach of giving the user all the information and letting them make schedules for irrigation dynamically, instead of having the system do it automatically. These products provide very convenient user interface compared to traditional systems (thanks to smartphones, PCs, smart assistants). They are connected on one end to the waterline, and to valves on the other end which are connected to sprinklers. Access to dedicated cloud servers enable server-side authentication as well as the weather service functionality. Cloud servers also enable the C&C (Computer and Communication) functionality: the ability to interface with the system through a proprietary app or computer program[4].

Figure 9: rainmachine[78]

Figure 10: BlueSpray[79]

Figure 11: greenIQ[80]

These products are very cheap (in the 150-200USD range). Even with the added functionality of automating irrigation scheduling through access to weather forecasts, these products all require outside connection to enable functionality through C&C. This analysis can infer through this that personal data harvesting is a part of this product family's business model, as the type of data collected is very precise, localized, and plentiful, and can be useful for targeted advertising thanks to plant health data (for example advertising pesticides and fertilizers based on plant and soil sensor data).

Cybersecurity

Why is this type of IoT product interesting to attackers? The answer is that it is easy to attack, and that attacks can cause greater damage than most other cyberattacks, since these irrigation systems are directly in control of real infrastructure, and directly connected to critical infrastructure. Attacks on such systems can cause financial harm and resource wastage from the overconsumption of water[81] – according to Lorenzo Francheshi-Bicchierai from vice.com. Another reason this is an interesting target for attackers is the price of the system; as it is quite cheap and companies start spending more and more money on cyber security[8], attackers might assume that little to no investment was made into cybersecurity countermeasures to prevent attacks, and that little care was taken to protect networks operated to enable the functionality of such products.

How were these attacks performed by security researchers? According to IOT Village: DEFCON, a cybersecurity research conference, there were

two main techniques utilized: for GreenIQ and Rainmachine, the controller's firmware was directly extracted (GreenIQ had its firmware directly extracted from the raspberry pi controller that the product is built around, and RainMachine had it extracted from its controller through a UART to USB cable). The second technique employed by the security researchers was used for the third product. For BlueSpray the researchers identified, intercepted and captured network traffic and analyzed it using Wireshark, a piece of software used for network analysis. Through this analysis, the researchers inferred the mechanisms and software controlling the BlueSpray controller[82].

Having access to the controllers' firmware is the first step of the attack, as it allows Ben Nassi and the authors of "Piping Botnet – Turning Green Technology into a Water disaster"[82], to understand how to communicate with the device, and helps them in breaking authentication protocols, enabling them to interfere or intrude on the same network. This is indeed the researchers' next step: performing spoofing attacks (masquerading as another person/computer by falsifying data to gain illegitimate advantage or access). This is sometimes referred to as a "Man in the middle" or MITM attack. The goal of this attack is to change the input of the smart irrigation system to water the plants according to the attacker's wishes and not the system owner's wishes or the automatic decisions made by the system. The MITM attack takes advantage of the way the network interfaces between the IoT product's manufacturers' cloud servers and the user. For GreenIQ, the protocol used for this communication is HTTP; as outlined in an earlier section, HTTP does not have built in security, and is widely recognized as being very unsecure by network security experts today, to the point where the researchers, as well as the public at the security conference, found the mere presence of HTTP as a network protocol in this system to be amusing[4]. The MITM attack takes advantage of this vulnerability. Indeed, the way the network communication is supposed to work with an actual user is that a cloud server run by the manufacturer mediates a connection between the server and the user to update the watering plan, or update the user on sensor data. These sessions are initiated every minute. During this exchange, there is an HTTP request and response – the response that is requested is a timestamp authenticated with the device ID – when this response reaches the server it updates the watering plan after checking the timestamp and comparing it to the one on the request. In the MITM attack, the attackers spoof the system configuration with knowledge gained from extracting or extrapolating the firmware – this enables them to get a bot on the network that

intercepts network communication from the cloud server, and sends its own fake responses. In this sense, the MITM attack hijacks the request from the cloud server – it then sends a new update time that is newer than the time that is stored locally on the irrigation device. It then sends a fake XML file for a watering plan that consumes water all day using DNS spoofing through a fake DNS server.

This paper goes into detail into these attacks to show that they are not technically complex. It can be argued that these techniques have been around for a long time, are well understood and quite basic. This goes to show that these IoT products' manufacturers and designers have not invested much time and resources into cybersecurity of these systems.

Internal or External Data breaches
No data breaches have not been found within BlueSpray, rainmachine and greenIQ – that is not to say it is not possible or that small scale breaches have not happened within these manufacturer's servers or databases containing data that they have collected. The investigation into such breaches has not been fruitful as of the writing of this paper.

Data Privacy
A user's or consumer's privacy is based on how their private data is handled. Data is collected and stored due to its value. The main reason for the importance of keeping personal data private is the consequences if it is not, referring back to Section 2.5.

The cases in this paper have a varying level of data privacy described in their privacy policies. Every case looked at use third party services and partners. The Ring, Tile and Rainmachine cases do this with data analysis companies as is seen in the privacy policies.

Rainmachine
In their privacy policy rainmachine state what data they collect from their users. They collect: email addresses, first names and last names, phone numbers, addresses and cookies and usage data. Even though, they also share these data with third party companies, they state in their privacy policy that "third parties have access to your Personal Data only to perform these tasks on our behalf and are obligated not to disclose or use it for any other purpose"[83]. These "tasks" are only further explained as for analytic purposes, yet they do not state that this is the only thing they give to third party

companies. In their privacy policy, they mention some analytical companies they use i.e. Google Analytics, Firebase and AFS Analytics, however, they do that state that these companies are the only ones they share data with[83].

GreenIQ
GreenIQ do not clearly state in their terms of service what data they collect on you. They do, however, reserve the right to use any data created by the user, and to share it with third parties. The following is quoted from their Terms of Service: "When you use the GreenIQ Service, you create data ("Data") that is related to programming your Garden Computer. You hereby provide GreenIQ a limited non-exclusive, nontransferable license to use, upload, display, copy, manipulate and store Data solely in connection with providing the Service to you"[84].

GreenIQ do not mention anywhere what data is collected, what is shared with third parties nor does it mention who the third parties are. This is arguably the least transparent of all our use cases.

4.3 Unforced Web-connectivity IoT Products: Cases

4.3.1 Tile

Product Description and Business Model
An example of an unforced use case in IoT, as we defined it, is the Tile product. Tile is a device that is put on easy to misplace personal items such as: Keys, phone, bags, Tv remotes or anything else that one would have a tendency to lose. The device is then connected to the users phone through an app and if the item is lost the device is then trackable via a bluetooth connection. The tile casts a signal spanning a 100 foot radius, but that does not mean that the device has to be within a 100 foot radius, because the user would then be able to check on the app when it was last updated and where, if that is not enough then the user can activate the tile community which then broadens the search area by making every other phone on tile help finding the lost item. If the item is within the 100 foot radius, then the tile rings. If someone has lost their phone then they can use their

Figure 12: Tile[85]

tile devices to find that phone, it makes the phone ring no matter if it has been silenced. It is possible for a user to upgrade tile products to a premium subscription, this allows the user to be notified if the user leaves the product behind as well as location history and so on[86],[85].

Data collection
According to the terms of service and the privacy policy of Tile, Tile stores the location of the users phone. They explain that this is done in order to be able to find the desired items at a later point. This information is stored away from the users account data as to not associate these with each other, they call this process pseudonymization. According to the data security firm Protegrity, Pseudonymization substitutes the identity of the data subject in such a way that additional information is required to re-identify the data subject[87]. They also collect "your device's model, operating system type and version, and the dates and times of your requests"[88].

Further, from Tile's privacy policy: "How We Use De-Identified Information may share with third parties, including advertisers and service providers, anonymized, aggregated data we collect about you and other users, such as de-identified demographic information, de-identified Location Information, and information about the computer or device from which you access the Services, or the results of hashing your email address"[88]. In addition to that Tile uses and processes the personal information and location information of its users to create anonymous, statistical and aggregated data reports such that no individual user is identified[88]. This is puzzling, as it is hard to see why and how fully anonymized data can be useful to advertisers.

Related to cybersecurity: "We Take Security Seriously We implement various systems, applications and procedures to secure your Personal Information, in order to reduce the risks of theft, damage, loss of information, or unauthorized access, disclosure, modification or use of information. Please be aware, however, that these measures cannot absolutely guarantee the security of your Personal Information. Therefore, although we take great efforts to protect your Personal Information, we cannot guarantee and you cannot reasonably expect that the databases will be immune from any wrongdoings, malfunctions, unlawful interceptions or access, or other kinds of abuse and misuse"[88]. This liability statement is not as troubling as Ring's for example. This seems very standard.

Cybersecurity
Bluetooth is used to connect one device to another. In order to initiate the connection between two bluetooth devices, a piconet master is used to initiate the connection and the other are slave devices. Piconet is an ad-hoc network that connects devices using Bluetooth technology[89]. Once connected, the data between bluetooth devices is transmitted in the form of packets. Today, there are a few known methods for hacking into the bluetooth's security measures. A method called "bluebugging" is important to know when talking about a device connected to the phone via bluetooth like Tile. Using bluebugging allows the hacker to hack into the owner's phone. This can be dangerous in cases where people have their data saved along with the location of their belongings. Any hacker having access to this information can take control over the owner's messages, calls and the location of all the devices that the owner has connected to the Tile app[90].

According to the Mozilla foundation Tile has a bug bounty program, which means that if someone finds a security issue with tile, they would possibly get paid, if this information is disclosed to the company[91].

Data privacy
In the case of the Tile devices, they mention the ways in which the users' data are used and they specify what data they are going to use. They mainly focuses on two specific information types they share i.e. the user's locations and their devices' location[88]. Tile describes three different types of third party companies that they share their data with, namely:

- Web Analysis Service i.e Google Analytics, they do however state they make sure the third party company is not able to identify the user based on the shared information.
- Social plugins and Social widgets i.e Social networks, which the user is able to connect to their Tile app. They state that if the users do not wish for the social network to be able to gain their data, then they need not connect their social networks to the Tile app.
- Third-party payment processing i.e Apple app store, Google app store and Amazon app store. Tile states that these companies might collect financial information about the user.

Tile does not however state that these specific companies are the only ones used[88]. Nor does give its users the chance to opt out. The chance that if

the user does not wish for his/her data to be stored, the database will not store it. Although it is an unforced case because Tiles does not connect to internet, if the users do not wish for their data to be recorded they would have to give up the device. There is no other way to use the device while making sure nothing is being stored for long-term. So although they are being transparent with which data they use and why they use it, they are offering no alternative for someone who does not want his/her data to be tracked.

4.4 Comparing the use Cases

When comparing the three it is clear that GreenIQ are the ones that are the least transparent with their privacy policy. A user can not see what data is collected about them, nor which third parties they share the collected data with. Ring and Rainmachine have similar privacy policies in that they mention what types of information is shared, as well as what companies they share the data with for the vague term "analytic purposes".

These forced use cases clearly share portions of the user's data that the user might not be plainly aware of. From a consequentialists point of view, this has a very likely potential consequence that the user might decide to not purchase these items, and therefore, from the business actors perspective, they would suffer a decrease in sales.

For Tile, it seems apparent that they do not collect more data than they need. When looking at which third parties Tile shares user data with, if users choose to save their payment details within the app, these data are shared in order to process payments. Users are aware of this.

They do also mention that they share the location data for analytical purposes. This does not seem surprising either because the main data that is collected by Tile is location data. This unforced external network connectivity use case shows, that for this type of IoT device not a lot of data is collected, nor shared with third parties. From a consequentialist approach the average consumer would trust this company with less effort, keeping in mind that the abilities of Tile are also limited to only location services.

5 Discussion

The focus in this section is to look at the ethical considerations and consequences of mishandling personal data in the IoT trend in order to discuss the connection between these with different aspects of the IoT. These aspects are related to technology such as authentication, cybersecurity and data handling with regards to data privacy. Further relevant aspects for the discussion relate to considerations for businesses such as transparency and applications of using personal data. Other aspects relate to philosophical and regulatory questions such as personal data ownership and personal and corporate responsibility.

5.1 Authentication and Cybersecurity

Authentication is a central aspect of the IoT. It is important that every device on an IoT network is identified and legitimate. Authentication and network authorization - as is looked at earlier in this paper with the hacking case for the smart irrigation systems, unauthorized access to an IoT network can enable hackers full control of the functionality of the product, to the point of locking legitimate users out of such networks[24].

Authentication is a complicated task. It is simplified by allowing IoT products to access outside networks for the purpose of accessing an authentication server. Usually, these are decentralized databases with which IoT products can cross-reference network IDs such as IP addresses (using DNS servers), or other hardware identifiers. This outside connection can create other risks for the users[24]. As mentioned in Section 3.5.3, an outside connection is not always required to ensure device and node authentication over IoT networks. Authentication can take place locally on the network using lightweight encryption schemes, even with resource limited hardware. Encryption is required for authentication. For ensuring that the connected devices can be trusted to actually be what it is. Moreover, devices on IoT need to all be equipped with a unique identity that can be authenticated when the device attempts to connect to a central server or gateway[92]. This unique ID is how the system can track the devices and communicate securely with it. Looking at the example of Tile, it can be seen just how important authentication is. For the connected devices to be able to notified only when the owner wants it to, the device containing the app needs to be authenticated. If that was not the case, anyone would be able to track just about

every other device that is connected to Tile.

The consequences of forcing outside connections in the IoT networks is that user and consumer data are put more at risk, if the proper cybersecurity measures are not taken. As seen in the analysis of the smart irrigation systems, this can lead to real-world consequences in the form of resource wastage through loss of control of the system. Therefore, the ethical consequences on users can be negative if network security is not emphasized and properly implemented. This is due to the existence of lightweight encryption schemes which means that outside connections are not required for authentication in the IoT but are a simpler solution. Not using lightweight encryption and instead relying on central authentication authorities through forcing the IoT networks and products to connect to these outside servers leads to consequences for businesses as well, in that there then exists a heightened need for network security.

Lack of investment in implementing localized lightweight encryption in the IoT leads to larger financial requirements of maintaining outside central authentication authorities in server costs, as well as the cost of implementing good cybersecurity in these networks. If neither of these steps are taken, the ethical consequences for users are negative: their data is more at risk, and as outlined in Section 2.5 this can have negative ethical consequences for them.

5.2 Transparency

This paper takes the stance that it is important to mention transparency when it comes to data handling and cybersecurity issues in order to best make an ethical judgment on certain IoT products. Can transparency in the way the IoT manufacturers handle its users' data and protects their networks help with the ethical problems that are outlined? How openly does a manufacturer share its strategy and tools related to cybersecurity and data handling can have a great effect on the ethics of their products. This may seem like a deontological or duty ethics argument and it can be a straight forward one. However, it can also be argued that by knowing the intentions of the manufacturer, practical and technical details of personal data handling along with cybersecurity tools, efforts, problems and implementation can help not only assuage the users but also show that the manufacturer is taking these issues seriously. It can also help security researchers and personal data regulators better understand the problems and better work with the industry to help remedy such problems.

A company or firm providing its users complete transparency equips its users with full confidence because it allows the user to gain insight into the company, its devices, processes and it also demonstrates that the company is acting responsibly[93]. Users need to be equipped with all the necessary information so they can make informed decision before making connected products a part of their lives. Lack of transparency means that it is close to impossible for the users to know just what kind of data their connected products are capable of collecting. This makes it almost impossible for consumers to make an informed decision about which the IoT products they should invite into their lives. The users' lack of awareness of what data is collected on them, leads to less pressure on the companies to do the ethical thing. Therefore, the companies do not value the cybersecurity of the data as much. This then leads to the data being at greater risk for unwanted data breaches. This is, therefore, a consequence of lack of transparency.

From the point of view of the companies however, being transparent may seem as a downside. Being transparent for them means that the users will know if their data are being used and being so aware might make users reluctant in buying those products.

5.3 Benefits of Data Applications vs Personal Data Privacy

As mentioned in Section 3.3, data about users is collected at any time they are on the internet or use applications. Companies may argue that it is an advantage for the customer, since they get to see advertisement based on their personal data, but the companies have indeed the better advantages with personal data. Companies can use the data in so many more ways than consumers can. Analyzing data can help companies to make faster and better decisions about improving efficiency, effectiveness to fulfill customer expectations or the production of new products and services, as mentioned in the Ring's privacy police in Section 4.2.1.

The benefits of the applications of personal data clearly are not balanced between users and businesses, as seen in the example from the Netflix documentary mentioned in Section 2.5. Indeed, these ethical consequences can be negative for individual users and consumers. In that case, individuals' personal data were analyzed to better target people with political ads to convince them to vote in ways they otherwise would not. These applications

require large datasets comprised of personal data in order to target individuals. This also brings into question personal data ownership. As mentioned earlier in Section 2.5, this paper considers information gained from the analysis of personal data to be personal data itself – as this information can be used to target individuals in a way that would not be possible without the data. Due to this consideration, the applications for personal data that directly create such information without informing the concerned individuals in order to target them with marketing or political ads, are viewed as unethical from the point of view of individuals due to the negative consequences that would affect them. From a business standpoint, due to the lack of consequences of using personal data to add value and create new information, even if that information can be considered personal data itself, these applications do not create negative consequences for the business and are therefore not unethical. This may require a change in regulations to force consequences in order to force the consequences onto the businesses.

All of this seems necessary for companies, but is it ethical to collect someone's personal data in favor to strengthen a business, from a consequentialist point of view? Compared to how little advantages customers have and how much some business earn by selling personal data, it does not seem ethically right to sell others personal data. One may argue that customers can read the privacy policies before using websites or products, but the users get somehow forced to accept these policies, as some websites won't show all their content. The cookies the websites use can be modified by the user, so it only saves what the user wants it to save, but there is always some forced information the user has to disclose. The only way not to accept privacy policies with products like our use cases, is to not use them at all, which is not useful for the customer or the business.

5.4 Data Anonymization and Pseudonymization

As it is outlined earlier in this paper, ethical consequences of the IoT relate to the mishandling of personal data with respects to data collection (how it is done, the awareness or lack thereof of the individual who creates the data etc) as well as data storage (hoarding data in centralized locations with varying degrees of security measures). These ethical consequences come about due to the personal aspect of data and information collected. Without the personal identifiable nature of these data, none of the ethical consequences outlined in this paper come into play. It is therefore interesting to look at technologies

and techniques that remove personally identifiable information from collected data. In this section of the discussion, pseudonymization, a technique used by the Tile product as outlined in the Tile case study from Section 4.3.1, is looked at, as well as data. The workings of these technologies is discussed as well as the ethical consequences of implementing such technologies.

According to google, "Anonymization is a data processing technique that removes or modifies personally identifiable information; it results in anonymized data that cannot be associated with any one individual"[94]. Google strives to achieve data anonymization through generalizing data and adding noise to it. Generalizing data refers to removing a portion of the data or replacing it with a common value[94]. According to google, "we may use generalization to replace segments of all area codes or phone numbers with the same sequence of numbers." Google further explains what is achieved through this anonymization process: "Generalization allows us to achieve k-anonymity, an industry-standard term used to describe a technique for hiding the identity of individuals in a group of similar persons. In k-anonymity, the k is a number that represents the size of a group. If for any individual in the data set, there are at least k-1 individuals who have the same properties, then we have achieved k-anonymity for the data set. For example, imagine a certain data set where k equals 50 and the property is zip code. If we look at any person within that data set, we will always find 49 others with the same zip code. Therefore, we would not be able to identify any one person from just their zip code"[94].

According to google, k-anonymity is not enough – if one knows an individual is part of a data set which contains individuals who share the same sensitive attribute, information about these individuals may be revealed. Google goes on to state that to "mitigate this risk, we may leverage l-diversity, an industry-standard term used to describe some level of diversity in the sensitive values. For example, imagine a group of people searched for the same sensitive health topic (e.g. flu symptoms) all at the same time. If we look at this data set, we would not be able to tell who searched for the topic, thanks to k-anonymity. However, there may still be a privacy concern since everyone shares a sensitive attribute (i.e. the topic of the query). L-diversity means the anonymized data set would not only contain flu searches. Rather, it could include other searches alongside the flu searches to further protect user privacy." L-diversity can be leveraged to aid in anonymizing data sets by making sure that querying them cannot result in results that contain only one sensitive attribute.

Another technique used by Google is differential privacy, which aims to add noise to data: "Differential privacy (also an industry-standard term) describes a technique for adding mathematical noise to data. With differential privacy, it's difficult to ascertain whether any one individual is part of a data set because the output of a given algorithm will essentially appear the same, regardless of whether any one individual's information is included or omitted. For example, imagine we are measuring the overall trend in searches for flu across a geographic region. To achieve differential privacy, we add noise to the data set. This means we may add or subtract the number of people searching for flu in a given neighborhood, but doing so would not affect our measurement of the trend across the broader geographic region"[94].

Data anonymization has beneficial ethical consequences for individuals and can be a technique used by the IoT product manufacturers to further this goal if they are implemented in personal data storage. Indeed, according to google, anonymizing data can help enable detection of "security threats, like phishing and malware sites, all while protecting user identities"[94], as well as enable Google to "safely share anonymized data externally, making it useful for others without putting the privacy of our users at risk"[94].

It is important to note that anonymizing data can lead to that data being less useful in some applications. Applications that require generalized data, such as search autocompletion, are unaffected[94] – but different types of value added data analysis such as that used in targeted advertisements are rendered impossible by data anonymization (if data retailers do not know specifics about individuals and only generalized data through anonymized data sets, they cannot target individuals specifically). This is due to the fact that, according to the data security firm Protegrity, data anonymization "irreversibly destroys any way of identifying the data"[87]. The ethical consequences of this on businesses can be negative as it can affect their bottom line. According to Protegrity, "Pseudonymization substitutes the identity of the data subject in such a way that additional information is required to re-identify the data subject"[87]. A technique that leads to pseudonymization of data is one called "tokenization" which "provides a consistent token for each unique name and requires access to additional information"[87].

One of the drawbacks of this technique is if an attacker gets access to the database linking these tokens back to the original data, then that attacker has access to the data. A positive consequence for businesses of using this technique is that data can be more safely shared with third parties, and the usefulness of the data for analysis and further applications is not lost.

5.5 Comparing Forced and Unforced Web Connectivity

Products with forced and unforced web connectivity differ greatly in how they handle personal data. As defined in Section 4.1, forced refers to whether the IoT forces an internet connection and thereby the collection and dissemination of potentially personal information and data outside of the local and controlled part of the IoT. A forced use case in IoT products is more vulnerable to data mishandling, as seen in our analysis of the smart irrigation systems. A forced use case is open to more possibilities both in utilization and in vulnerabilities against cyber-attacks. These vulnerabilities are not as present in the unforced use case, due to data not being disseminated on networks not controlled by the user.

It is however also important to remember that there is a huge difference in what these devices are supposed to do. It is valid that an alarm system (Ring) or irrigation system need more data than a tracking app (Tile), but it also proves that the IoT devices do not necessarily have to involve so much data that users may start wondering if it is worth the security risk. In this sense, the functionality of the IoT determines not only whether or not the product may be of a forced or unforced aspect (as defined in this paper) but also the level of potential vulnerability due to data mishandling (the more data collected, even for legitimate purposes, the more valuable the dataset, and the more likely it will be targeted by attackers). The amount of investment in cybersecurity both from the network side and also from the data aggregation and storage side is therefore heightened – this can be considered a negative ethical consequence for businesses of the forced connectivity cases. For the user and consumer, the lack of investment in these fields leads to negative ethical consequences, as outlined in Section 2.5. Other investments in techniques such as data anonymization or pseudonymization also see heightened importance in forced use cases due to the nature of data being more likely to be mishandled in such cases. Due to the higher risk, getting rid of personal identifiable information in data reduces negative ethical consequences for the user.

Looking at the different business models of these types of use cases may also be enlightening. A question to pose here is that is it ethical for companies to "double dip"? Is it okay to charge the consumer for a product and sell their personal data in return? From a consequentialists' point of view, double dipping could be regarded as ethically correct. This is because the company

needs to be able to uphold the quality of the product because a lot of the IoT products need a working server from the company for the product to work at their full capability. However, that is not the case from the users' point of view, since this will mean a loss of data privacy, which as determined in Section 2.5 has negative ethical consequences for the user.

6 Conclusion

This paper outlines ethical considerations of data handling and mishandling in IoT and the consequences thereof, and then analyzes cases that illustrate the handling of these data in real-world contexts. It does so in a comprehensive way, exploring consequences for different actors, from users and consumers, to businesses to governments and regulators. In answering the research question this paper claims that handling personal data for the purpose of gaining an edge analytically or financially or in the pursuit of power over individuals is unethical when looking at the consequences on these individuals.

Based on the analysis, this paper concludes the extent to which the ethics of personal data handled with IoT may vary depending on the product and on the point of view taken (i.e. company's or user's). In some cases, the companies' actions might seem ethical from their point of view, but that this might not be the case if they look at it from the consumers' point of view. This paper looks and analyses the claims made by the companies in their terms of services and what it means for the users.

The conclusion here is that although the ethical consequences on businesses of the mishandling of personal data in the context of the IoT are less clear (as some of these consequences are positive and therefore not unethical), if one were to take a utilitarian perspective to analyse the ethics of the situation as a whole, then the consequences for individuals have the most weight, seeing as businesses are comprised of individuals. Employing a perspective from utilitarian ethics, a field of consequential ethics which concerns itself with making ethical and moral determinations based on the outcome for the greatest number of people, strengthens this paper's conclusion if it is to be adopted.

Answering the sub-questions posed in Section Section 1.3, it can be concluded that the forced/unforced distinction affects the ethical concerns of our research question in the way described in the earlier comparison of

forced/unforced cases. The conclusion drawn is that forced network connectivity in the IoT product is either outright less ethical than unforced, or if it is not, a heightened requirement for cybersecurity or anonymizing data exists. The second sub-question is answered in Section 2.4. The third sub-question is answered in Section 2.5. The discussion outlines two techniques in data anonymization and pseudonymization, which although they are not technologies, serve to answer the latest sub-question.

7 Perspective Discussion

7.1 Considerations on Data Ownership

Data ownership is a new approach to reduce the misuse of personal data. This approach is so recent that there is no common definition for data ownership. Ownership itself can be described as the fact that an owner legally posses a "thing", like for example a car or house and that this owner has the right to transfer this possession to others[95]. With this definition of ownership, data ownership can be described as having the legal right and complete control over a single piece or set of data elements, in the case of this project this set of data elements is the personal data of a person[96].

The most important and interesting question about data ownership is what the consequences for consumers, businesses and governments are. Today, companies earn money by selling the data they collect. First-party data is collected from websites or devices, which the customer directly interacts with, as mentioned in Section 3.3. This data comes directly from the customer, so companies know that it is of high quality and accurate. As soon as the companies collect personal data, they own it. The accuracy of the collected data is valuable to other companies as well, so they buy the data. This process now makes the purchased data to second-party data, because it is purchased directly from the company that owns it. The sources for second-party data are similar to those for first-party data, e.g. activity on websites or mobile app usage. Usually large data aggregators buy first-party data, because they know it is of high quality and the data aggregators can have control over what they buy and how the information gets used. All the data the aggregators collect, is gathered in large data sets and sold as third-party data, where after the data is used to expand companies audience or increase the precision of their targeting[97].

Within the last years, people got more and more aware of that companies have a lot of personal data on every individual and that the companies use it to there advantage, by selling the data and improve their advertising, as mentioned before in this section. So, after people allow companies to collect their data, they can't decide what happens with the data, neither how they use it for advertising nor to which third-party companies their data is sold. If consumer own their data, they could avoid the misuse of their data and decide them self how to use it. Consumer could decide if they want to sell their data to companies or even donate it for medical research. But before a private person has the possibility to manage their data, someone has to collect the data and these data collection companies won't work without payment. So, consumer might have to pay those firms or make some kind of a contract before even receiving their own data.

Companies might have more disadvantages, if users manage their own data, because the most companies either sell user data, or use them for advertising. So if they would not be able to sell the data anymore, they might lose a big source of their annual income. So, assuming that the companies even have to buy the data from the consumer, they even have to spend more money on advertising, which would make the loss of money even bigger compared to what it is, when companies manage the consumers data.

It is not only companies that collect and profit from user data, governments collect and use data as well. In today's world, governments do not trust their citizens, so they want to monitor them as much as they can or are allowed to. The best example for that, is Chinas surveillance state. China has about 200 million cameras that use facial recognition, to watch their citizens[38]. Governments might use data in a good way as well. An example for that are governmental medical research facilities, that might use data to find new treatments for diseases. Without the user data, or with an lower amount of data, these facilities might not be able to work efficiently on finding cures for diseases.

7.2 Regulations and Enforceability

In the EU the regulation of data protection is described in GDPR(General Data Protection Regulation). GDPR was made to protect peoples right to protection of their personal data[98]. This regulation encapsulates 99 articles. In article 5 GDPR states that it is only the necessary data that should be processed, this is also called data minimisation. This article and many more

are supposed to help protect the European peoples data[99], and more over this law is supposed to add a level of transparency for the consumer[100], since companies have to tell the consumer what information they process. That is however not what happens always. These regulations open up for cybersecurity risks of illegal activities like;"identity theft, cyberattacks, online espionage, theft of intellectual property, fraud, unlawful sale of drugs, human trafficking, and other criminal behavior"[101].

GDPR is enforced by a data protecting officer, that all companies must hire according to Section 4, article 38 in the regulation[98]. According to enforcementtracker.com 121 fines have been imposed since the regulations initiation back in May 2018[102].

7.3 Question of Personal vs Corporate Responsibility and Ethics (a Deontological Approach)

This is an interesting discussion to participate in, although it does not necessarily fit properly with an discussion on consequentialist ethics as it fits a more deontological ethical framework; therefore it is discussed in the perspectives discussion.

It is important to discuss the question of responsibility: if a user is able to affect the IoT technology, and through knowledge and technical expertise, limit the potential ethical problems of such technology without affecting functionality of the IoT products, is it the user's responsibility to seek that knowledge and use it? Or is it the company's responsibility either to educate their customers, educate regulators, or steer clear of ethical issues that might be solve able with the right knowledge?

Either argument is correct with different ethical frameworks, though a deontological position looks at duty ethics, and therefore the largest responsibility or duty arguably lies with the business. There is something to be said about personal responsibility, in the sense that any amount of knowledge the individual seeks in gaining technical expertise helps in achieving better control over ethical concerns and avoiding negative ethical consequences of mishandling data.

Yet, the average person cannot be expected to begin to pursue such knowledge if they are not interested in it, without first being made aware of why such knowledge and expertise is important. Because of this argument, a larger part of the responsibility, it is argued, falls on the businesses and corporations who handle data and have an obligation to have the technical expertise to stay competitive.

7.4 Limitations of this Paper's Approach

This paper's consequentialist approach has limitations, as critiqued in "Peril v. Promise: IoT and the Ethical Imaginaries", an article by Funda Ustek-Spilda, Alison Powell, Irina Shklovski and Sebastián Lehuedé. According to this article, "Technology reporting of IoT is characterized by a dichotomous imaginary of the future. On the one hand, it features anxieties about the consequences of pervasive connectivity, on the other hand, it envisions of a future where all technologies will be seamlessly connected to provide the most efficient and productive services. The unpredictability of how IoT technologies will evolve and how sociotechnical decision-making processes will change have led to a priori assumptions being made about the impossibility of identifying all ethical issues that might arise from IoT. As a consequence, in the literature, we see that discussions about ethics revolve around a limited focus on security and privacy. Moreover, these issues are understood as technical problems that are technical, so fixable, if only the 'appropriate' solutions are adopted. Issues pertaining to equity, equality and trustability arising from the adoption of IoT on the other hand, are vaguely categorized as 'social ethics' and the underlying ethical, social and economic issues are ignored, so are the situated contexts within which developers and designers of IoT technologies work"[103]. This article makes the informed claim that consequentialist ethics as related to IoT and technology exist in a world where there is a lack of understanding of how "local culture and network society influence the understanding and movement of particular social values among IoT developers, beyond the technical considerations of privacy and security"[103].

This article proposes a practical framework for ethics: "We propose to go beyond the consequentialist/utilitarian points of view, by bringing together three ethical frameworks that we think fit better with the problems at hand. These include virtue ethics, capabilities approach and care ethics. Virtue ethics focuses on individual's process of attempting to live a good life;

capabilities approach examines their ability to act, including to choose an alternative given the existing structural constraints and opportunities; and care ethics takes into account the shifting obligations and responsibilities of individuals as they are positioned in a web of sociotechnical networks." [103] The authors of this article claim that "Bringing these three approaches together enables us to acknowledge that ethics as a process is not exclusively dependent on subjectivities of individuals (e.g. their principles and actions), but acknowledges the situatedness of ideals and actions within structural conditions that can limit and shape them, and the demands and obligations that arise from these conditions."[103]

This framework, had it been researched earlier in the writing of this paper, may have been a more interesting one to use in approaching the topic of IoT and personal data.

References

[1] Kommunikation Internet Tingenes - Gratis billeder på Pixabay. URL https://pixabay.com/da/illustrations/kommunikation-internet-1927697/.

[2] Somayya Madakam, R. Ramaswamy, and Siddharth Tripathi. Internet of Things (IoT): A Literature Review (Date accessed 04-12-2019). *Journal of Computer and Communications*, 03(05):164–173, 2015. doi: 10.4236/jcc.2015.35021. URL http://www.scirp.org/journal/jcchttp://dx.doi.org/10.4236/jcc.2015.35021http://dx.doi.org/10.4236/jcc.2015.35021.

[3] Renée Lynn Midrack. What Is a Smart Refrigerator? (Date accessed 05-12-2019), 2019. URL https://www.lifewire.com/smart-refrigerator-4158327.

[4] Ben Nassi. IoT VILLAGE - Attacking Commercial Smart Irrigation Systems (Date accessed 29-11-2019), 2018. URL https://www.youtube.com/watch?v=M5S2-PdvGc4.

[5] Basel Solaiman and Eloi Bosse. *Information Fusion and Analytics for Big Data and IoT (Date accessed 04-12-2019)*. 2016. URL https://books.google.dk/books?hl=da&lr=&id=WaKPCwAAQBAJ&oi=fnd&pg=PR7&dq=information+overload+iot&ots=8ZOB3x85L5&sig=ngzb7hGyyQTLMRQn3_Lt8kfyDXU&redir_esc=y#v=onepage&q=information%20overload%20iot&f=false.

[6] ABOUT — heartrunner.com (Date accessed 04-12-2019). URL https://heartrunner.com/about/.

[7] Fritz Allhoff and Adam Henschke. The Internet of Things: Foundational ethical issues. *Internet of Things*, 2018. ISSN 25426605. doi: 10.1016/j.iot.2018.08.005.

[8] RSAC Contributer. The Future of Companies and Cybersecurity Spending (Date accessed 05-12-2019), 2019. URL https://www.rsaconference.com/industry-topics/blog/the-future-of-companies-and-cybersecurity-spending.

[9] Study Regulations for International Bachelor in Natural Sciences. Technical report, Roskilde University, 2019.

[10] Gary Allemann. Data Management – are you seeing the real value? — Big Data (Date accessed 04-12-2019), 2018. URL https://www.gigabitmagazine.com/big-data/data-management-are-you-seeing-real-value.

[11] Julia Angwin, Surya Mattu, and Terry Parris Jr. Facebook Doesn't Tell Users Everything It Really Knows About Them (Date accessed 29-11-2019), 2016. URL shorturl.at/ekvT4.

[12] Karim(Director) Amer and Jehane(Director Noujaim. The Great Hack, 2019.

[13] JAMES H. MOOR. WHAT IS COMPUTER ETHICS? (Date accessed 13-11-2019). *Metaphilosophy*, 16(4):266–275, 10 1985. doi: 10.1111/j.1467-9973.1985.tb00173.x. URL http://doi.wiley.com/10.1111/j.1467-9973.1985.tb00173.x.

[14] Kay G Schulze, Frances Grodzinsky, Frances " Grodzinsky, and Frances S Grodzinsky. Teaching Ethical Issues in Computer Science: What Worked and What Didn't (Date accessed 13-11-2019). Technical report, 1996. URL http://digitalcommons.sacredheart.edu/computersci_fac.

[15] Walter Sinnott-Armstrong. Consequentialism (Date accessed: 05-12-2019), 2019. URL https://plato.stanford.edu/entries/consequentialism/.

[16] Julia Driver. The History of Utilitarianism (Date accessed 04-12-2019), 2014. URL https://plato.stanford.edu/entries/utilitarianism-history/.

[17] Larry Alexander and Michael Moore. Deontological Ethics (Date accessed 05-12-2019), 2016. URL https://plato.stanford.edu/entries/ethics-deontological/.

[18] Security breaches (Date Accessed 09-12-2019). URL https://ico.org.uk/for-organisations/guide-to-pecr/communications-networks-and-services/security-breaches/.

[19] Richard O. Mason, Carr P. Collins, and Edwin L. Cox. Four Ethical Issues of the Information Age. Technical report.

[20] Carole Cadwalladr. It's not about privacy – it's about power — TED Talk, 2019. URL https://www.ted.com/talks/carole_cadwalladr_it_s_not_about_privacy_it_s_about_power.

[21] Why We Are Unique (Date Accessed 09-12-2019), . URL https://www.ponemon.org/about-ponemon.

[22] Kenneth Kiesnoski. 5 of the biggest data breaches ever (Date Accessed 09-12-2019). URL https://www.cnbc.com/2019/07/30/five-of-the-biggest-data-breaches-ever.html.

[23] Vangie Beal. What is Record? Webopedia Definition (Date Accessed 09-12-2019). URL https://www.webopedia.com/TERM/R/record.html.

[24] Shancang Li and Li Da Xu. *Securing the Internet of Things*. 2017. ISBN 9780128045053. doi: 10.1002/9781119187202.ch11.

[25] Quentin Stafford-Fraser. Trojan Room Coffee Pot Biography (Date accessed 28-11-2019). URL https://www.cl.cam.ac.uk/coffee/qsf/coffee.html.

[26] Story of Evolution of IoT — Adoption of internet of Things (Date Accessed 09-12-2019). URL shorturl.at/aqX23.

[27] What Do the Next Five Years Hold For the Internet of Things (IoT)? (Date Accessed 07-12-2019), . URL https://www.e-zigurat.com/innovation-school/blog/what-do-the-next-five-years-hold-for-the-iot/.

[28] WebWise Team. What are cookies? (Date accessed 06-12-2019), 2012. URL http://www.bbc.co.uk/webwise/guides/about-cookies.

[29] How do websites track users? — Technologies and methods (Date accessed 06-12-2019), . URL https://www.cookiebot.com/en/website-tracking/.

[30] Michal Wlosik and Michael Sweeney. What's the Difference Between First-Party and Third-Party Cookies? (Date accessed 06-12-2019). URL https://clearcode.cc/blog/difference-between-first-party-third-party-cookies/.

[31] Margaret Rouse. What is third-party cookie? (Date accessed 06-12-2019), 2014. URL https://whatis.techtarget.com/definition/third-party-cookie.

[32] Sean Keach. What is a cookie, should you accept them, and if you don't what happens? (Date accessed 06-12-2019). URL shorturl.at/mosHI.

[33] Mark Alan Richards. Facebook JavaScript SDK is often illegal (Date Accessed 14-12-2019). URL https://markssoftware.com/2018/06/23/facebook-javascript-sdk-is-often-illegal.html.

[34] Paige Boshell. The Power of Place: Geolocation Tracking and Privacy (Date accessed 09-12-2019), 2019. URL https://businesslawtoday.org/2019/03/power-place-geolocation-tracking-privacy/#_ftnref10.

[35] What categories of my Facebook data are available to me? (Date accessed 09-12-2019), 2019. URL https://www.facebook.com/help/930396167085762.

[36] Datagennemsigtighed (Date accessed 10-12-2019). URL https://safety.google/privacy/data/.

[37] Shaufikah Shukri, Latifah Munirah Kamarudin, Mohd Hafiz, and Fazalul Rahiman. Device-Free Localization for Human Activity Monitoring (Date accessed 05-12-2019). Technical report, 2018. URL www.intechopen.com.

[38] Sophie Perryer. Surveillance cameras have become one of China's most valuable exports – here's why (Date accessed 16-12-2019), 2019. URL shorturl.at/bJOY2.

[39] Nacer Khalil, Mohamed Riduan Abid, Driss Benhaddou, and Michael Gerndt. Wireless Sensor Network for Internet of Things. Technical report, 2016.

[40] Lukas Reinfurt, Uwe Breitenbücher, Michael Falkenthal, Frank Leymann, Andreas Riegg, and Daimler AG Uwe Breitenbücher. Internet of Things Patterns. *EuroPLoP'*, 16:21, 2016. doi: 10.1145/3011784.3011789. URL http://dx.doi.org/10.1145/3011784.3011789.

[41] Bernard Mar. How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read (Date accessed 25-11-2019), 2018. URL shorturl.at/sxCRT.

[42] Liza Agrba. Consumer Data is More Valuable Than Oil. Do You Know Your Digital Worth? (Date accessed 25-11-2019). URL shorturl.at/tEKLV.

[43] Luke Edwards. How do companies make money from your data? (Date accessed 26-11-2019). URL https://www.pocket-lint.com/apps/news/130366-how-do-companies-make-money-from-your-data.

[44] Qubole. The Impact of Big Data on the Digital Advertising Industry (Date accessed 26-11-2019), 2014. URL https://www.qubole.com/blog/big-data-advertising-case-study/https://www.qubole.com/blog/big-data-advertising-case-study/.

[45] Johanna Rivard. Why Your Website Needs To Be Responsive (Date accessed 27-11-2019). *Marketing Insider Group*, 2019. URL https://marketinginsidergroup.com/content-marketing/marketing-needs-data-driven/https://marketinginsidergroup.com/content-marketing/marketing-needs-data-driven/.

[46] Eugen Knippel. What is Data-Driven Marketing? The Definitive Guide (Date accessed 27-11-2019). URL https://www.adverity.com/data-driven-marketing/.

[47] Andrijana Horvat, Giulia Granato, Vincenzo Fogliano, and Pieternel A. Luning. Understanding consumer data use in new product development and the product life cycle in European food firms – An empirical study. *Food Quality and Preference*, 76:20–32, 9 2019. ISSN 09503293. doi: 10.1016/j.foodqual.2019.03.008.

[48] Keith D. Foote. A Brief History of the Internet of Things (Date accessed: 04-12-2019), 2016. URL https://www.dataversity.net/brief-history-internet-things/#.

[49] Vishal Sharma, Hrishikesh Thakur, Pankaj Gaud, Hrishikesh Yadav, and Mahavir Devmane. IoT Enabled Smart-Home. Technical Report 2, 2017.

[50] X. Krasniqi and E. Hajrizi. Use of IoT Technology to Drive the Automotive Industry from Connected to Full Autonomous Vehicles. *IFAC-PapersOnLine*, 2016. ISSN 24058963. doi: 10.1016/j.ifacol.2016.11.078.

[51] Alok Kulkarni and Sampada Sathe. Healthcare applications of the Internet of Things: A Review. Technical report. URL `www.ijcsit.com`.

[52] Jonathan Gregory. The Internet of Things: Revolutionizing the Retail Industry. Technical report, 2015.

[53] Andreas Kamilaris, Feng Gao, Francesc X Prenafeta-Boldú, and Muhammad Intizar Ali. *Agri-IoT: A Semantic Framework for Internet of Things-enabled Smart Farming Applications*. 2016. ISBN 9781509041305.

[54] Harmon Leon. Cybersecurity Expert Warns of Increasing Vulnerabilities in Devices (Date accessed 27-11-2019), 2019. URL `shorturl.at/io123`.

[55] Rebecca Joseph. Wi-Fi baby monitor hacked: Parents wake up to voice threatening to kidnap their child (Date accessed 27-11-2019), 2018. URL `https://globalnews.ca/news/4785542/wifi-baby-monitor-hacked-kidnap/`.

[56] Matthew Burrough and Jonathan Gill. Smart Thermostat Security: Turning up the Heat. Technical report.

[57] New Google Patent Could Turn Your Bathroom Mirror Into A Medical Device (Date accessed 04-12-2019), 2018. URL `https://www.cbinsights.com/research/google-patent-smart-home-medical-device/`.

[58] What is a Local Area Network (LAN)? - Definition from Techopedia (Date accessed 28-11-2019), . URL `https://www.techopedia.com/definition/5526/local-area-network-lan`.

[59] What is a Virtual Private Network (VPN)? - Definition from Techopedia (Date Accessed 06-12-2019), . URL https://www.techopedia.com/definition/4806/virtual-private-network-vpn.

[60] Bradley Mitchell. Wireless Standards: 802.11a, 802.11b/g/n and 802.11ac (Date accessed 04-12-2019), 2019. URL https://www.lifewire.com/wireless-standards-802-11a-802-11b-g-n-and-802-11ac-816553.

[61] Glenn Schatz. The Complete List Of Wireless IoT Network Protocols (Date accessed 29-11-2019), 2016. URL https://www.link-labs.com/blog/complete-list-iot-network-protocols.

[62] OSI Model vs TCP/IP Model — Top 7 Useful Differences To Learn (Date Accessed 06-12-2019). URL https://www.educba.com/osi-model-vs-tcp-ip-model/.

[63] What is TCP/IP? - Definition from Techopedia(Date Accessed 29-11-2019), . URL https://www.techopedia.com/definition/2460/transmission-control-protocolinternet-protocol-tcpip.

[64] What is Internet Protocol (IP)? - Definition from Techopedia (Date accessed 29-11-2019), . URL https://www.techopedia.com/definition/5366/internet-protocol-ip.

[65] What is Hypertext Transfer Protocol (HTTP)? - Definition from Techopedia (Date accessed 29-11-2019), . URL https://www.techopedia.com/definition/2336/hypertext-transfer-protocol-http.

[66] What is HTTPS? - Definition from Techopedia (Date accessed 28-11-2019), . URL https://www.techopedia.com/definition/5361/hypertext-transport-protocol-secure-https.

[67] What is Bluetooth? - Definition from Techopedia (Date accessed 28-11-2019), . URL https://www.techopedia.com/definition/26198/bluetooth.

[68] What is the Global System for Mobile Communications (GSM)? - Definition from Techopedia (Date accessed 28-11-2019), . URL https://www.techopedia.com/definition/5062/global-system-for-mobile-communications-gsm.

[69] What is the Domain Name System (DNS)? - Definition from Techopedia (Date accessed 29-11-2019), . URL https://www.techopedia.com/definition/24201/domain-name-system-dns.

[70] Mohamed Amine Ferrag, Leandros A. Maglaras, Helge Janicke, Jianmin Jiang, and Lei Shu. Authentication Protocols for Internet of Things: A Comprehensive Survey (Date accessed 29-11-2019), 2017. URL https://www.researchgate.net/publication/320078644_Authentication_Protocols_for_Internet_of_Things_A_Comprehensive_Survey.

[71] What is Encryption? - Definition from Techopedia(Date accessed 29-11-2019), . URL https://www.techopedia.com/definition/5507/encryption.

[72] Ring Terms of Service - Legal — Ring (Date accessed 28-11-2019). URL https://shop.ring.com/pages/terms.

[73] A. J. Dellinger. A Security Flaw Leaves Ring Doorbells and Cameras Vulnerable to Spying — Digital Trends, 2019. URL https://www.digitaltrends.com/home/ring-video-doorbell-security-flaw-hack/.

[74] Alfred Ng. Ring doorbells had vulnerability leaking Wi-Fi login info, researchers find (Date Accessed 06-12-2019). URL shorturl.at/mIQY5.

[75] David Logde. Steal your Wi-Fi key from your doorbell? IoT WTF! — Pen Test Partners. URL https://www.pentestpartners.com/security-blog/steal-your-wi-fi-key-from-your-doorbell-iot-wtf/.

[76] Dani Deahl. Ring let employees watch customer videos, claim reports (Date Accessed 10-12-2019), 2019. URL https://www.theverge.com/2019/1/10/18177305/ring-employees-unencrypted-customer-video-amazon.

[77] Matt Drange and Reed Albergotti. At Ring's R and D Team, Security Gaps and Rookie Engineers (Date Accessed 10-12-2019). Technical report, The Information, 2018. URL https://www.theinformation.com/articles/at-rings-r-d-team-security-gaps-and-rookie-engineers.

[78] RainMachine - Forecast Smart WiFi Irrigation Controllers (Date Accessed 07-12-2019). URL https://www.rainmachine.com/.

[79] BlueSpray - Web Based, Wireless (Wifi) Irrigation Controller (Date Accessed 07-12-2019). URL https://www.bluespray.net/.

[80] Vandingsstyring greenIQ Smart Garden Hub (Date Accessed 07-12-2019). URL shorturl.at/cqr17.

[81] Lorenzo Franceschi-Bicchierai. Hackers Could Cause Havoc By Pwning Internet-Connected Irrigation Systems (Date accessed 04-12-2019), 2018. URL https://www.vice.com/en_us/article/xwk5n7/hacking-internet-connected-irrigation-systems.

[82] Ben Nassi, Moshe Sror, Ido Lavi, Yair Meidan, Asaf Shabtai, and Yuval Elovici. Piping Botnet-Turning Green Technology into a Water Disaster. Technical report.

[83] Privacy Policy - RainMachine - Forecast Samrt Wi-Fi Irrigation Controllers (Date accessed 29-11-2019), . URL https://www.rainmachine.com/privacy/.

[84] GreenIQ -Terms-of-service.

[85] Tile - For Finding Anything - IoT - Internet of Things (Date accessed 29-11-2019), . URL https://iot.do/devices/tile-for-finding-anything.

[86] Learn How Tile's Bluetooth Tracking Device & Tracker App Helps You Find Your Lost Things — Tile (Date accessed 29-11-2019). URL https://www.thetileapp.com/en-us/how-it-works.

[87] Clyde Williamson. Pseudonymization vs. Anonymization and How They Help With GDPR (Date accessed 16-12-2019), 2017. URL https://www.protegrity.com/blog/pseudonymization-vs-anonymization-help-gdpr.

[88] Privacy Policy — Tile (Date accessed 29-11-2019), . URL https://www.thetileapp.com/en-us/privacy-policy.

[89] What is Piconet? - Definition from Techopedia (Date Accessed 09-12-2019), . URL https://www.techopedia.com/definition/5081/piconet.

[90] Vinayak P Musale and S S Apte. Security Risks in Bluetooth Devices. Technical Report 1, 2012. URL www.bluetooth.com,.

[91] Tile Mate (Date accessed 10-12-2019), . URL shorturl.at/xMRTZ.

[92] Why is Device Authentication Necessary for IoT? — Internet of Things Security — Thales eSecurity (Date Accessed 16-12-2019), . URL https://www.thalesesecurity.com/faq/internet-things-iot/why-device-authentication-necessary-iot.

[93] Dan Timpson. Transparency, responsibility and accountability in the age of IoT (Date Accessed 10-12-2019). URL shorturl.at/fkmtU.

[94] How Google anonymizes data – Privacy and Terms – Google (Date accessed 16-12-2019), . URL https://policies.google.com/technologies/anonymization?hl=en.

[95] What does ownership mean? (Date accessed 12-12-2019), . URL https://www.definitions.net/definition/ownership.

[96] What is Data Ownership? - Definition from Techopedia (Date accessed 12-12-2019), . URL https://www.techopedia.com/definition/29059/data-ownership.

[97] 1st Party Data, 2nd Party Data, and 3rd Party Data (Date accessed 12-12-2019), 2019. URL https://www.lotame.com/1st-party-2nd-party-3rd-party-data-what-does-it-all-mean/#what-is-3rd-party-data.

[98] General Data Protection Regulation (GDPR) – Official Legal Text (Date accessed 27-11-2019). URL https://gdpr-info.eu/.

[99] Andy Crabtree, Tom Lodge, James Colley, Chris Greenhalgh, Kevin Glover, Hamed Haddadi, Yousef Amar, Richard Mortier, Qi Li, John Moore, Liang Wang, Poonam Yadav, Jianxin Zhao, Anthony

Brown, Lachlan Urquhart, and Derek McAuley. Journal of Reliable Intelligent Environments Building accountability into the Internet of Things: the IoT Databox model (Date accessed 27-11-2019). doi: 10.1007/s40860-018-0054-5. URL https://doi.org/10.1007/s40860-018-0054-5.

[100] Andrew Denley, Mark Foulsham, and Brian Hitchen. *GDPR – How to Achieve and Maintain Compliance*. Routledge, 1 2019. doi: 10.4324/9780429449970.

[101] Roslyn Layton. Statement before the Senate Judiciary Committee On the General Data Protection Regulation and California Consumer Privacy Act: Opt-ins, Consumer Control, and the Impact on Competition and Innovation. Technical report, 2019.

[102] GDPR Enforcement Tracker - list of GDPR fines (Date accessed 02-12-2019). URL https://www.enforcementtracker.com/#.

[103] Irina Shklovski, Alison Powell, and Sebastián Lehuedé. *Peril v. Promise: IoT and the Ethical Imaginaries Funda Ustek-Spilda* Virt-EU: Values and Ethics in Innovation for Responsible Technology in Europe Department of Media and.* 2019. ISBN 9781450359719.